



Multidimensional diversity in video recommender systems: a holistic framework of literature gaps and future directions

İlknur Dönmez^a, Faruk Bulut^{b,c,*}

^a Istanbul Health and Technology University, Faculty of Engineering and Natural Sciences, Computer Engineering, Istanbul, Türkiye

^b School of Computer Science and Electronic Engineering, University of Essex, Colchester, England, UK

^c Department of Software Engineering, Istanbul Aydın University, Istanbul, Türkiye

ARTICLE INFO

Keywords:

Video recommendation systems (VRS)
Literature review
Diversity
Bias
Multi-objective optimization
Human-centered VRS

ABSTRACT

Video recommender systems shape content discovery, user engagement, and information access across digital platforms such as YouTube, Netflix, and TikTok. Although recent advances in deep learning, graph-based models, and multimodal recommendation have improved personalization, key challenges remain, including limited diversity metrics, weak geometric analysis of embedding spaces, inadequate multimodal datasets, and insufficient integration of fairness and transparency. This review synthesizes recent research on diversity-aware video recommender systems through a structured literature review. The screening process identified 572 records from six scholarly sources; after title–abstract screening and full-text eligibility assessment, 101 studies were included in the qualitative synthesis. The findings show that diversity is often treated as a narrow, one-dimensional objective, while broader system-level, user-centered, and societal perspectives remain underdeveloped. The review therefore argues for multidimensional and human-centered recommendation frameworks that incorporate geometric diversity modeling, multi-objective optimization, long-term user satisfaction, and user-controllable design. The study provides a conceptual basis for developing more robust, fair, transparent, and sustainable video recommender systems.

1. Introduction

Video recommender systems are central to digital platforms because they shape what users watch, how they access information, and which creators gain visibility. While these systems achieve high personalization accuracy and engagement, they also raise concerns about limited diversity, algorithmic bias, unfair exposure, echo chambers, and broader societal effects. This study examines these issues through the lens of multidimensional diversity.

Existing research has addressed diversity through classical reranking, graph-based models, deep learning, generative approaches, and optimization techniques. However, many studies still rely on narrow metrics, give limited attention to embedding-space geometry, use datasets that poorly reflect real-world short-video platforms, and treat ethical issues such as fairness and transparency as secondary concerns. These limitations suggest that future recommender systems should move beyond accuracy-driven ranking and support broader, fairer, and more sustainable information distribution.

Video recommender systems are widely used on platforms such as

YouTube, Netflix, and TikTok (Wei et al., 2023), and they also support education, online learning, and sports-related applications (Ma, 2025; Yang & Zhao, 2023). Given the amount of time users spend in digital environments (Clayborne et al., 2025; Frielingsdorf et al., 2025; Yang et al., 2025), these systems should be viewed not only as technical tools but also as mechanisms that influence social and informational structures (Niu et al., 2025; Tran et al., 2025).

Most recommender systems optimize engagement signals such as clicks, watch time, and repeated interaction (Afreen et al., 2025; Beynon et al., 2024). Although collaborative filtering, content-based filtering, and deep learning models improve personalization, they can also intensify familiarity bias, popularity bias, filter bubbles, misinformation exposure, and creator-level inequality (Gao et al., 2023). Addressing these challenges requires both technical and societal perspectives.

This paper makes four contributions. First, it organizes fragmented diversity-related concepts into a multidimensional framework covering popularity, content, producer, viewpoint, temporal, and semantic diversity. Second, it compares this framework with established concepts such as novelty, serendipity, fairness, exposure diversity, and echo

* Corresponding author.

E-mail addresses: ilknur.donmez@istun.edu.tr (İ. Dönmez), farukbulut@aydin.edu.tr, faruk.bulut@essex.ac.uk (F. Bulut).

<https://doi.org/10.1016/j.iswa.2026.200724>

Received 10 July 2026; Received in revised form 17 August 2026; Accepted 19 August 2026

Available online 20 August 2026

2667-3053/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

chamber mitigation. Third, it synthesizes recent work on reranking, debiasing, Large Language Model (LLM)-based reranking, deep and generative models, graph-based approaches, and multimodal systems. Finally, it identifies future directions in geometric diversity modeling, multi-objective optimization, long-term user satisfaction, and human-centered recommendation design.

2. Review methodology

This review followed a structured evidence-synthesis procedure inspired by PRISMA 2020 transparency principles. Because the included studies vary in datasets, metrics, architectures, and objectives, the study does not conduct a quantitative meta-analysis. Instead, it provides a qualitative and conceptual synthesis of methodological patterns, limitations, and future directions in diversity-aware video recommender systems.

The literature search covered six resources (ACM Digital Library, IEEE Xplore, Scopus, Web of Science, Google Scholar, and arXiv) for studies published between 2020 and early 2026. Older studies were included only when they introduced foundational concepts directly relevant to diversification or reranking. Reference lists of influential surveys and highly cited studies were also screened to identify additional work.

The search strategy combined three keyword groups using Boolean operators: recommendation domain terms (video recommendation, video recommender system, micro-video recommendation, short-video recommendation, and multimodal recommendation); diversity-related terms (diversity, novelty, serendipity, popularity bias, fairness, exposure fairness, echo chamber, filter bubble, and viewpoint diversity); and methodological terms (reranking, learning to rank, debiasing, graph neural network, embedding, multimodal fusion, transformer, diffusion model, generative recommendation, and large language model reranking). Search strings were adapted to each database while preserving these core concepts.

The final database search was conducted between 10 January 2026 and 25 January 2026. Duplicate records and clearly irrelevant records were removed before title–abstract screening through a combination of reference-manager duplicate detection, DOI/title matching, and manual verification. Two reviewers independently screened the records and coded the eligible studies; disagreements were resolved through discussion until consensus was reached. Complete database-specific search strings are provided in the supplementary material to support reproducibility.

Studies were included if they focused on video, short-video, or micro-video recommendation; evaluated a general recommendation method in a video-like or multimodal setting; or made a substantive contribution to diversity, fairness, bias mitigation, reranking, or exposure control. Studies were excluded if they addressed only advertising optimization, general information retrieval without recommendation, non-recommender video analysis, or purely accuracy-oriented recommendation with no link to diversity, fairness, bias, novelty, or exposure. Non-scholarly sources were used only for background context.

- **Recommendation context:** short-form video, long-form video, micro-video, multimodal recommendation, or general recommender-system context with clear relevance to video recommendation.
- **Model family:** classical reranking, debiasing/familiarity-bias mitigation, LLM-based reranking, deep or generative models, graph-based approaches, or multimodal recommendation.
- **Diversity construct:** the specific form of diversity, fairness, novelty, bias mitigation, exposure control, or echo-chamber reduction addressed by the study.
- **Evaluation basis:** datasets, modalities, metrics, baselines, and whether temporal dynamics or real-world platform characteristics were considered.

Table 1

Literature sources.

Stage	Count
Records identified through database searching (ACM, IEEE, Scopus, WoS, Google Scholar, arXiv)	572
Records removed before screening (duplicates, non-scholarly records, and clearly irrelevant records)	352
Records screened (title/abstract)	220
Records excluded at title/abstract screening (not recent or not aligned with the diversity perspective)	119
Full-text articles assessed for eligibility	101
Studies included in the qualitative synthesis	101

Table 2

Literature sources according to their categories.

Model family	Primary focus	Num	Reference IDs
Classical reranking approaches	Post processing diversity constraints (e.g., maximal marginal relevance style)	10	(Carraro & Bridge, 2024; Karo et al., 2024; Kelly & Bridge, 2006)
Learning to alleviate familiarity bias (LAFB) / debiasing	Bias correction at scoring/ranking stage; Diversity-aware recommendation; Exploration vs Exploitation; Debiasing techniques; Contrastive Learning	16	(Jin et al., 2025; Liu, 2023; Ren et al., 2026; Yang & Geng, 2026)
LLM based reranking	Prompted list refinement / controllable constraints	12	(Carraro & Bridge, 2024; Ding, 2025; Liu et al., 2026; Martina et al., 2025)
Deep learning and generative models	Representation learning; diffusion/generative diversity	12	(Gu et al., 2026; Yang & Zhao, 2023; Yang et al., 2025; Zhao et al., 2023)
Graph and graph network-based approaches	Echo chamber / polarization analysis and mitigation	15	(Liu et al., 2025; Niu et al., 2025; Rawka & Jadeja, 2022; Tran et al., 2025)
Multimodal video recommender systems	Joint use of text/image/audio/video signals	14	(Li et al., 2024; Liu et al., 2024; Ma, 2025; Ni et al., 2023; Wang et al., 2024; Yang & Zhao, 2023; Yang et al., 2025)

- **Ethical and societal considerations:** whether the study explicitly discussed fairness, transparency, privacy, safety, creator exposure, misinformation, or broader social impact.

Titles and abstracts were screened first, followed by full-text review of the remaining records. Each study was coded using the categories listed above. If a study contributed to multiple areas, it was assigned to its primary category, while secondary contributions were noted for synthesis.

Quality and relevance were assessed according to four criteria: clarity of the research objective, relevance to diversity-related constructs, adequacy of datasets and evaluation metrics, and explicit discussion of limitations or ethical considerations. Because the review synthesizes heterogeneous conceptual and technical studies, no single risk-of-bias score was calculated; methodological strengths and limitations were evaluated qualitatively.

The synthesis was thematic. Studies were grouped into six model families: classical reranking, familiarity-bias mitigation and debiasing, LLM-based reranking, deep learning and generative models, graph- and network-based approaches, and multimodal video recommender systems. They were then compared across diversity modeling, data representation, evaluation methodology, and ethical or societal framing. The recurring limitations informed the literature gaps in Section 5 and the

future research agenda in Section 6.

Table 1 summarizes the screening flow. In total, 572 records were identified through database searching across ACM Digital Library, IEEE Xplore, Scopus, Web of Science, Google Scholar, and arXiv. Before title–abstract screening, 352 records were removed because they were duplicates, non-scholarly sources, or clearly irrelevant to the review scope. The remaining 220 records were screened by title and abstract. At this stage, 119 records were excluded because they were not recent, not aligned with the diversity perspective, outside the recommender-system scope, or insufficiently connected to video, multimodal, or platform recommendation settings. The remaining 101 full-text articles were assessed for eligibility, and all 101 studies were included in the qualitative synthesis.

Table 2 provides representative examples for each category. The categories are not mutually exclusive, and a single study may contribute to more than one model family, such as multimodal deep learning or LLM-based and classical reranking.

3. Comparative framework for diversity dimensions in video recommender systems

Diversity in video recommendation is a multidimensional construct rather than a single list-level property. It concerns not only whether recommended videos differ from one another, but also whether recommendation systems distribute attention across content types, creators, viewpoints, popularity levels, semantic regions, and time. This section therefore defines the main diversity dimensions used in the review and positions them within a unified evaluation framework.

3.1. Operational dimensions

For user set U , video catalog V , and recommendation list $R_{u,k}$ generated for user u , diversity can be operationalized through complementary dimensions. Intra-list diversity measures dissimilarity among items in the same list; novelty captures the extent to which recommended items are unfamiliar or less popular; serendipity combines unexpectedness with relevance; calibration evaluates whether the recommendation list reflects the user's historical preference distribution; provider fairness concerns the distribution of exposure across creators; exposure fairness examines quality disparities across user groups; viewpoint diversity measures the breadth of stances or perspectives; and temporal diversity captures how recommendations evolve across sessions.

These dimensions are related but not interchangeable. Traditional diversity and ILD (Intra-List Diversity) focus on item dissimilarity; novelty focuses on user unfamiliarity; serendipity requires both unexpectedness and usefulness; fairness-oriented dimensions concern the allocation of exposure; and viewpoint and temporal diversity address broader informational and longitudinal effects. Multidimensional diversity integrates these perspectives into a common framework for evaluating whether a video recommender system is relevant, varied, fair, and sustainable over time.

Intra-list diversity measures the average dissimilarity among videos in a recommendation list. Higher values indicate broader content spread, but the result depends strongly on the underlying similarity measure and item representation.

$$ILD(L_u) = \frac{2}{|L_u|(|L_u| - 1)} \sum_{i \in L_u} \sum_{j \in L_u, j \neq i} (1 - \text{sim}(\phi, (i)\phi(j)))$$

In video recommendation, ILD is useful for detecting repetitive lists, but it does not guarantee meaningful semantic, viewpoint, or provider-level diversity.

Novelty captures the degree to which recommended videos are unfamiliar to the user or less dominant in the system. Popularity-based novelty is simple to compute, but it is only an approximation of

genuine user-level unfamiliarity.

$$\text{Novelty}(L_u) = \left(\frac{1}{|L_u|} \right) \sum_{i \in L_u} -\log_2 P(i)$$

where $P(i)$ is the marginal popularity of item i , typically estimated as the proportion of interactions received by i relative to all item interactions. Higher novelty values indicate that the recommendation list contains less popular and potentially less familiar items. Popularity-based novelty is computationally lightweight and widely used, though it is an imperfect substitute for genuine user-level unfamiliarity.

Serendipity refers to recommendations that are both unexpected and relevant. It is valuable for discovery-oriented experiences, but its evaluation often requires user feedback or reliable proxies for unexpectedness.

$$\text{Seren}(i, u) = \text{Rel}(i, u) \cdot 1[\text{sim}(\phi, (i)\bar{\phi}(H_u)) < \tau]$$

where $\bar{\phi}(H_u)$ is the centroid of the user's historical interaction embeddings and τ is a dissimilarity threshold. The binary indicator ensures that only items sufficiently distant from past behavior are counted, while the relevance term $\text{Rel}(i, u)$ filters out items that are merely random.

Calibration (Cal) evaluates whether the distribution of recommended categories is proportionally aligned with the user's observed preferences. It prevents diversity-oriented ranking from drifting too far from user interests.

$$\text{Cal}(L_u) = 1 - \text{KL}(p_u \parallel q_{L_u})$$

where p_u is the empirical category distribution derived from the user's interaction history H_u and q_{L_u} is the category distribution of L_u . Calibration is especially important when ILD is optimized aggressively, a highly diverse list may inadvertently overrepresent categories the user rarely engages with, reducing perceived quality.

Provider fairness (PF) concerns the distribution of exposure across content creators or providers. It is particularly important on video platforms where recommendation visibility directly affects creator reach and platform-level inequality.

$$\text{PF} = 1 - \text{Gini}(\{e_p\}_{p \in \mathcal{P}})$$

where $e_p = \sum_{u \in U} \sum_{i \in L_u} 1[i \in \mathcal{S}_p]$ is the total exposure accumulated by provider p . A Gini coefficient near zero indicates near-equal exposure distribution; values approaching one signal that exposure is concentrated among a small number of dominant providers.

Exposure fairness evaluates whether different user groups receive recommendation lists of comparable quality. Its use requires careful attention to privacy, group definitions, and the ethical handling of sensitive attributes.

$$\text{NDCG}(L_{G_a}) - \text{NDCG}(L_{G_b}) \leq \epsilon, \forall G_a, G_b \subseteq \mathcal{U}$$

The threshold ϵ defines acceptable disparity. Exposure fairness is closely related to broader algorithmic fairness literature, but its application to recommendation introduces unique challenges, particularly around the availability and appropriateness of protected attribute data.

Viewpoint diversity measures the range of stances or perspectives represented in a recommendation list. It is central to reducing filter bubbles and echo chambers in informational or opinion-driven video domains.

$$\text{VD}(L_u) = H(\{v_i\}_{i \in L_u})$$

where $v_i \in V$ is the viewpoint or stance label of item i and $H(\cdot)$ denotes Shannon entropy. Low viewpoint diversity is associated with filter bubble and echo chamber formation, phenomena of significant concern in the context of algorithmic content curation.

Temporal diversity captures whether recommendations change

Table 3

Comparative overview of diversity dimensions in video recommender systems.

Dimension	Primary Metrics	Representative Datasets	Relevant Model Families	Key Limitations
Intra-List Diversity	ILD, α -NDCG, ERR-IA	MovieLens-20 M, YouTube-8 M, KuaiRec	MMR, DPP, GNN-RS	Embedding-sensitive; coarse genre taxonomy may be insufficient for video
Novelty	MSI, Long-tail rate, EPC	Netflix Prize, ML-1 M, MIND	Popularity-debiased MF, Causal RS	May reduce short-term satisfaction; amplifies cold-start for niche content
Serendipity	EILD, Unexpectedness@K	MovieTweatings, TikTok dataset	Contextual Bandits, VAE-CF, RL-RS	Subjective ground truth; human annotation required at scale
Calibration	KL-Calibration, JS Divergence	MovieLens-20 M, Amazon Video	Post-hoc re-ranking, Constrained MF	Requires predefined taxonomy; sensitive to interaction sparsity
Provider Fairness	Gini of exposure, NDCE	KuaiRec, Amazon Product	Fairness-constrained ranking, ILP re-ranking	Conflicts with relevance; provider identity often unavailable in public data
Exposure Fairness	Δ -NDCG, Demographic Parity	ML-1 M (w/ demographics), Last.fm	Adversarial debiasing, Counterfactual RS	Protected attributes raise privacy concerns; group definitions are context-dependent
Viewpoint Diversity	Fragmentation index, RBS, Echo Chamber Score	AllSides, PolitiFact-Video	Knowledge-graph RS, Stance-aware NLP-RS	Viewpoint labeling requires domain expertise; mixed-stance content is difficult to classify
Temporal Diversity	LTC, Inter-session Jaccard, Repeat Rate	KuaiRand, Netflix Prize (temporal)	SASRec, BERT4Rec, RL-RS	Session boundary definitions vary; high turnover can undermine personalization stability

meaningfully across sessions. It complements list-level measures by evaluating whether users are exposed to a broader range of videos over time.

$$TD(I_u^t, I_u^{t+1}) = 1 - \frac{|I_u^t \cap I_u^{t+1}|}{|I_u^t \cup I_u^{t+1}|}$$

This Jaccard-based formulation captures ‘session-level’ list turnover. Complementary measures can track coverage across a broader temporal window, evaluating how much of the catalog a user is exposed to over an extended period.

3.2. Comparative framework

Table 3 consolidates these dimensions into a compact comparative framework. It links each dimension to common metrics, representative datasets, relevant model families, and practical limitations, thereby supporting consistent comparison across studies.

3.3. Relationships and trade-offs among dimensions

The dimensions are mutually dependent and often involve trade-offs. Increasing ILD may reduce calibration; promoting novelty can conflict with short-term relevance; provider fairness may require exposure constraints that alter ranking utility; viewpoint diversity can affect engagement; and temporal diversity must balance exploration with recommendation stability. These tensions indicate that diversity-aware recommendation should be treated as a multi-objective design problem rather than a single-metric optimization task.

These relationships can be jointly formalized as a constrained multi-objective optimization problem:

$$\max_{\pi} \sum_{u \in \mathcal{U}} \text{Rel}(\pi, (u) u) \text{ s.t. } \text{ILD}(\pi(u)) \geq \delta_1, \text{EFOACal}(\pi(u)) \leq \delta_2, \text{PF}(\pi) \geq \delta_3$$

where π is the recommendation policy and the threshold parameters define acceptable levels of diversity, fairness, and calibration. Because these thresholds depend on platform goals, user expectations, and societal context, no universal optimum exists.

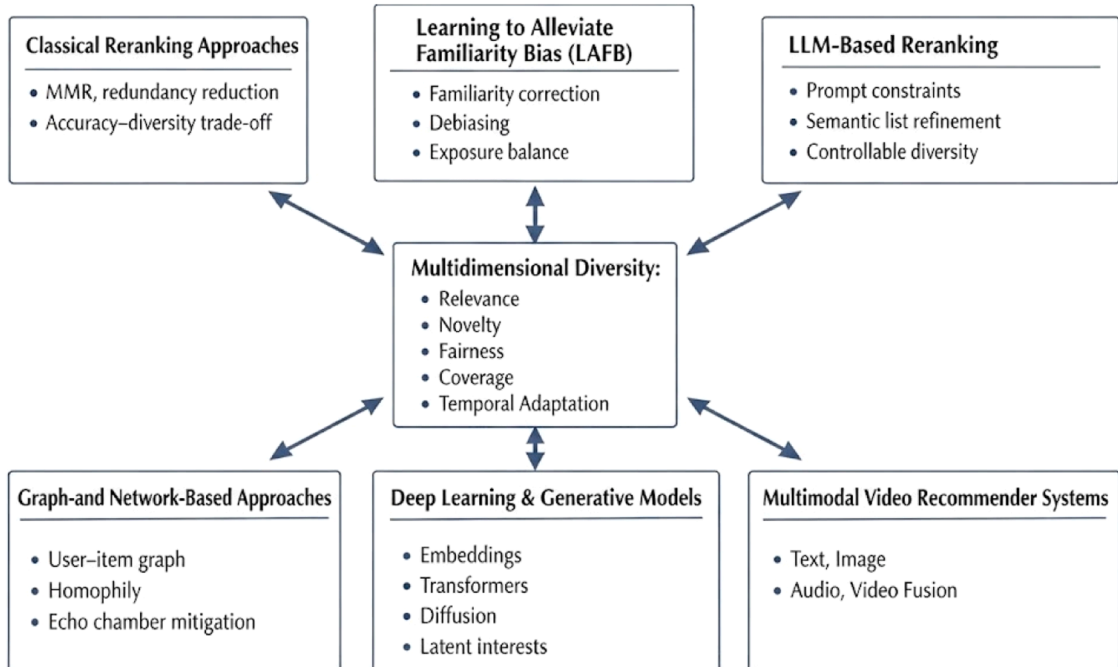


Fig. 1. Overview of current video recommendation models contributing to multidimensional diversity.

These trade-offs connect the framework to recurring challenges in recommender systems, including the accuracy-diversity trade-off, popularity bias, cold start, feedback loops, and echo chambers (Deldjoo & Di Noia, 2025; Hassani Saissi et al., 2025; Klimashevskaya et al., 2024; Ma et al., 2024; Yuan & Hernandez, 2023; Zhao et al., 2023). The framework therefore provides a structured basis for linking technical diversity metrics with broader platform-level concerns.

3.4. Limitations of the framework

The framework has three main limitations. First, most metrics are list-level, whereas users experience diversity longitudinally across repeated sessions. Second, many measures are sensitive to the embedding function or metadata used to represent videos, which limits comparability across studies. Third, several dimensions require provider identities, viewpoint labels, or group attributes that are rarely available in public datasets. These limitations show that the framework should be understood as an organizing structure rather than a complete measurement solution.

4. Current video recommendation models contributing to diversity

Current diversity-oriented video recommendation approaches can be grouped into six main categories: classical reranking, familiarity-bias mitigation, LLM-based reranking, deep and generative models, graph-based approaches, and multimodal systems. Fig. 1 presents this categorization in detail.

4.1. Classical reranking approaches

Reranking introduces diversity after a base recommender has generated candidate items (Carraro & Bridge, 2024). It reorders the list to reduce redundancy while preserving relevance, making it a practical post-processing strategy for existing collaborative filtering or deep learning pipelines. Maximal Marginal Relevance (MMR) is a widely used reranking method that balances relevance with dissimilarity to already selected items. By penalizing redundancy, it produces recommendation lists that remain relevant but less repetitive, as shown in the equation:

$$MMR(v) = \lambda \cdot Rel(v) - (1 - \lambda) \cdot \max_{v' \in S} Sim(v, v')$$

Here, $Rel(v)$ denotes the relevance of an item to the user, $Sim(v, v')$ represents the similarity between recommended items, S denotes the current recommendation set, and $\lambda \in [0, 1]$ is a parameter that controls the accuracy-diversity trade-off. Specifically, $\lambda \rightarrow 1$ prioritizes accuracy, whereas $\lambda \rightarrow 0$ prioritizes diversity. Beyond MMR, reranking methods reduce redundancy, balance exposure, and improve complementarity after relevance-based candidate generation. These methods are especially useful in short-video platforms, where similar and popular content can dominate top-ranked results (Li et al., 2024; Ren et al., 2026; Wang et al., 2024; Yang & Geng, 2026). Recent work further extends reranking through contextual distillation, conversational refinement, and coverage-oriented ranking for ambiguous queries (Liang et al., 2025; Li et al., 2024; Martina et al., 2025). Similar redundancy-reduction principles are also used in text summarization, showing the robustness of MMR-style diversification (Karo et al., 2024).

4.2. Learning to alleviate familiarity bias (LAFB)

Familiarity bias occurs when users are repeatedly exposed to already known content, reducing the visibility of novel or less popular items (Liu, 2023). Learning to Alleviate Familiarity Bias (LAFB) addresses this problem through a lightweight, model-agnostic correction mechanism applied at the post-ranking stage (Liao et al., 2022). The base recommender first produces a relevance score $s(u, v)$. A familiarity vector $f(u,$

$v)$, derived from historical interactions such as exposure frequency, subscriptions, watch time, and recency, then estimates how familiar the user is with the item.

This vector is transformed into a personalized debiasing factor through a learned function, expressed as:

$$d(u, v) = g(f(u, v))$$

Here, $g(\cdot)$ produces an item-specific correction coefficient that suppresses overly familiar content while preserving the visibility of less familiar or novel items. The correction factor is then combined with the base score to generate a calibrated final ranking. Its main advantage is that it improves diversity without changing the recommender architecture. By operating on output scores, LAFB maintains personalization while reducing repetitive content and increasing exposure to new or less visible creators.

Recent studies frame familiarity and popularity bias as ranking-stage problems shaped by repeated exposure, feedback imbalance, and biased sampling (Chen et al., 2023). Score calibration and post-ranking correction therefore offer practical tools for reducing overexposure while maintaining recommendation quality. Debiasing is also linked to exploration, fairness-aware recommendation, selection-bias mitigation, and contrastive learning, all of which support broader discovery and more equitable exposure (Dong et al., 2024; Li et al., 2025; Lin et al., 2024; Liu et al., 2024; Shang et al., 2024; Shi et al., 2023; Wu et al., 2024; Yang et al., 2023; Yang et al., 2025; Zhang et al., 2025).

4.3. LLM-Based reranking

LLM-based reranking improves recommendation lists through natural-language prompts rather than manually designed rules (Liu et al., 2026). Candidate items can be reordered according to objectives such as reducing repetition, increasing coverage, or avoiding similar consecutive items (Gao et al., 2025; Jin et al., 2025; Ren et al., 2025; Sharifmoghaddam et al., 2025). This approach is relevant for video recommendation because textual and multimodal descriptions can help LLMs distinguish items semantically (Ni et al., 2023). However, it also faces limitations related to computational cost, calibration, noisy inputs, and possible recency bias (Dehghankar & Asudeh, 2025; Fang et al., 2025; Xiong et al., 2025).

Unlike classical reranking, LLM-based reranking treats ranking as a language-based reasoning task. The candidate list and prompt are provided to the model, which generates a revised ranking that considers relevance and diversity (Ding, 2025). Video embeddings can further enrich this process by adding semantic content information (Zhao et al., 2023). Formally, a conventional recommender first generates a candidate set, which is converted into a textual representation and passed to the LLM with diversity constraints, as shown in the equations below. Here, C denotes the candidate item set, and Π represents the LLM-generated ranking:

$$\Pi = LLM(C, \text{Prompt})$$

$$\Pi \approx \arg \max (Relevance + \lambda \cdot Diversity)$$

The formulation indicates that LLM-based ranking does not explicitly optimize a well-defined objective function; rather, it implicitly approximates this objective through diversity constraints expressed in natural language within the prompt. The key distinction is that diversity is encoded through prompt instructions rather than explicit similarity penalties. This makes LLM-based reranking flexible and interpretable, but less directly optimized for ranking than methods such as MMR or greedy reranking.

4.4. Deep learning and generative models

Deep learning and generative models support diversity by learning complex user-item patterns in high-dimensional representation spaces

(Chaudhari et al., 2024). They can model multiple latent interests, temporal dynamics, and multimodal signals more effectively than traditional methods. Their main contribution is richer representation learning: user, item, and contextual embeddings capture broader preference patterns, which is especially valuable for video recommendation with behavioral, semantic, and multimodal signals (Martins et al., 2020; Wang & Zhao, 2022; Wu et al., 2022). Generative and diffusion-based models further broaden recommendation by modeling multiple possible user interests rather than a single deterministic outcome, improving coverage and reducing repetition while preserving relevance (Gu et al., 2026; Li et al., 2024; Liu, 2025; Liu & Zhang, 2025; Treuillier et al., 2026).

In Neural Collaborative Filtering, user and video embeddings are combined through a neural network to estimate a prediction score. This allows the model to capture nonlinear interaction patterns and multiple latent interests, indirectly supporting diversity. Transformer-based sequential models represent user behavior as a time series and use self-attention to capture dependencies across past interactions. This supports both short- and long-term interest modeling.

As a result, recommendations can reflect evolving interests rather than repeatedly reinforcing a single preference profile. Diffusion-based models such as DimeRec reconstruct meaningful representations from noise and can generate multiple plausible recommendation outcomes (Li et al., 2024). The forward noising process is shown as:

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \varepsilon, \\ \varepsilon \sim N(0, I)$$

Here, x_0 is the original data sample, x_t is the noisy state, and ε is Gaussian noise. By learning the reverse denoising process, diffusion models represent user preferences as a distribution over possible content directions, which naturally supports diversity. Overall, Neural Collaborative Filtering (NCF), transformer-based models, and diffusion models contribute to diversity by learning nonlinear interactions, temporal preference shifts, and distributed preference representations.

4.5. Graph- and network-based approaches

Graph-based approaches model users, content, and interactions as networks $G=(V,E)$. Nodes (V) represent users or items, while edges (E) capture interactions such as comments, replies, or co-engagement. Community detection and graph neural networks help identify echo chambers, polarization, and concentrated exposure patterns in video recommendation environments (Gao et al., 2023; Rawka & Jadeja, 2022).

These methods reveal community structures, repeated exposure, and restricted information flow (Avalle et al., 2024; Khanam et al., 2023; Maghsoudi et al., 2025). They are especially relevant for short-video platforms, where algorithms may reinforce homogeneous viewing behavior (Du et al., 2025; Gao et al., 2023). Graph-based mitigation strategies include heterogeneous graph modeling, graph contrastive learning, multi-view representation learning, graph reconstruction, V and sequential graph modeling (Amara et al., 2025; Han et al., 2022; He et al., 2023; Kim et al., 2024; Li et al., 2024; Liu et al., 2022; Liu et al., 2025).

Two common metrics are homophily and selective exposure. Homophily captures the tendency of similar users to connect or interact, as expressed:

$$P(\text{connection}|\text{similar attribute}) \gg P(\text{connection}|\text{different attribute})$$

When this condition holds, strong homophily indicates a higher likelihood of echo chamber formation. Selective exposure measures repeated exposure to content aligned with a user's existing views. Together, these metrics quantify community homogeneity and network-level polarization, making graph-based analysis useful for evaluating and mitigating echo chambers.

Graph-based video recommendation systems represent users, videos, creators, tags, comments, and co-engagement patterns as heterogeneous networks. After interaction and similarity signals are integrated into the graph, graph neural networks and message-passing methods learn low-dimensional node embeddings that preserve both structural and contextual information. These embeddings are then used to generate Top-K recommendations that balance relevance with diversity, exposure, and fairness constraints.

4.6. Multimodal video recommender systems

Multimodal video recommender systems move beyond user-item interactions by incorporating text, images, audio, and raw video. These signals provide richer content representations and help systems understand what a video is about, not only how popular it is.

Text captures explicit topics, visual frames provide scene and style information, audio captures speech or music, and video streams capture temporal content dynamics (Liu et al., 2024; Liu et al., 2024; Lubos et al., 2023; Malitesta et al., 2025; Mu & Wu, 2023). Multimodal fusion combines these signals into shared representations, improving personalization, robustness, and context-aware diversity (Lei et al., 2023; Pomo et al., 2025; Ye et al., 2025).

MicroLens is a representative large-scale dataset for this paradigm, containing approximately 9 million interactions, 1 million users, and 91,000 micro-videos, with title, thumbnail, audio, and raw video information (Ni et al., 2023). It enables models to learn from content-level signals rather than relying only on IDs or pre-extracted features.

Beyond accuracy and diversity, multimodal systems also raise trust and privacy considerations. Recent work therefore links multimodal recommendation with safer, more transparent, and ethically aligned system design (Wang et al., 2024).

5. Gaps in the literature

Despite progress in reranking, deep learning, graph-based, and multimodal approaches, video recommender systems still face important gaps in how diversity is measured, modeled, and aligned with ethical concerns. These gaps concern three main areas: diversity metrics, datasets, and ethical integration. Fig. 3 summarizes their relationship with future research directions.

Insufficiency of Diversity Metrics: Most studies measure diversity through simple list-level or category-based metrics. These measures are useful but insufficient for capturing producer, viewpoint, semantic, and temporal diversity. A more expressive approach is to model diversity continuously in the embedding space by computing the average pairwise distance between recommended items, as shown in the equation below:

$$\text{diversity}(S) = \frac{1}{N(N-1)} \sum_{i \neq j} d(v_i, v_j)$$

This treats diversity as a continuous property of the representation space rather than as a discrete category count. However, this formulation should be interpreted cautiously: large embedding distances do not necessarily imply meaningful topical, viewpoint-level, or fairness-related diversity.

Lack of Geometric Diversity Modeling: Embedding-based geometric approaches remain underused, even though modern recommender systems rely heavily on learned user, item, and multimodal representations. The spatial distribution of recommended items therefore provides a natural way to evaluate diversity. Without geometric analysis, recommendation lists may appear diverse by category while remaining concentrated in a narrow semantic region. Diversity should therefore be evaluated both at the list level and within the embedding space.

Data Limitations: Many studies still rely on movie-centric datasets such as MovieLens and Netflix, which do not reflect the multimodal,

Table 4

Future research directions.

Research Directions	Core Approach	Formalization / Metric	Key Idea
Geometric Diversity Modeling	Embedding space analysis of video representations	Pairwise distance, cluster entropy, manifold structure, and optimal transport	Diversity as spatial dispersion in representation space
Multi-Objective Optimization	Balance between accuracy, diversity, and fairness	Weighted objective over accuracy, diversity, fairness, and novelty	Balances competing recommendation objectives rather than optimizing accuracy alone
Long-Term User Satisfaction	Reinforcement learning and MDP-based sequential recommendation	Discounted long-term utility objective	Optimizes sustained user benefit rather than short-term clicks or watch time
Multimodal Diversity Modeling	Joint modeling of text, visual, audio, video, and interaction signals	Modality-specific embeddings and cross-modal fusion	Captures diversity across content modalities rather than relying only on user-item interactions
Human-Centered Design	Human-in-the-loop and user-controllable recommendation systems	User-adjustable control parameter α for the accuracy-diversity trade-off	Allows users to actively tune recommendation behavior and exploration level

fast-changing, and highly interactive nature of short-video platforms such as TikTok, YouTube Shorts, and Douyin. Future datasets should better capture text, visual, audio, video, temporal, and interaction signals. Without such datasets, proposed methods remain difficult to validate in realistic recommendation settings.

Ethical and Societal Issues: Ethical and societal concerns remain weakly integrated into engineering-oriented recommendation frameworks. Key issues include bias, over promotion of popular content, echo chambers, misinformation, privacy, and transparency. Thus, diversity should be treated not only as an optimization goal but also as part of a broader framework for fairness, safety, transparency, and privacy.

6. Future research directions

Accordingly, next-generation recommender systems should optimize multiple objectives simultaneously, including accuracy, diversity, fairness, and novelty, while also supporting user control and promoting the long-term health of the information ecosystem. This shift calls for recommendation frameworks that are multidimensional, geometry-aware, temporally adaptive, and human-centered.

Taken together, these research directions indicate a shift from recommendation systems that primarily answer, “What should the user watch?” toward multi-objective systems that also address “What information distribution is most beneficial for users and society?” Achieving this goal requires redefining recommendation as more than a user-item matching problem. Future systems should jointly optimize the geometric distribution of recommendations in embedding spaces, long-term user satisfaction, and interactive decision-making that aligns with user preferences while promoting a healthier and more balanced information ecosystem.

Table 4 summarizes these future research directions by presenting each direction, its core approach, its main formalization or metric, and its key conceptual contribution.

6.1. Geometric modeling of diversity

Future studies should move beyond categorical notions of diversity and investigate the geometry of recommendation spaces. Building on the average pairwise-distance formulation introduced in Section 5, geometric diversity can be further extended through cluster entropy, manifold coverage, and optimal transport. Instead, it explains how geometric analysis can be extended through complementary measures such as cluster entropy, manifold coverage, and optimal transport.

Here, the item embedding vectors are compared using a distance metric such as cosine or Euclidean distance. When item embeddings are normalized, cosine distance is commonly used to measure dissimilarity, whereas Euclidean distance is sensitive to embedding scale and therefore typically requires normalization or metric learning. However, average pairwise distance captures only one aspect of diversity. Future research should therefore consider richer geometric measures, including cluster entropy, manifold structure, and optimal transport.

Cluster entropy can measure how evenly recommended videos are

distributed across semantic clusters. Low entropy would indicate that recommendations are concentrated in a small number of clusters, whereas high entropy would suggest broader semantic coverage.

Manifold-based diversity can evaluate whether recommendations cover multiple regions of a latent semantic manifold rather than remaining confined to a narrow area of user interest.

Optimal transport-based diversity can compare the recommendation distribution with a desired target distribution and quantify the cost of transforming one distribution into the other. This may help identify whether the recommendation list is overly concentrated in particular semantic or popularity regions.

Therefore, geometric modeling enables diversity to be understood not only as the presence of different categories but also as balanced coverage of the representation space. This perspective can support more robust evaluation and optimization of recommendation lists. Diversity-aware recommendation should also be formulated as a multi-objective optimization problem:

$$\max_{\pi} \mathcal{L} = \lambda_1 R_{acc} + \lambda_2 R_{div} + \lambda_3 R_{fair} + \lambda_4 R_{novel}$$

where R_{acc} : accuracy (relevance / Click-Through Rate (CTR)); R_{div} : diversity; R_{fair} : fairness (e.g., content creator equity) and R_{novel} : novelty / exploration rate.

Potential future research directions include generating Pareto-optimal recommendation lists, developing dynamic weighting schemes that adapt objectives according to user and temporal context, designing gradient-based multi-objective learning-to-rank models, and integrating LLM-based reranking into multi-objective optimization frameworks. Combining methods such as LAFB, diffusion models, and LLM-based reranking within this formulation may lead to more balanced and robust recommendation systems.

6.2. Long-Term user satisfaction

Most existing recommender systems are optimized for short-term signals such as clicks and watch time. However, this optimization paradigm can reinforce echo chambers, filter bubbles, and excessive specialization. Future recommender systems should therefore prioritize long-term user knowledge development rather than immediate engagement alone. This problem can be formulated as a time-dependent optimization objective, as shown below:

$$\max_{\pi} \sum_{t=0}^T \gamma^t \cdot U(u_t, v_t)$$

where γ is the temporal discount factor, T is the time set, and U is the long-term user utility function.

Future research directions include modeling user behavior as a Markov decision process (MDP), applying reinforcement learning-based recommender systems, defining information diversity gain as an explicit reward signal, incorporating echo chamber mitigation directly into reinforcement learning reward functions, and optimizing the user learning curve as a long-term objective.

This perspective shifts recommender systems from click-optimization mechanisms toward information-ecosystem optimization systems.

6.3. Human-centered design

Future recommender systems should not rely solely on algorithmic optimization. Instead, human-in-the-loop and user-in-the-loop architectures that provide explicit user control will become increasingly important. In this context, the recommendation policy can be parameterized as shown below:

$$\pi^* = \arg \max_{\pi} (\alpha \cdot R_{acc} + (1 - \alpha) \cdot R_{div})$$

Here, α represents a user-controllable parameter that enables direct adjustment of the trade-off between accuracy and diversity. Like reinforcement learning paradigms, user feedback can continuously refine recommendation quality, diversity, and balance.

Potential research directions include diversity sliders, user-adjustable exploration modes, explainable recommendation interfaces, active learning from user feedback, and personalized diversity policies. These approaches would allow users not only to receive recommendations but also to shape the recommendation process itself.

7. Contributions of the study

This study provides a systematic review of diversity in video recommender systems from a unified and multidisciplinary perspective. Its main contributions can be summarized as follows:

First, the study conceptualizes diversity in video recommender systems as a multidimensional construct that includes content, producer, viewpoint, semantic, temporal, and popularity-related dimensions. This provides a broader framework than traditional list-level or category-based diversity measures.

Second, it synthesizes recent research on classical reranking, debiasing methods, LLM-based reranking, deep learning, generative models, graph-based approaches, and multimodal systems, highlighting how each contributes to diversity from a different methodological angle.

Third, the review identifies key gaps in existing studies, particularly the lack of geometric diversity modeling, limitations of available datasets, and insufficient integration of ethical and societal issues.

Based on these gaps, it proposes future research directions centered on geometric diversity, multi-objective optimization, long-term user satisfaction, and human-centered recommender-system design.

8. Discussion

Diversity in video recommendation is multi-dimensional rather than monolithic. The comparative framework introduced in this work clarifies eight operationally distinct dimensions—Intra-List Diversity, Novelty, Serendipity, Calibration, Provider Fairness, Exposure Fairness, Viewpoint Diversity, and Temporal Diversity, and situates each within a coherent analytical structure. By supplying formal definitions, primary metrics, representative datasets, suitable model families, and explicit limitations, the framework moves beyond the terminological conflation that has long characterised the literature.

Several practical insights follow. First, optimising a single dimension in isolation is rarely sufficient and often counterproductive. Aggressive maximisation of Intra-List Diversity, for example, can distort category proportions and thereby degrade Calibration; popularity-driven Novelty improvements tend to concentrate exposure among dominant providers, harming Provider Fairness. These tensions are not merely empirical observations but structural trade-offs that can be expressed as a constrained multi-objective problem. The appropriate balance among objectives is inherently platform-specific and must be negotiated according to the system's goals and values rather than treated as a universal

optimisation target.

Second, metric choice remains highly sensitive to representation. Nearly all quantitative measures depend on the underlying item embedding function ϕ . Changes in video representation can alter reported diversity scores without any corresponding change in user experience, rendering cross-study comparisons unreliable unless embedding details are carefully controlled and reported.

Third, evaluation remains incomplete. Most available metrics operate at the list level, yet users experience diversity longitudinally across sessions and over time. Temporal Diversity partially addresses this gap, but richer longitudinal measures are still needed. Likewise, dimensions such as Viewpoint Diversity and Provider Fairness require structured metadata (stance labels, provider identities) that are scarce in public benchmarks, creating a systematic mismatch between societally important criteria and those that can be readily evaluated.

From a methodological perspective, the reviewed approaches differ in their practical strengths and limitations. Reranking methods are practical and model-agnostic but can be limited by shallow similarity metrics; debiasing methods directly address popularity and familiarity bias but may affect relevance; graph-based approaches better capture exposure and echo-chamber structure but require richer interaction data; multimodal methods improve semantic diversity but increase computational and privacy costs; and LLM-based reranking improves controllability but remains expensive and less directly optimized.

As summary, these observations underscore that diversity-aware recommendation is not a purely algorithmic problem. It is a design problem that requires explicit prioritisation among competing objectives, careful attention to representation choices, and evaluation protocols that capture both immediate list properties and longer-term exposure patterns. The framework presented here is intended as a practical reference that enables more deliberate, reproducible, and balanced system design.

Recent studies in broader machine intelligence further illustrate the relevance of architecture optimization, federated learning, multimodal perception, and resource-constrained deployment for designing scalable intelligent systems (Agarwal & Gupta, 2026; Agarwal et al., 2025; Aggarwal et al., 2025; Arya et al., 2026; Bai et al., 2026; Bmotra et al., 2025; Garg, 2025; Gupta et al., 2025; Khan & Singh, 2025; Khan et al., 2025; Kumar et al., 2026; Kumar & Lehal, 2026; Kumar, 2025; Malik & Habib, 2025; Myakala et al., 2025; Pereira et al., 2025; Ranjit et al., 2026; Sarvakar & Rana, 2025; Singh et al., 2025; Sharma et al., 2025; Tiwari et al., 2025; Wang & Zhao, 2026). These developments provide methodological context for future diversity-aware recommender-system research.

9. Conclusion

This study reviewed diversity in video recommender systems through a multidimensional perspective that connects content, producer, viewpoint, temporal, semantic, and popularity-related dimensions.

Although existing approaches (reranking, deep learning, graph-based models, and multimodal methods) have achieved significant progress, they remain insufficient to fully capture this multidimensional structure. The study further identifies key gaps in the literature, including insufficient diversity metrics, lack of geometric modeling, limited availability of realistic multimodal datasets, and inadequate integration of ethical considerations into system design. These limitations highlight the need to reconsider recommender systems not only from a technical performance perspective but also in terms of their societal impact.

Based on the proposed future research directions, it is emphasized that recommender systems should evolve toward multi objective optimization frameworks, geometric modeling of diversity in embedding spaces, long term user satisfaction objectives, and human centered design paradigms. This transformation redefines recommender systems

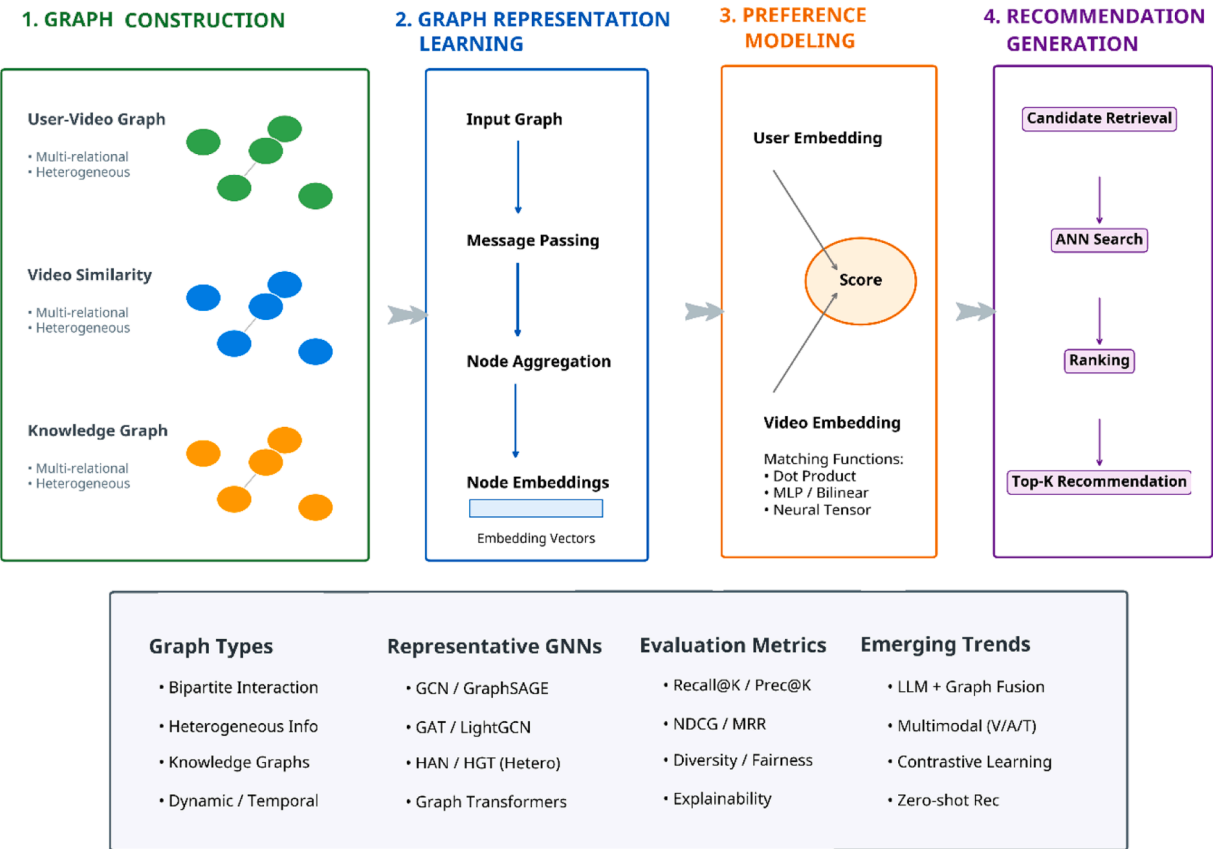


Fig. 2. Graph and network based video recommendation framework.

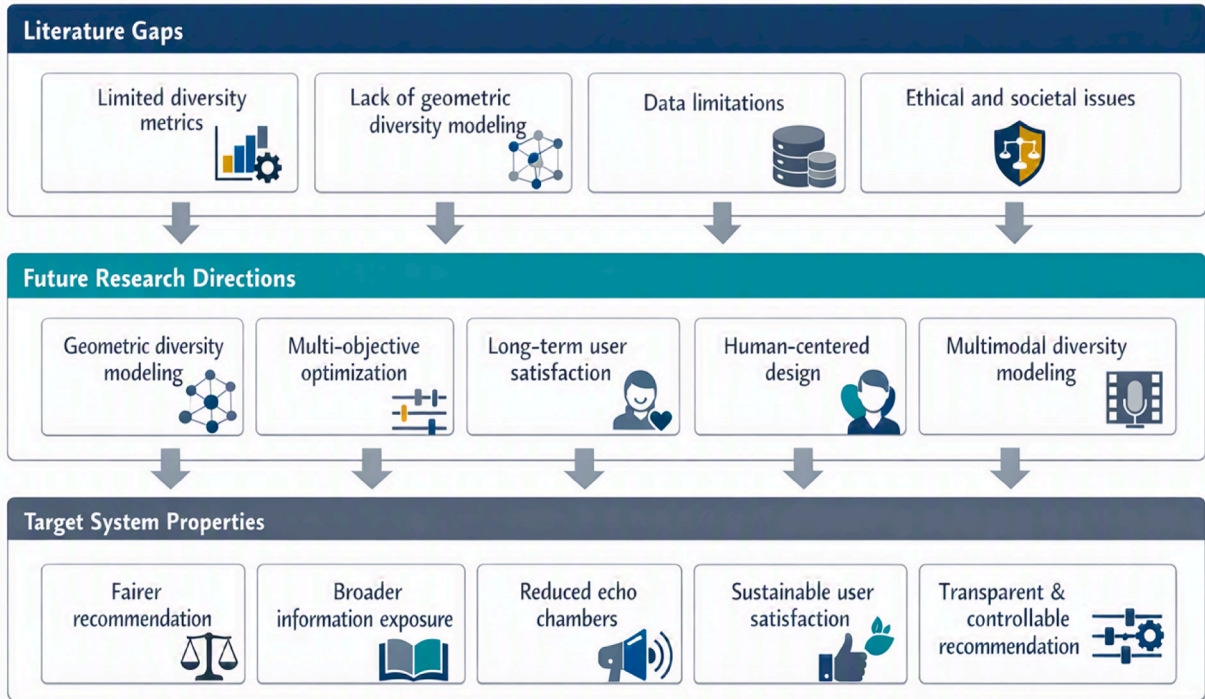


Fig. 3. Gaps in the literature and future research concepts.

from a classical user–item matching problem into a broader engineering problem that governs information distribution. The stages are shown in detail in Fig. 2.

Future work should develop geometry-aware metrics, multi-objective optimization, long-term satisfaction models, and human-centered systems. In this sense, multidimensional diversity functions

as a unifying lens for designing fairer, more transparent, and more sustainable video recommender systems.

Declaration of ethical standards

The authors declare that this review complies with ethical standards for authorship, originality, citation, and academic integrity. Because it uses publicly available literature and involves no human participants, animals, personal data, or experiments, ethics committee approval was not required.

Use of AI tools

AI tools were used only for language editing and grammatical refinement. The scientific content, analysis, interpretations, and conclusions remain the authors' responsibility.

Credit authorship contribution statement

Conceptualization: Literature Review, current VR systems, basic concepts, diversity, gaps, future direction; Author 1 and Author 2; Methodology: Author 1; Analysis and Interpretation: Author 1 and Author 2; Writing: Original Draft: Author 1; Writing Review and Editing: Author 1 and Author 2; Visualization: Author 1; Supervision: Author 2. Both authors approved the final manuscript.

Declaration of competing interest

The authors declare no competing financial interests, personal relationships, or conflicts of interest related to this work.

Acknowledgements

No funding, grant, or other financial support was received for this study. The Open Access publication of this article was funded by the University of Essex.

Data availability

No data was used for the research described in the article.

References

- Afreen, J., Mohaghegh, M., & Dobarjeh, M. (2025). Systematic literature review on bias mitigation in generative AI. *AI and Ethics*. <https://doi.org/10.1007/s43681-025-00721-9>
- Agarwal, S., Kandpal, M., Gupta, V. K., Aggarwal, S., & Jain, P. (2025). Secure and decentralized heart sound analysis using federated learning and blockchain technology. *ICCK Transactions on Machine Intelligence*, 2(1), 1–11. <https://doi.org/10.62762/TMI.2025.567350>
- Agarwal, S. K., & Gupta, Y. K. (2026). A Novel approach of progressive transfer learning for MRI brain tumor classification using VGG16 and MobileNet architectures. *ICCK Transactions on Machine Intelligence*, 2(1), 28–37. <https://doi.org/10.62762/TMI.2025.367009>
- Aggarwal, V., Aggarwal, S., Bhardwaj, A., & Gupta, V. K. (2025). Quantum computing essentials: Bridging theory and practice for new learners. *ICCK Transactions on Machine Intelligence*, 1(3), 117–126. <https://doi.org/10.62762/TMI.2025.173543>
- Amara, A., Taieb, M. A. H., & Aouicha, M. B. (2025). A multi-view GNN-based network representation learning framework for recommendation systems. *Neurocomputing*, 619, Article 129001. <https://doi.org/10.1016/j.neucom.2024.129001>
- Arya, S., Anju, & Onah, J. N. (2026). A data-driven framework for methane emission prediction using machine learning methods. *ICCK Transactions on Machine Intelligence*, 2(1), 53–64. <https://doi.org/10.62762/TMI.2025.782852>
- Avalle, M., Di Marco, N., Etta, G., Sangiorgio, E., Alipour, S., Bonetti, A., & Quattrociochi, W. (2024). Persistent interaction patterns across social media platforms and over time. *Nature*, 628(8008), 582–589. <https://doi.org/10.1038/s41586-024-07229-y>
- Bai, Y., Hu, J., Tursun, E., & Yimit, H. (2026). Data-driven operational assessment method and digital twin system for unmanned surface vehicles. *ICCK Transactions on Machine Intelligence*, 2(1), 38–52. <https://doi.org/10.62762/TMI.2025.444910>
- Beynon, A., Hendry, D., Lund Rasmussen, C., Rohl, A. L., Eynon, R., Thomas, G., Stearne, S., Campbell, A., Harris, C., Zabattiero, J., & Straker, L. (2024). Measurement method options to investigate digital screen technology use by children and adolescents: A narrative review. *Children*, 11(7), 754. <https://doi.org/10.3390/children11070754>
- Bmotra, M., Kumar, R., & Dhiman, G. (2025). Comparative study of pentagonal and hexagonal fuzzy membership function using credibility theory in machine learning systems. *ICCK Transactions on Machine Intelligence*, 1(3), 127–137. <https://doi.org/10.62762/TMI.2025.922612>
- Carraro, D., & Bridge, D. (2024). Enhancing recommendation diversity by reranking with large language models. *ACM Transactions on Recommender Systems*. <https://doi.org/10.1145/3700604>
- Chaudhari, A., Seddig, A. H., Sarlan, A., & Raut, R. (2024). A hybrid recommendation system: A review. *IEEE Access: Practical Innovations, Open Solutions*, 12, 157107–157126. <https://doi.org/10.1109/ACCESS.2024.3480693>
- Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, X. (2023). Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3), 1–39. <https://doi.org/10.1145/3564284>
- Clayborne, Z. M., Capaldi, C. A., & Mehra, V. M. (2025). Associations between digital media use behaviours, screen time and positive mental health in youth: Results from the 2019 Canadian health survey on children and youth. *BMC Public Health*, 25(1), 2303. <https://doi.org/10.1186/s12889-025-22874-2>
- Dehghankar, M., & Asudeh, A. (2025). Rank it, then ask it: Input reranking for maximizing the performance of llms on symmetric tasks. In , 2. *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V* (pp. 451–462). <https://doi.org/10.1145/3711896.3737099>
- Deldjoo, Y., & Di Noia, T. (2025). CFaiRLLM: Consumer fairness evaluation in large-language model recommender system. *ACM Transactions on Intelligent Systems and Technology*. <https://doi.org/10.1145/3725853>
- Ding, C. (2025). HMR-LLM: Hybrid multimodal sequential recommendation with LLM reranking. In. In *2025 IEEE International Conference on Big Data (BigData)* (pp. 5230–5237). IEEE. <https://doi.org/10.1109/BigData6926.2025.11401803>
- Dong, L., He, G., & Sun, A. (2024). Not all videos become outdated: Short-Video recommendation by learning to deconfound release interval bias. In *Proceedings of the 18th ACM Conference on Recommender Systems* (pp. 179–188). <https://doi.org/10.1145/3640457.3688113>
- Du, B., Ye, Z., Wu, Z., Jankowska, M., Ai, Q., & Liu, Y. (2025). Understanding the effect of opinion polarization in short video browsing. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 906–916). <https://doi.org/10.1145/3726302.3730107>
- Fang, H., Tao, S., Chen, N., Chang, K. X., & Sakai, T. (2025). Do large language models favor recent content? A study on recency bias in LLM-based reranking. In *Proceedings of the 2025 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region* (pp. 85–94). <https://doi.org/10.1145/3767695.3769493>
- Frielingsdorf, H., Pomichov, V., Rystedt, I., Lindstrand, S., Korhonen, L., & Henriksen, H. (2025). Associations of time spent on different types of digital media with self-rated general and mental health in Swedish adolescents. *Scientific Reports*, 15(1), 993. <https://doi.org/10.1038/s41598-024-83951-x>
- ... & Gao, C., Zheng, Y., Li, N., Li, Y., Qin, Y., Piao, J., & Li, Y. (2023a). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Transactions on Recommender Systems*, 1(1), 1–51. <https://doi.org/10.1145/3568022>
- Gao, J., Chen, B., Zhao, X., Liu, W., Li, X., Wang, Y., & Tang, R. (2025). Llm4rerank: Llm-based auto-reranking framework for recommendations. In *Proceedings of the ACM on Web Conference 2025* (pp. 228–239). <https://doi.org/10.1145/3696410.3714922>
- Gao, Y., Liu, F., & Gao, L. (2023b). Echo chamber effects on short video platforms. *Scientific Reports*. <https://doi.org/10.1038/s41598-023-33370-1>
- Garg, M. (2025). Enhanced reinforcement learning-based resource scheduling for secure blockchain networks in IIoT. *ICCK Transactions on Machine Intelligence*, 1(1), 29–41. <https://doi.org/10.62762/TMI.2024.529242>
- Gu, Y., Zou, L., & Li, C. (2026). Deep learning to rank in industrial search engines, recommender systems, and online advertising: An overview and new perspectives. *ACM Transactions on Information Systems*, 44(4), 1–52. <https://doi.org/10.1145/3797895>
- Gupta, S., Adhikari, U., Roy, D., & Hazra, S. (2025). IoT-integrated reinforcement learning-based mine detection system for military and humanitarian applications. *ICCK Transactions on Machine Intelligence*, 1(1), 17–28. <https://doi.org/10.62762/TMI.2025.235880>
- Han, J., Li, W., Cai, Z., & Li, Y. (2022). Multi-aggregator time-warping heterogeneous graph neural network for personalized micro-video recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 676–685). <https://doi.org/10.1145/3511808.3557403>
- Hassani Saissi, M., Idrissi, N., & Zellou, A. (2025). Understanding echo chambers in recommender systems: A systematic review. *International Journal of Advanced Computer Science & Applications*, 16(10). <https://doi.org/10.14569/ijacsa.2025.0161071>
- He, Y., Wu, G., Cai, D., & Hu, X. (2023). Meta-path based graph contrastive learning for micro-video recommendation. *Expert Systems with Applications*, 222, Article 119713. <https://doi.org/10.1016/j.eswa.2023.119713>
- Jin, C., Peng, H., Zhao, S., Wang, Z., Xu, W., Han, L., & Metaxas, D. N. (2025). Apeer: Automatic prompt engineering enhances large language model reranking. In *Companion Proceedings of the ACM on Web Conference 2025* (pp. 2494–2502). <https://doi.org/10.1145/3701716.3717574>
- Karo, I. M. K., Dewi, S., & Perdana, A. (2024). Implementasi text summarization pada review aplikasi digital library system menggunakan metode maximum marginal relevance. *JEKIN-Jurnal Teknik Informatika*, 4(1), 25–31. <https://doi.org/10.58794/jekin.v4i1.671>

- Kelly, J., & Bridge, D. (2006). Enhancing the diversity of conversational collaborative recommendations: A comparison. *Artificial Intelligence Review*, 25, 79–95. <https://doi.org/10.1007/s10462-007-9023-8>
- Khan, A., & Singh, J. (2025). A novel image captioning technique using deep learning methodology. *ICCK Transactions on Machine Intelligence*, 1(2), 52–68. <https://doi.org/10.62762/TMI.2025.886122>
- Khan, M., Kumar, R., & Dhiman, G. (2025). A hybrid machine learning fuzzy non-linear regression approach for neutrosophic fuzzy set. *ICCK Transactions on Machine Intelligence*, 1(1), 42–51. <https://doi.org/10.62762/TMI.2025.561363>
- Khanam, K. Z., Srivastava, G., & Mago, V. (2023). The homophily principle in social network analysis: A survey. *Multimedia Tools and Applications*, 82(6), 8811–8854. <https://doi.org/10.1007/s11042-021-11857-1>
- Kim, J., Kim, E., Yeo, K., Jeon, Y., Kim, C., Lee, S., & Lee, J. (2024). Content-based graph reconstruction for cold-start item recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1263–1273). <https://doi.org/10.1145/3626772.3657801>
- Klimashevskaya, A., Jannach, D., Elahi, M., & Trattner, C. (2024). A survey on popularity bias in recommender systems. *User Modeling and User-Adapted Interaction*, 34(5), 1777–1834. <https://doi.org/10.1007/s11257-024-09406-0>
- Kumar, A. (2025). Clinical text analytics: Techniques, deep learning models, and the future of medical text analytics. *ICCK Transactions on Machine Intelligence*, 1(3), 148–165. <https://doi.org/10.62762/TMI.2025.451731>
- Kumar, A., & Lehal, G. S. (2026). Detection of newspaper layouts using YOLO12. *ICCK Transactions on Machine Intelligence*, 2(2), 77–87. <https://doi.org/10.62762/TMI.2025.846033>
- Kumar, V., Saxena, A., & Chawla, L. (2026). Adaptive learning density estimators for Tsallis Entropy and Kapur entropy with applications in system training. *ICCK Transactions on Machine Intelligence*, 2(1), 12–27. <https://doi.org/10.62762/TMI.2025.317970>
- Lei, F., Cao, Z., Yang, Y., Ding, Y., & Zhang, C. (2023). Learning the user's deeper preferences for multi-modal recommendation systems. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(3s), 1–18. <https://doi.org/10.1145/3573010>
- Li, F., Huang, J., Tang, S., Han, B., Cao, H., Sui, H., & Kang, X. (2025). Contrastive prototype framework for calibrating video recommendation. In *Proceedings of the 33rd ACM International Conference on Multimedia* (pp. 6373–6382). <https://doi.org/10.1145/3746027.3755869>
- Li, F., Si, X., Tang, S., Wang, D., Han, K., Han, B., & Chen, H. (2024a). Contextual distillation model for diversified recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 5307–5316). <https://doi.org/10.1145/3637528.3671514>
- Li, N., Pan, Y., Gao, C., Jin, D., & Liao, Q. (2024b). Full-stage diversified recommendation: Large-scale online experiments in short-video platform. In *Proceedings of the ACM on Web Conference 2024* (pp. 4565–4574). <https://doi.org/10.1145/3589334.3648144>
- Li, W., Huang, R., Zhao, H., Liu, C., Zheng, K., Liu, Q., Mou, N., Zhou, G., Lian, D., Song, Y., Bao, W., Yu, E., & Ou, W. (2024c). DimeRec: A unified framework for enhanced sequential recommendation via generative diffusion models. In *Proceedings of the 18th ACM International Conference on Web Search and Data Mining*. <https://doi.org/10.1145/3701551.3703555>
- ... & Li, Z., Yang, C., Chen, Y., Wang, X., Chen, H., Xu, G., & Sheng, M. (2024d). Graph and sequential neural networks in session-based recommendation: A survey. *ACM Computing Surveys*, 57(2), 1–37. <https://doi.org/10.1145/3696413>
- Liang, R., Zhang, C., Li, L., Wang, J., Zhu, X., & Sun, A. (2025). Tvr-ranking: A dataset for ranked video moment retrieval with imprecise queries. In *Proceedings of the 2025 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region* (pp. 231–239). <https://doi.org/10.48550/arXiv.2407.06597>. arXiv preprint.
- Liao, M., Sundar, S. S., & Walther, J. B. (2022). User trust in recommendation systems: A comparison of content-based, collaborative and demographic filtering. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1–14). <https://doi.org/10.1145/3491102.3501936>
- Lin, J., Li, Q., Xie, G., Guan, Z., Jiang, Y., Xu, T., & Zhao, P. (2024). Mitigating sample selection bias with robust domain adaption in multimedia recommendation. In *Proceedings of the 32nd ACM International Conference on Multimedia* (pp. 7581–7590). <https://doi.org/10.1145/3664647.3680615>
- Liu, H., Li, C., & Tian, L. (2022). Multi-modal graph attention network for video recommendation. In *2022 IEEE 5th International Conference on Computer and Communication Engineering Technology (CCET)* (pp. 94–99). IEEE. <https://doi.org/10.1109/CCET55412.2022.9906399>
- Liu, J. (2023). Toward a two-sided fairness framework in search and recommendation. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval* (pp. 236–246). <https://doi.org/10.1145/3576840.3578332>
- Liu, K., Wang, J., & Zhang, X. (2024a). Debiased momentum contrastive learning for multimodal video similarity measures. *Neurocomputing*, 563, Article 126938. <https://doi.org/10.1016/j.neucom.2023.126938>
- ... & Liu, N., Hu, X. E., Savas, Y., Baum, M. A., Berinsky, A. J., Chaney, A. J., & Stewart, B. M. (2025). Short-term exposure to filter-bubble recommendation systems has limited polarization effects: Naturalistic experiments on YouTube. *Proceedings of the National Academy of Sciences*, 122(8), Article e2318127122. <https://doi.org/10.1073/pnas.2318127122>
- ... & Liu, Q., Hu, J., Xiao, Y., Zhao, X., Gao, J., Wang, W., & Tang, J. (2024b). Multimodal recommender systems: A survey. *ACM Computing Surveys*, 57(2), 1–17. <https://doi.org/10.1145/3695461>
- ... & Liu, Q., Zhu, J., Yang, Y., Dai, Q., Du, Z., Wu, X. M., & Dong, Z. (2024c). Multimodal pretraining, adaptation, and generation for recommendation: A survey. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 6566–6576). <https://doi.org/10.48550/arXiv.2404.00621>. arXiv preprint.
- Liu, Y. (2025). A generative and discriminative model for diversity-promoting recommendation. *Information Systems*, 128, Article 102488. <https://doi.org/10.1016/j.is.2024.102488>
- Liu, Y., & Zhang, Y. (2025). Diversity-promoting recommendation with dual-objective optimization and dual consideration. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/TKDE.2025.3543285>
- Liu, Z., Wen, Z., Tian, Z., Huang, Z., Zhu, M., Wei, Z., & Li, D. (2026). Rethinking the hidden risk of reranking: Achieving risk-aware reranking with information gain for RAG with LLMs. In *Proceedings of the ACM Web Conference 2026* (pp. 1887–1898). <https://doi.org/10.1145/3774904.3792085>
- Lubos, S., Felfernig, A., & Tautschnig, M. (2023). An overview of video recommender systems: State-of-the-art and research issues. *Frontiers in Big Data*, 6, Article 1281614. <https://doi.org/10.3389/fdata.2023.1281614>
- Ma, F. (2025). Learning behavior analysis and personalized recommendation system of online education platform based on machine learning. *Computers and Education: Artificial Intelligence*, 8, Article 100408. <https://doi.org/10.1016/j.caeai.2025.100408>
- Ma, H., Xie, R., Meng, L., Feng, F., Du, X., Sun, X., & Meng, X. (2024). Negative sampling in recommendation: A survey and future directions. *ACM Transactions on Information Systems*. <https://doi.org/10.1145/3793855>
- Maghsoudi, M., Valikhani, M. H., & Zohdi, M. H. (2025). Network-based video recommendation using viewing patterns and modularity analysis: An integrated framework. *IEEE Access: Practical Innovations, Open Solutions*, 13, 5660–5678. <https://doi.org/10.1109/ACCESS.2025.3526876>
- Malik, I. A., & Habib, G. (2025). AI enabled resource-constrained computing architectures for IoT devices. *ICCK Transactions on Machine Intelligence*, 1(3), 138–147. <https://doi.org/10.62762/TMI.2025.225921>
- Malitesta, D., Cornacchia, G., Pomo, C., Merra, F. A., Di Noia, T., & Di Sciascio, E. (2025). Formalizing multimedia recommendation through multimodal deep learning. *ACM Transactions on Recommender Systems*, 3(3), 1–33. <https://doi.org/10.1145/3662738>
- Martina, A. F. M., Petruzzelli, A., Musto, C., de Gemmis, M., Lops, P., & Semeraro, G. (2025). DistillRecDial: A knowledge-distilled dataset capturing user diversity in conversational recommendation. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems (RecSys '25)*, Prague, Czech Republic, September 22–26, 2025 (pp. 726–735). <https://doi.org/10.1145/3705328.3748161>
- Martins, G. B., Papa, J. P., & Adeli, H. (2020). Deep learning techniques for recommender systems based on collaborative filtering. *Expert Systems*, 37(6), Article e12647. <https://doi.org/10.1111/exsy.12647>
- Mu, Y., & Wu, Y. (2023). Multimodal movie recommendation system using deep learning. *Mathematics*, 11(4), 895. <https://doi.org/10.3390/math11040895>
- Myakala, P. K., Kamatala, S., & Bura, C. (2025). Privacy-preserving federated learning for IoT botnet detection: A federated averaging approach. *ICCK Transactions on Machine Intelligence*, 1(1), 6–16. <https://doi.org/10.62762/TMI.2025.796490>
- Ni, Y., Cheng, Y., Liu, X., Fu, J., Li, Y., He, X., Zhang, Y., & Yuan, F. (2023). A content-driven micro-video recommendation dataset at scale. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management*. <https://doi.org/10.1145/3746252.3761655>
- Niu, S., Li, T., & Chi, M. (2025). A literature review of ethical considerations in recommender systems for user-generated content in human-computer interaction. *ACM Transactions on Recommender Systems*. <https://doi.org/10.1145/3770747>
- Pereira, M. S., de Azevedo, T. M., Ribeiro, C. T., Freire, A. I., Francisco, M. B., Pereira, J. L. J., & de Paiva, A. P. (2025). Discriminating planted capsicum Spp. Varieties via machine learning and multivariate data reduction. *ICCK Transactions on Machine Intelligence*, 1(3), 166–185. <https://doi.org/10.62762/TMI.2025.385133>
- Pomo, C., Attimonelli, M., Danese, D., Narducci, F., & Di Noia, T. (2025). Do recommender systems really leverage multimodal content? A comprehensive analysis on multimodal representations for recommendation. In *Proceedings of the 34th ACM International Conference on Information and Knowledge Management* (pp. 2377–2387). <https://doi.org/10.1145/3746252.3761398>
- Ranjit, K. C., Shrestha, B., & Ghimire, U. (2026). Optimizing CNN architectures for steering angle prediction for self-driving vehicles in unstructured roads: A comparative study of activation functions and model complexity. *ICCK Transactions on Machine Intelligence*, 2(2), 88–99. <https://doi.org/10.62762/TMI.2025.759110>
- Rawka, T., & Jadeja, M. (2022). A survey on graph neural network based video recommendation system. *Deep learning for social media data analytics* (pp. 141–155). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-10869-3_8
- ... & Ren, R., Wang, Y., Zhou, K., Zhao, W. X., Wang, W., Liu, J., & Chua, T. S. (2025). Self-calibrated listwise reranking with large language models. In *Proceedings of the ACM on Web Conference 2025* (pp. 3692–3701). <https://doi.org/10.1145/3696410.3714658>
- Ren, Z., Wu, Y., Lu, J., Ary, A., Liu, Y., Wei, L., & Heldt, L. (2026). Learning to alleviate familiarity bias in video recommendation. arXiv preprint [doi: 10.48550/arXiv.2602.07987](https://arxiv.org/abs/2602.07987)
- Saravakar, K., & Rana, K. (2025). A hybrid framework combining CNN, LSTM, and transfer learning for emotion recognition. *ICCK Transactions on Machine Intelligence*, 1(2), 103–116. <https://doi.org/10.62762/TMI.2025.572412>
- Shang, Y., Gao, C., Chen, J., Jin, D., & Li, Y. (2024). Improving item-side fairness of multimodal recommendation via modality debiasing. In *Proceedings of the ACM Web Conference 2024* (pp. 4697–4705). <https://doi.org/10.1145/3589334.3648156>
- ... & Sharifmoghaddam, S., Pradeep, R., Slavescu, A., Nguyen, R., Xu, A., Chen, Z., & Lin, J. (2025). Rankllm: A python package for reranking with llms. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in*

- Information Retrieval (pp. 3681–3690). <https://doi.org/10.48550/arXiv.2505.19284>. arXiv preprint.
- Sharma, S., Devgan, K., Jangra, D., Bhardwaj, A., & Aggarwal, S. (2025). Neuro-inspired alert system for air quality prediction using ensemble preprocessing and SNN classification. *ICCK Transactions on Machine Intelligence*, 1(2), 69–79. <https://doi.org/10.62762/TMI.2025.403059>
- Shi, X., Liu, Q., Xie, H., Wu, D., Peng, B., Shang, M., & Lian, D. (2023). Relieving popularity bias in interactive recommendation: A diversity-novelty-aware reinforcement learning approach. *ACM Transactions on Information Systems*, 42(2), 1–30. <https://doi.org/10.1145/3618107>
- Singh, S., Aggarwal, S., Singh, A., & Sharma, A. (2025). Integrating artificial intelligence and machine learning in autism detection via gut microbiome analysis. *ICCK Transactions on Machine Intelligence*, 1(2), 90–102. <https://doi.org/10.62762/TMI.2025.682666>
- Tiwari, S., Kumar, D., Mahajan, A., & Sachar, S. (2025). Emotion detection from speech using CNN-BiLSTM with feature rich audio inputs. *ICCK Transactions on Machine Intelligence*, 1(2), 80–89. <https://doi.org/10.62762/TMI.2025.306750>
- Tran, G., Jung, J., & Han, J. (2025). Mitigating social polarization in video sharing platforms using unbiased recommendation systems: A case study of South Korea political YouTube channels. *Technology in Society*. <https://doi.org/10.1016/j.techsoc.2025.102881>
- Treuillier, C., Castagnos, S., Dufraisse, E., Özgöbek, Ö., & Brun, A. (2026). Shift your focus for the greater good: Improving fairness at no cost for accuracy and diversity in news recommender systems. *ACM Transactions on Recommender Systems*. <https://doi.org/10.1145/3790097>
- Wang, D., & Zhao, X. (2022). Affective video recommender systems: A survey. *Frontiers in Neuroscience*, 16, Article 984404. <https://doi.org/10.3389/fnins.2022.984404>
- Wang, S., Zhang, X., Wang, Y., & Ricci, F. (2024a). Trustworthy recommender systems. *ACM Transactions on Intelligent Systems and Technology*, 15(4), 1–20. <https://doi.org/10.1145/3627826>
- Wang, X., & Zhao, R. (2026). FedTLRec: Federated recommendation with transformer-based parameter aggregation and LoRA compression. *ICCK Transactions on Machine Intelligence*, 2(1), 65–76. <https://doi.org/10.62762/TMI.2025.882476>
- Wang, Y., Xia, H., & Liu, Y. (2024b). Cmc-mm: Multi-modal recommendation model with cross-modal correction. *Journal of Intelligent Information Systems*, 62(5), 1187–1211. <https://doi.org/10.1007/s10844-024-00848-x>
- Wei, W., Zhao, S., & Zou, D. (2023). Recommendation systems: A survey and new perspectives. *World Scientific Annual Review of Artificial Intelligence*, 1. <https://doi.org/10.1142/s2811032323300013>, 2330001:1–2330001:24.
- Wu, H., Song, R., Ma, L., Liu, Z., Yang, J., & Qiao, Y. (2024). The debiasing recommendation algorithm based on pseudo-label uncertainty verification. In *Proceedings of the 2024 10th International Conference on Communication and Information Processing* (pp. 668–674). <https://doi.org/10.1145/3708657.3708764>
- Wu, L., He, X., Wang, X., Zhang, K., & Wang, M. (2022). A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 35(5), 4425–4445. <https://doi.org/10.1109/TKDE.2022.3145690>
- Xiong, Y., Tu, X., & Zhao, W. (2025). AFR-Rank: An effective and highly efficient LLM-based listwise reranking framework via filtering noise documents. *Information Processing & Management*, 62(6), Article 104232. <https://doi.org/10.1016/j.ipm.2025.104232>
- Yang, F., & Geng, X. (2026). Dynamic debiasing for popularity bias in recommendation systems: A reinforcement learning approach. *Journal of Intelligent Information Systems*. <https://doi.org/10.1007/s10844-026-01034-x>
- Yang, H., Fu, L., Lu, Q., Fan, Y., Zhang, T., & Wang, R. (2025a). Research on the design of a short video recommendation system based on multimodal information and differential privacy. In *Proceedings of the 2025 4th International Conference on Cyber Security, Artificial Intelligence and the Digital Economy (CSAIDE 2025)*. <https://doi.org/10.1145/3729706.3729714>
- Yang, Y., Huang, C., Xia, L., Huang, C., Luo, D., & Lin, K. (2023). Debaised contrastive learning for sequential recommendation. In *Proceedings of the ACM Web Conference 2023* (pp. 1063–1073). <https://doi.org/10.1145/3543507.3583348>
- Yang, Y., Liu, Y., & Naghizadeh, P. (2025b). Adaptive bounded exploration and intermediate actions for data debiasing. *INFORMS Journal on Computing*. <https://doi.org/10.1287/ijoc.2024.0651>
- Yang, Y., & Zhao, Y. (2023). Personalized sports health recommendation system assisted by Q-learning algorithm. *International Journal of Human-Computer Interaction*, 41(4), 1889–1901. <https://doi.org/10.1080/10447318.2023.2295693>
- ... & Ye, Y., Zheng, Z., Shen, Y., Wang, T., Zhang, H., Zhu, P., & Xiong, H. (2025). Harnessing multimodal large language models for multimodal sequential recommendation. In , 39. *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 13069–13077). <https://doi.org/10.1609/aaai.v39i1.2.33426>.
- Yuan, H., & Hernandez, A. A. (2023). User cold start problem in recommendation systems: A systematic review. *IEEE Access : Practical Innovations, Open Solutions*, 11, 136958–136977. <https://doi.org/10.1109/ACCESS.2023.3338705>
- Zhang, F., Luo, W., & Yang, X. (2025). Mitigating selection bias in recommendation systems through sentiment analysis and dynamic debiasing. *Applied Sciences*, 15(8), 4170. <https://doi.org/10.3390/app15084170>
- Zhao, X., Wang, M., Zhao, X., Li, J., Zhou, S., Yin, D., Li, Q., Tang, J., & Guo, R. (2023). Embedding in recommender systems: A survey. arXiv preprint [doi: 10.48550/arXiv.2310.18608](https://arxiv.org/abs/10.48550/arXiv.2310.18608).
- Zhao, Y., Wang, Y., Liu, Y., Cheng, X., Aggarwal, C., & Derr, T. (2023b). Fairness and diversity in recommender systems: A survey. *ACM Transactions on Intelligent Systems and Technology*, 16, 1–28. <https://doi.org/10.1145/3664928>