ACM Transactions on Recommender Systems

Special Issue on Challenges in Modern Multimodal Recommender Systems

Guest Editors:
- **Yubin Kim**, Vody, yubink.cs@gmail.com
- **Daniele Malitesta**, Université Paris-Saclay, d.malitesta@gmail.com
- **Alberto Carlo Maria Mancino**, Politecnico di Bari, alberto.mancino@poliba.it
- **Claudio Pomo**, Politecnico di Bari, claudio.pomo@poliba.it
- **Shah Nawaz,** Johannes Kepler University Linz, shah.nawaz@jku.at

Recommender systems (RS) have been reshaped in recent years by *multimodal* approaches that leverage multiple types of data and content. **In this Special Issue, the term "multimodal" encompasses any combination of two or more different data modalities – including, but not limited to, text, images, audio, video, structured metadata, knowledge graphs, and behavioral or sensor signals.** Modern multimodal recommender systems (MMRS) now underpin major application domains, including e-commerce, entertainment, news and media platforms, social networks, and personalized content delivery. These systems are increasingly powered by deep learning, large foundation models, and generative AI, which enable rich multimodal representations and complex user–item interactions at scale.

Despite this progress, **core challenges in designing, deploying, and evaluating MMRS remain**. Such systems must cope with **noisy, incomplete, or biased multimodal data**, including cases of missing or misaligned modalities (e.g., an item lacking an image or a user profile missing certain data). Models face difficulties in **effectively fusing heterogeneous modalities and remain** robust against distribution shifts or modality drops in dynamic real-world environments. Moreover, the boundary between *retrieval* and *recommendation* needs careful consideration: many recommender pipelines include a multimodal retrieval stage (candidate generation) followed by personalized ranking. This special issue emphasizes **personalized recommendation** scenarios that leverage multimodal content, distinguishing them from generic multimedia retrieval. We welcome research that addresses both multimodal candidate retrieval and personalized ranking, as they jointly contribute to better recommendations.

Another critical challenge lies in the **evaluation and feedback loops** for MMRS. Traditional accuracy metrics on static test splits often fail to capture robustness, fairness, and user-centric outcomes. In practice, **online evaluation methods** (such as A/B testing or multi-armed bandits) and feedback loop analyses are needed to understand how multimodal recommenders perform with real users over time, including issues such as exposure bias and system-induced feedback effects. Likewise, ensuring **data quality** across modalities is paramount: noise, modality imbalance, and skewed representations can significantly impact fairness and user trust. Addressing **ethical and rights considerations** – privacy of user data, copyright compliance for content (e.g., images or videos), content moderation, and misuse prevention – is essential for safe deployment of multimodal recommenders.

Within this context, the **Special Issue on "Challenges in Modern Multimodal Recommender Systems"** aims to advance the state of the art by bringing together new insights on these problems. We invite original contributions that **analyze and characterize key challenges across the MMRS lifecycle**, propose **novel algorithms and architectures** to tackle multimodal recommendation tasks, and **bridge theoretical advances with real-world applications**. The scope of the special issue is broad: we welcome both **foundational research** (e.g., new models, learning methods, theory) and **applied/industrial work** (e.g., deployment case studies, systems engineering for MMRS), with a strong emphasis on **data quality, robustness, fairness, and rigorous evaluation**. Interdisciplinary works spanning recommender systems, information retrieval, multimedia processing, machine learning, data mining, human–computer interaction, and AI ethics are particularly encouraged. Our goal is to foster MMRS that are not only more powerful in recommendation accuracy, but also more **reliable, fair, transparent, and user-centric** in practice.

Topics

We welcome submissions on **topics including, but not limited to, the following**:

- **Data-centric challenges in MMRS:** Noisy, incomplete, or conflicting multimodal signals; handling missing modalities and imputation strategies; misalignment between different modal content and user intent; bias in multimodal data sources (text, images, audio, video) and its impact on recommendations; methods and tools for assessing and improving multimodal data quality.
- **Modeling and algorithmic innovations:** Multimodal representation learning techniques and robust fusion strategies; architectures for sparse, cold-start, or long-tail recommendation scenarios using multimodal data; models that remain **robust to distribution shifts** or corrupted inputs; model-agnostic approaches for denoising or debiasing multimodal feature sets.
- **Evaluation and benchmarking:** Development of metrics and evaluation protocols for multimodal recommenders **beyond accuracy** (e.g., robustness, diversity, fairness, transparency, user satisfaction); construction, curation, and sharing of realistic multimodal recommendation datasets; benchmarking methodologies and open-source evaluation toolkits for MMRS.
- **Online evaluation and feedback loops:** Techniques for evaluating MMRS in live environments, such as A/B testing frameworks and multivariate bandit experiments; analysis of **feedback loops**, popularity bias, and exposure bias unique to multimodal recommenders; methods to mitigate feedback-induced bias and to incorporate online user feedback safely into model updates.
- **Deployment, monitoring, and MLOps:** Efficiency and scalability considerations for large-scale multimodal recommender deployment (e.g. reducing inference latency, system optimization); **drift detection** and handling evolving data distributions or user preferences over time; managing modality availability issues at runtime (graceful degradation when a modality is missing or unreliable); continuous monitoring, auditing, and logging of recommender performance (including fairness and compliance audits) in production.
- **Human-centered and ethical aspects:** User studies on trust, satisfaction, and user perception of multimodal recommendations; techniques for **multimodal explanations and visualizations** to increase transparency; fairness-aware and privacy-preserving MMRS (ensuring recommendations do not reinforce biases or violate user privacy); **safety and content rights** in multimodal content (e.g. content moderation filters, copyright and licensing considerations, prevention of malicious use of generated content); interactive and human-in-the-loop recommendation interfaces that leverage rich multimodal content to give users more control.

- **Generative and foundation models for MMRS:** Applications of large language models, vision–language models, and other foundation models in recommendation (e.g. using generative models to create or enrich item content, or to better understand user preferences); retrieval-augmented generation for recommendations; using generative AI to produce personalized content (such as explanations, summaries, or even synthetic recommended items) that enhance user experience in multimodal settings.

Important Dates
- Submission deadline: June 1, 2026
- First-round review decisions: August 1, 2026
- Deadline for revision submissions: October 1, 2026
- Notification of final decisions: February 1, 2027

Submission Information

This special issue welcomes a range of submission types, including **technical research papers**, **theoretical or methodological studies**, **applied case studies from industry**, **survey and review articles**, and **perspective/opinion pieces**. Submissions should clearly fall within the scope of the special issue and address one or more of the topics listed above. We particularly encourage contributions that **bridge the theoretical and practical aspects** of multimodal recommendation—for example, papers that introduce new algorithms accompanied by evaluations on real-world data, or papers that derive insights from industrial deployments with rigorous analysis. Emphasis should be given to **data quality, robustness, fairness, and comprehensive evaluation** to align with the theme of this special issue.

We also invite authors of outstanding papers from relevant conferences (e.g., **RecSys, SIGIR, KDD, The Web Conference (WWW), WSDM, CIKM, UMAP, CHI**, etc.) to submit **extended versions** of their work. An extended submission must contain **at least 30% new material** beyond the original publication – for instance, additional technical contributions, expanded experiments, new analysis, or theoretical extensions that significantly strengthen the work. Authors should explain in their cover letter how the journal version extends prior work. All submissions must be **original** and not under consideration for publication elsewhere.

Manuscripts should be prepared according to the **ACM TORS author guidelines** (see: https://dl.acm.org/journal/tors/author-guidelines) and submitted through the ACM TORS online submission system via **Manuscript Central** (at https://mc.manuscriptcentral.com/tors). When submitting, please select the option for the "*Special Issue on Challenges in Modern Multimodal Recommender Systems*."

For questions and further information, please contact the guest editors (Email: **d.malitesta@gmail.com**).