First International Workshop on Data Quality-Aware Multimodal Recommendation (DaQuaMRec @ RecSys2025)

Claudio Pomo[1], Dietmar Jannach[2], Yubin Kim[3], Daniele Malitesta[4], Alberto Carlo Maria Mancino[1], Julian McAuley[5], Alessandro B. Melchiorre[6,7] and Shah Nawaz[7]

[1]*Polytechnic University of Bari, Bari, Italy*
[2]*University of Klagenfurt, Klagenfurt, Austria*
[3]*Vody Inc., Los Angeles, USA*
[4]*Université Paris-Saclay, CentraleSupélec, Inria, Gif-sur-Yvette, France*
[5]*University of California, San Diego, La Lolla, USA*
[6]*Criteo AI Lab, Paris, France*
[7]*Institute of Computational Perception, Johannes Kepler University Linz, Linz, Austria*

Abstract

This preface introduces the Proceedings of the "First International Workshop on **Da**ta **Qua**lity-Aware **M**ultimodal **Rec**ommendation" (**DaQuaMRec**), that was co-located with the 19th ACM Conference on Recommender Systems (RecSys 2025) and held in Prague (Czech Republic) on September 22, 2025. DaQuaMRec offered a dedicated satellite event at RecSys 2025 for researchers and practitioners interested in data challenges arising in multimodal recommendation, such as noisy, missing, or corrupted data modalities, modality misalignment, bias and fairness in multimodal data, definition and evaluation of data quality in multimodal recommendation, and many other related topics. DaQuaMRec featured two keynote speeches, an invited talk, two paper sessions (with five presented papers), and a conclusive discussion panel. The workshop website is accessible at: https://sites.google.com/view/daquamrec2025.

Keywords

Recommendation, Multimodal Deep Learning, Data Quality-Aware Machine Learning, Noisy Multimodal Data, Missing Modalities, Multimodal Bias, Multimodal Fairness, Modalities Misalignment, Multimodal Data Quality Evaluation

1. Introduction

Multimodal recommender systems transform how we experience multimedia digital services, powering smarter suggestions in fashion, social media, music, food, and beyond. Combining data from images, text, and audio, these systems outperform traditional, single-modality approaches, enabling richer user profiling and more accurate recommendations. Driven by advances in deep learning and the rise of large foundation models, the multimodal recommendation is progressing remarkably. Yet, beneath the surface of these powerful models lies a crucial, often-overlooked challenge: ***data quality***.

DaQuaMRec, the "First International Workshop on **Da**ta **Qua**lity-Aware **M**ultimodal **Rec**ommendation" brought this foundational concern to the forefront, as a dedicated workshop event co-located with the 19th ACM Conference on Recommender Systems [1] (RecSys 2025[1]). While state-of-the-art

First International Workshop on Data Quality-Aware Multimodal Recommendation (DaQuaMRec @ RecSys 2025), September 22–26, 2025, Prague, Czech Republic
claudio.pomo@poliba.it (C. Pomo); dietmar.jannach@aau.at (D. Jannach); yubink.cs@gmail.com (Y. Kim); d.malitesta@gmail.com (D. Malitesta); alberto.mancino@poliba.it (A. C. M. Mancino); jmcauley@ucsd.edu (J. McAuley); a.melchiorre@criteo.com (A. B. Melchiorre); shah.nawaz@jku.at (S. Nawaz)
0000-0001-5206-3909 (C. Pomo); 0000-0002-4698-8507 (D. Jannach); 0000-0001-5033-2677 (Y. Kim); 0000-0003-2228-0333 (D. Malitesta); 0000-0001-8027-9475 (A. C. M. Mancino); 0000-0003-0955-7588 (J. McAuley); 0000-0003-1643-1166 (A. B. Melchiorre); 0000-0002-7715-4409 (S. Nawaz)

[1]https://recsys.acm.org/recsys25/.

multimodal recommendation models continue to evolve, their performance is tightly bound to the quality of the data they rely on. Noisy inputs [2, 3], missing modality information [4, 5], misaligned data [6, 7], and embedded biases [8, 9] can all degrade system performance and lead to inaccurate or unfair recommendations [10, 11]. In its first edition, the DaQuaMRec workshop offered a dedicated space to explore all such pressing issues. Its core mission was to foster deep, focused discussions and catalyze new research aimed at understanding, evaluating, and improving data quality in multimodal recommendation scenarios and settings. In comparison to similar recent workshops held at ACM Multimedia 2023 [12], SIGIR 2024 [13], and CIKM 2024 [14], as well as the paper session "Multimedia Recommendation" hosted at RecSys 2023[2], the DaQuaMRec workshop provided a careful and timely focus on the *quality* aspect of multimodal data in personalized recommendation.

In this preface, we summarize the main contributions and take-home messages presented, discussed, and published within the workshop proceedings. In the next sections, we first highlight the main objectives of the workshop and its topics of interest. Second, we provide a general overview on the workshop program as held within the RecSys 2025 venue. Then, we indicate all useful links for the dissemination, promotion, and collection of the workshop's materials, alongside the organizers and the program committee. Finally, we point the interested readers to a Special Issue on similar topics to those of DaQuaMRec, hosted in the journal "ACM Transactions on Recommender Systems" (ACM TORS).

2. Objectives

Multimodal recommender systems [15, 16, 17, 18, 19, 20] have taken over various domains in personalized recommendation, from fashion [21, 22, 23, 24] and music [25, 26] to food [27, 28] recommendation. By incorporating multiple data sources (called *modalities*) such as images, text, and audio, these models have demonstrated remarkable profiling capabilities, being able to learn more tailored user profiles and provide higher recommendation performance than previous methodologies leveraging single modalities [29, 23, 25, 30]. Alongside the latest advances in deep learning [31, 32] and increased multimodal data availability, multimodal recommendation has known an unprecedented growth, lately exploiting the representational power of large multimodal models [33, 34, 35]. Although effective, their progress may be still limited by a critical (and often disregarded) issue: ***data quality***[3].

Under this perspective, **the First International Workshop on Data Quality-Aware Multimodal Recommendation (DaQuaMRec)** aimed to discuss, analyze, and suggest possible research directions to address this largely-overlooked aspect. Indeed, the performance and reliability of multimodal recommender systems are highly dependent on the quality of the multimodal data they are trained on [36, 37]. Many data quality issues could arise in the context of multimodal recommendation. For instance, they include (but are not limited to) scenarios involving noisy data [3, 2, 38], incomplete/missing multimodal information [5, 4, 39, 40], but also modality misalignment [6, 7, 41] and biased information naturally encoded within multimodal recommendation data [8, 9, 42]. If not promptly recognized, analyzed, and solved, such data quality problems can lead to severe performance degradation in multimodal recommender systems, eventually resulting in suboptimal and potentially unfair personalized suggestions [10, 43, 44, 45].

Thus, DaQuaMRec was originally designed as an ad-hoc venue to discuss and propose solutions to the key challenges surrounding data quality. Its primary objective was to encourage focused discussions and promote research aimed at understanding, assessing, and improving data quality in multimodal recommendation. Rather than concentrating solely on the development of increasingly sophisticated models, the workshop took a step back and sought to examine the fundamental data-related issues that form the foundation of all multimodal recommendation efforts. In this respect, the venue featured talks regarding the fundamental role of multimodality in recommendation (the academic keynote), discussing the impact of data quality in multimodal recommendation within various contexts, such as fashion, tourism, and video recommendation. Additionally, it investigated the contribution of advanced models

[2] https://recsys.acm.org/recsys23/session-16/.
[3] https://www.dagstuhl.de/26281.

to address the data quality aspect, such as single-branch architectures (the invited talk) and intelligent agents for recommendation.

Conclusively, this forum should not only contribute to foster fruitful discussions around data quality problems under the academic perspective, but also through an industrial [46] lens, as discussed during the industrial keynote by Albatross AI and the discussion panel involving esteemed research scientists from Aampe, Spotify Sweden, and Google.

3. Topics of Interest

As stated in its call for papers, DaQuaMRec focused on works addressing the topics listed below, as well as other related ones:

- **Foundations of Data Quality in Multimodal Recommendation:** Theoretical frameworks, definitions, and metrics for assessing data quality in multimodal settings.
- **Detecting and Mitigating Noisy or Corrupted Multimodal Data:** Techniques for identifying and handling noise, outliers, and corrupted information in visual, textual, or other modalities.
- **Handling Missing or Incomplete Modalities:** Strategies for recommendation when one or more modalities are partially or entirely missing for certain items.
- **Addressing Modality Misalignment:** Methods to detect and correct semantic inconsistencies between different data streams (e.g., an image and its textual description).
- **Bias and Fairness in Multimodal Data:** Investigating, measuring, and mitigating societal biases (e.g., gender, race) present in multimodal datasets for recommendation.
- **Evaluation and Benchmarking of Data Quality:** Novel protocols, datasets, and benchmarks for evaluating the impact of data quality on recommendation performance.
- **Data-Centric and Human-in-the-Loop Approaches:** Systems and methodologies that prioritize data improvement, including crowdsourcing and active learning, to enhance recommendation quality.
- **Applications and Case Studies:** Real-world applications and in-depth case studies demonstrating the challenges and successes of managing data quality in live multimodal recommender systems.

4. Program

DaQuaMRec was held in Prague (Czech Republic) on September 22, 2025. The workshop featured the following paper presentations/discussions:

- **Keynote #1: Aixin Sun (NTU Singapore)**, title: "Multimodality in Recommender Systems: Does It Help, and Should We Expect an Answer?"
- **Keynote #2: Malte Lichtenberg (Albatross AI)**, title: "Sequential Recommenders and Multimodal Inputs: Mitigating Data Quality Issues in Industry-scale Recommenders"
- **Paper Session #1**
  - **G. Rippberger and J. Neidhardt**: "Comparative Analysis of Fashion Captioning for Multimodal Fashion Recommendation"
- **Invited Talk: Marta Moscati (JKU Linz)**, title: "Single-Branch Architectures for Recommendation"

- **Paper Session #2**
  - **Z. Wang, W. Höpken, and D. Jannach**, "Data Quality Challenges in Multimodal Tourism Recommender Systems"
  - **M. Valentini, A. Ferrara, and T. Di Noia**, "Exploring the Impact of Data Quality on Agentic Recommender Systems"
  - **E. Purificato**, "Inside the Frame: A Plan for Audio-Visual Feature Analysis of Video Recommendations for Children"
  - **S. Malani, Y. Zhang, L. Liu**, "Minimize Negative Experiences in Video Recommendation Systems with Multimodal Large Language Models" (invited from the main RecSys 2025 conference)
- **Discussion Panel**, panelists: Olivier Jeunen (Aampe), Henrik Lindström (Spotify), Suman Malani (Google, Inc)

5. Organizers and Program Committee

DaQuaMRec 2025 was organized by:

- **Claudio Pomo** (Politecnico di Bari) and **Daniele Malitesta** (Université Paris-Saclay): main contacts and organizers
- **Dietmar Jannach** (University of Klagenfurt), **Julian McAuley** (University of California), and **Shah Nawaz** (Institute of Computational Perception, Johannes Kepler University Linz): academic advisors
- **Yubin Kim** (Vody Inc.): industry advisor
- **Alberto Carlo Maria Mancino** (Politecnico di Bari): publicity and proceedings
- **Alessandro B. Melchiorre** (Criteo AI Lab): logistics

The program committee of the workshop was composed by: Aditya Chichani (Walmart), Marta Moscati (Johannes Kepler University Linz), Roger Zhe Li (Huawei), Salvatore Bufi (Politecnico di Bari), Alejandro Bellogin (Universidad Autonoma de Madrid), Bruno Sguerra (Deezer Research), Giandomenico Cornacchia (Amazon), Matteo Attimonelli (Politecnico di Bari), Tracy Holloway King (Adobe), Felice Antonio Merra (Cognism).

6. What's Next: Special Issue at ACM Transactions on Recommender Systems (ACM TORS)

Following on the success of this first edition of the DaQuaMRec workshop, it is planned to keep the discussion forum active in the future. The first objective is to repeat the event at other top-tier venues in the field. Additionally, on a more concrete level, a Special Issue on Challenges in Modern Multimodal Recommender Systems is now organized (as of February 2026) in the journal "ACM Transactions on Recommender Systems" (ACM TORS).

References

[1] M. Bieliková, P. Kordík, M. Schedl, M. de Gemmis, S. Pera, R. Alves, O. Jeunen, V. Ostuni (Eds.), Proceedings of the Nineteenth ACM Conference on Recommender Systems, RecSys 2025, Prague, Czech Republic, September 22-26, 2025, ACM, 2025.

[2] J. Zhang, Y. Zhu, Q. Liu, M. Zhang, S. Wu, L. Wang, Latent structure mining with contrastive modality fusion for multimedia recommendation, IEEE Trans. Knowl. Data Eng. 35 (2023) 9154–9167.
[3] P. Yu, Z. Tan, G. Lu, B. Bao, Multi-view graph convolutional network for multimedia recommendation, in: ACM Multimedia, ACM, 2023, pp. 6576–6585.
[4] C. Ganhör, M. Moscati, A. Hausberger, S. Nawaz, M. Schedl, A multimodal single-branch embedding network for recommendation in cold-start and missing modality scenarios, in: RecSys, ACM, 2024, pp. 380–390.
[5] D. Malitesta, E. Rossi, C. Pomo, T. D. Noia, F. D. Malliaros, Do we really need to drop items with missing modalities in multimodal recommendation?, in: CIKM, ACM, 2024, pp. 3943–3948.
[6] Y. Li, Q. Zhao, C. Lin, J. Su, Z. Zhang, Who to align with: Feedback-oriented multi-modal alignment in recommendation systems, in: SIGIR, ACM, 2024, pp. 667–676.
[7] G. Xv, X. Li, R. Xie, C. Lin, C. Liu, F. Xia, Z. Kang, L. Lin, Improving multi-modal recommender systems by denoising and aligning multi-modal content and user feedback, in: KDD, ACM, 2024, pp. 3645–3656.
[8] X. Liu, Z. Tao, J. Shao, L. Yang, X. Huang, Elimrec: Eliminating single-modal bias in multimedia recommendation, in: ACM Multimedia, ACM, 2022, pp. 687–695.
[9] D. Malitesta, G. Cornacchia, C. Pomo, T. D. Noia, On popularity bias of multimodal-aware recommender systems: A modalities-driven analysis, in: MMIR@MM, ACM, 2023, pp. 59–68.
[10] Y. Shang, C. Gao, J. Chen, D. Jin, Y. Li, Improving item-side fairness of multimodal recommendation via modality debiasing, in: WWW, ACM, 2024, pp. 4697–4705.
[11] J. Li, S. Wang, Q. Zhang, S. Yu, F. Chen, Generating with fairness: A modality-diffused counterfactual framework for incomplete multimodal recommendations, in: WWW, ACM, 2025, pp. 2787–2798.
[12] W. Ji, Y. Wei, Z. Zheng, H. Fei, T.-s. Chua, Deep multimodal learning for information retrieval, in: Proceedings of the 31st ACM International Conference on Multimedia, MM '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 9739–9741. URL: https://doi.org/10.1145/3581783.3610949. doi:10.1145/3581783.3610949.
[13] X. Zhu, A. Dhua, D. Gray, I. Z. Yalniz, T. Yu, M. Elhoseiny, B. Plummer, Multimodal representation and retrieval [mrr 2024], 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, 2024, p. 3047–3050. URL: https://doi.org/10.1145/3626772.3657987. doi:10.1145/3626772.3657987.
[14] A. Chichani, S. Kallumadi, T. H. King, A. Lopatenko, 1st workshop on multimodal search and recommendations (cikm mmsr '24), in: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, CIKM '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 5577–5579. URL: https://doi.org/10.1145/3627673.3679082. doi:10.1145/3627673.3679082.
[15] D. Malitesta, G. Cornacchia, C. Pomo, F. A. Merra, T. D. Noia, E. D. Sciascio, Formalizing multimedia recommendation through multimodal deep learning, Trans. Recomm. Syst. 3 (2025) 37:1–37:33.
[16] Q. Liu, J. Hu, Y. Xiao, X. Zhao, J. Gao, W. Wang, Q. Li, J. Tang, Multimodal recommender systems: A survey, ACM Comput. Surv. 57 (2025) 26:1–26:17.
[17] H. Zhou, X. Zhou, Z. Zeng, L. Zhang, Z. Shen, A comprehensive survey on multimodal recommender systems: Taxonomy, evaluation, and future directions, CoRR abs/2302.04473 (2023).
[18] J. Xu, Z. Chen, S. Yang, J. Li, W. Wang, X. Hu, S. Hoi, E. C. H. Ngai, A survey on multimodal recommender systems: Recent advances and future directions, CoRR abs/2502.15711 (2025).
[19] Q. Liu, J. Zhu, Y. Yang, Q. Dai, Z. Du, X. Wu, Z. Zhao, R. Zhang, Z. Dong, Multimodal pretraining, adaptation, and generation for recommendation: A survey, in: KDD, 2024.
[20] L. Pan, Z. Pan, F. Cai, H. Chen, Multimodal recommender systems: A survey of representation, modeling, and optimization, Inf. Fusion 128 (2026) 103991.
[21] Z. Liu, P. Luo, S. Qiu, X. Wang, X. Tang, Deepfashion: Powering robust clothes recognition and retrieval with rich annotations, in: CVPR, IEEE Computer Society, 2016, pp. 1096–1104.
[22] X. Han, Z. Wu, Y. Jiang, L. S. Davis, Learning fashion compatibility with bidirectional lstms, in:

ACM Multimedia, ACM, 2017.
[23] X. Chen, H. Chen, H. Xu, Y. Zhang, Y. Cao, Z. Qin, H. Zha, Personalized fashion recommendation with visual explanations based on multimodal attention network: Towards visually explainable recommendation, in: SIGIR, ACM, 2019, pp. 765–774.
[24] W. Kang, C. Fang, Z. Wang, J. J. McAuley, Visually-aware fashion recommendation and design with generative image models, in: ICDM, IEEE Computer Society, 2017, pp. 207–216.
[25] S. Oramas, O. Nieto, M. Sordo, X. Serra, A deep multimodal approach for cold-start music recommendation, in: DLRS@RecSys, ACM, 2017, pp. 32–37.
[26] A. van den Oord, S. Dieleman, B. Schrauwen, Deep content-based music recommendation, in: NIPS, 2013, pp. 2643–2651.
[27] W. Min, S. Jiang, R. C. Jain, Food recommendation: Framework, existing solutions, and challenges, IEEE Trans. Multim. 22 (2020) 2659–2671.
[28] W. Min, S. Jiang, J. Sang, H. Wang, X. Liu, L. Herranz, Being a supercook: Joint food attributes and multimodal content modeling for recipe retrieval and exploration, IEEE Trans. Multim. 19 (2017) 1100–1113.
[29] R. He, J. J. McAuley, VBPR: visual bayesian personalized ranking from implicit feedback, in: AAAI, AAAI Press, 2016, pp. 144–150.
[30] T. Baltrusaitis, C. Ahuja, L. Morency, Challenges and applications in multimodal machine learning, in: The Handbook of Multimodal-Multisensor Interfaces, Volume 2 (2), Association for Computing Machinery, 2018, pp. 17–48.
[31] X. Chen, L. Yao, J. J. McAuley, G. Zhou, X. Wang, Deep reinforcement learning in recommender systems: A survey and new perspectives, Knowl. Based Syst. 264 (2023) 110335.
[32] R. Nahta, G. S. Chauhan, Y. K. Meena, D. Gopalani, Deep learning with the generative models for recommender systems: A survey, Comput. Sci. Rev. 53 (2024) 100646.
[33] Z. Zhao, W. Fan, J. Li, Y. Liu, X. Mei, Y. Wang, Z. Wen, F. Wang, X. Zhao, J. Tang, Q. Li, Recommender systems in the era of large language models (llms), IEEE Trans. Knowl. Data Eng. 36 (2024) 6889–6907.
[34] J. Lin, X. Dai, Y. Xi, W. Liu, B. Chen, H. Zhang, Y. Liu, C. Wu, X. Li, C. Zhu, H. Guo, Y. Yu, R. Tang, W. Zhang, How can recommender systems benefit from large language models: A survey, ACM Trans. Inf. Syst. 43 (2025). URL: https://doi.org/10.1145/3678004. doi:10.1145/3678004.
[35] S. Geng, J. Tan, S. Liu, Z. Fu, Y. Zhang, VIP5: towards multimodal foundation models for recommendation, in: EMNLP (Findings), Association for Computational Linguistics, 2023, pp. 9606–9620.
[36] D. Zha, Z. P. Bhat, K. Lai, F. Yang, Z. Jiang, S. Zhong, X. B. Hu, Data-centric artificial intelligence: A survey, ACM Comput. Surv. 57 (2025) 129:1–129:42.
[37] Y. Wang, K. Ding, X. Liu, J. Kang, R. A. Rossi, T. Derr, Data quality-aware graph machine learning, in: CIKM, ACM, 2024, pp. 5534–5537.
[38] F. Mo, L. Xiao, Q. Song, X. Gao, E. Liang, Multimodal graph neural network for recommendation with dynamic de-redundancy and modality-guided feature de-noisy, CoRR abs/2411.01561 (2024).
[39] C. Ganhör, M. Moscati, A. Hausberger, S. Nawaz, M. Schedl, Single-branch network architectures to close the modality gap in multimodal recommendation, ACM Transactions on Recommender Systems (2025).
[40] H. Bai, L. Wu, M. Hou, M. Cai, Z. He, Y. Zhou, R. Hong, M. Wang, Multimodality invariant learning for multimedia-based new item recommendation, in: SIGIR, ACM, 2024, pp. 677–686.
[41] X. Luo, J. Cao, T. Sun, J. Yu, R. Huang, W. Yuan, H. Lin, Y. Zheng, S. Wang, Q. Hu, C. Qiu, J. Zhang, X. Zhang, Z. Yan, J. Zhang, S. Zhang, M. Wen, Z. Liu, K. Gai, G. Zhou, QARM: quantitative alignment multi-modal recommendation at kuaishou, CoRR abs/2411.11739 (2024).
[42] W. Yang, Z. Fang, T. Zhang, S. Wu, C. Lu, Modal-aware bias constrained contrastive learning for multimodal recommendation, in: ACM Multimedia, ACM, 2023, pp. 6369–6378.
[43] J. Li, S. Wang, Q. Zhang, S. Yu, F. Chen, Generating with fairness: A modality-diffused counterfactual framework for incomplete multimodal recommendations, CoRR abs/2501.11916 (2025).
[44] Z. Lin, Y. Tan, Y. Zhan, W. Liu, F. Wang, C. Chen, S. Wang, C. Yang, Contrastive intra- and inter-modality generation for enhancing incomplete multimedia recommendation, in: ACM Multimedia,

ACM, 2023, pp. 6234–6242.
[45] J. Yi, Z. Chen, Multi-modal variational graph auto-encoder for recommendation systems, IEEE Trans. Multim. 24 (2022) 1067–1079.
[46] A. Singh, T. Vu, N. Mehta, R. Keshavan, M. Sathiamoorthy, Y. Zheng, L. Hong, L. Heldt, L. Wei, D. Tandon, et al., Better generalization with semantic ids: A case study in ranking for recommendations, in: Proceedings of the 18th ACM Conference on Recommender Systems, 2024, pp. 1039–1044.