De Gruyter Brill

Advancing Scholarly Communication through Artificial Intelligence-Based Multimodal Modeling

GUEST EDITORS

Pit Pichappan
Digital Information Research Labs, India
pichappan@dirf.org

Simon James Fong
University of Macau, Macao
ccfong@um.edu.mo

SCOPE OF THE ISSUE

The rapid evolution of information technology necessitates a corresponding pace in the evolution of communication technologies, as well as an increasing number of research activities across various domains. Therefore, enormous amounts of knowledge are being disseminated across multiple domains, including the life sciences and social sciences, on a daily basis. The increasing complexity of scientific communication, caused by the heterogeneous nature of data and formats, has been identified as a characteristic of scholarly communication. The foremost characteristics of scholarly communication have been identified as being multimodal, semantically heterogeneous, distantly related to each other, dynamic, context-dependent, time-dependent, space-dependent, and highly complex. Any discussion on multimodal models for scholarly communication can sense the impact of AI. Multimodal AI mimics human understanding by analysing information from various perceptions.

DESCRIPTION

The relevance of that communication resembles the distributions that explain various observable phenomena; hence, we can pick up the patterns and inherent regularities that define those entities of information. Regularly, the search systems are proficiency-based in textual queries. In contrast, the slightly advanced multimodal ones provide a jointly embedding space of text, images, and other media for next-gen user experiences. This leads to greater user engagement. Multimodal AI facilitates more natural and intuitive interactions between humans and systems, while also presenting new challenges that necessitate intensive research.

The relevance of that communication resembles the distributions that explain various observable phenomena; hence, we can pick up the patterns and inherent regularities that define those entities of information. Regularly, the search systems are proficiency-based in textual queries. In contrast, the slightly advanced multimodal ones provide a jointly embedding space of text, images, and other media for next-gen user experiences. This leads to greater user engagement. Multimodal AI facilitates more natural and intuitive interactions between humans and systems, while also presenting new challenges that necessitate intensive research.

The same input queries yield different outputs on different platforms with varying semantics. The input queries do not have a rigorous match with the textual descriptors searched in the platform. The semantic variation raises the search pattern. To address these disadvantages, multimodal models were implemented, although they also have their downsides. The evaluation model cannot employ a human-in-the-loop approach due to the size of the corpus, and some outputs are too complex and demanding for humans to evaluate. It necessitates the AI intervention in the multimodal data processing.

As these models are further advanced, the systems can better detect concept embedding in the text. To further capitalize on gains in problem specificity through the multimodal approach, the established concept embedding is further enhanced using AI. Well-tuned and scaled multimodal generative-pretrained models can only ensure the context of that trained output. Effective multimodal models can take text, image, and video as inputs and outputs. When AI is infused, it will enhance human judgment and skills in handling multimodal data. AI processed, multimodal model development creates alternative solutions for integrating text and images in text-oriented systems.

We aim to document the core concepts, technologies, tools, and concerns, as well as the architectures, training strategies, application scenarios, datasets, evaluation benchmarks, and existing challenges of multimodal models influenced by AI, in the forthcoming special issue. The proposed special issue addresses not only the themes below.

- In context learning
- Multimodal data processing
- Context preservation
- Text-image interfaces
- Unified large multimodal models
- Text and visual coding and decoding
- Text to Image models
- Visual question answering
- Diffusion models
- Contrastive Learning
- Modals for aligning text descriptions with other forms, such as images
- Obfuscated embeddings
- Multimodal data augmentation
- Cross-model generation and integration
- Multimodal deep generative models
- Multimodal interfaces
- Open-source AI-based tools for multimodal data
- AI-driven multimodal models
- Issues in multimodal AI, such as misuse and ethics

HOW TO SUBMIT

Deadline: December 20, 2025
Estimated number of manuscripts: 12 papers

Before submission authors should carefully read the Instruction for Authors. Manuscripts have to be written in LATEX, AMS-TEX, AMS-LATEX. We do not accept papers in Plain TEX format. For an initial submission, the authors are strongly advised to upload their entire manuscript, including tables and figures, as a single PDF file. Authors are strongly advised to submit the final version of the paper using the journal's LaTex Template.

All submissions to the Special Issue must be made electronically via the Editorial Manager submission and tracking review system. All manuscripts will undergo the standard peer-review process (single-blind, at least two independent reviewers).

The deadline for submissions is December 20, 2025, but individual papers will be reviewed and published online on an ongoing basis. When the submission of individual papers is completed, the review process begins immediately.

IMPORTANT DATES

Submission of papers: December 20, 2025
First Round Review: January 25, 2026
Second Round Review: February 20, 2026
Notification of Acceptance: March 31, 2026

In case of any question please contact Joanna Kosińska,
Managing Editor of Open Information Science at Joanna.Kosinska@degruyterbrill.com.