

Dnotitia Paper on Korean Cultural Dataset Accepted to ACM Multimedia 2025 Dataset Track

Dnotitia's HAN dataset, reflecting Korea's cultural heritage, was accepted to ACM Multimedia 2025, marking progress in inclusive, multicultural AI.

SEOUL, SOUTH KOREA, August 29, 2025 /EINPresswire.com/ -- [Dnotitia Inc.](#) (Dnotitia), a specialized company in long-term memory AI and semiconductor-integrated solutions, announced that the research project it participated in, focused on building a multimodal dataset reflecting Korean cultural heritage, has been accepted to the Dataset Track of [ACM Multimedia 2025](#), one of the world's most authoritative conferences in the multimedia field.

Since its inception in 1993, ACM Multimedia has grown into one of the world's premier conferences in multimedia, covering multimodal AI and next-generation media

technologies. The conference addresses the full spectrum of research across image, video, speech, and text processing and integration. Each year, thousands of papers are submitted, but only a small fraction are accepted, reflecting the event's highly competitive nature. The 33rd edition will be take place from October 27 – 31, 2025. in Dublin, Ireland.

The accepted paper introduces HAN (Heritage Augmented Narrative Visual-Language Description Dataset), a multimodal dataset designed to reflect Korea's cultural heritage and linguistic nuances. Going beyond simple translation, HAN adopts narrative-style captions that capture emotional context, social interactions, and cultural storytelling, helping AI systems



Dnotitia logo

ACM multimedia



Dublin, Ireland **27-31.10.2025**

ACM Multimedia logo

overcome performance limitations of existing multimodal models and generalize across diverse, multicultural environments

The HAN dataset goes beyond simple translation, adopting narrative-style captions that capture emotional context, social interactions, cultural background, and storytelling. This approach helps overcome the performance limitations of existing multimodal models, while alleviating bias issues common in large-scale image-text training and supporting the development of vision-language models that can generalize across multicultural environments.

The study is characterized by the systematic construction of a dataset extracted from 7,822 Korean broadcast programs, including 41,000 images and 410,000 Korean-English narrative captions. A key achievement is overcoming the limitations of English-centric datasets, thereby correcting linguistic imbalances and cultural bias while making multilingual and multicultural AI learning possible, including for low-resource languages such as Korean. In addition, by offering a scalable alternative to the traditionally high-cost, expert-dependent process of building cultural and language-based datasets, the project represents a significant breakthrough for future research.

In particular, by applying a narrative-style captioning approach, HAN richly captures the contextual meaning of cultural heritage, an innovation that has drawn attention from both academia and industry for combining innovative ideas with demonstrated practical value.

Furthermore, to validate the dataset's effectiveness, the research team conducted follow-up experiments using the diversity of narrative captions, which resulted in significant performance improvements over existing models. This confirmed that the HAN dataset is not just a data collection effort but a resource with practical value for both academic research and real-world applications – providing strong evidence of its broader impact.

As a foundational dataset, HAN is expected to contribute significantly to the global AI research ecosystem, with potential applications across multimodal AI, natural language processing, and digital archiving of cultural heritage.

"Just as K-pop and K-dramas have become part of daily life around the world, the time has come for AI models to also embody Korean culture," said Moo-Kyoung Chung, CEO of Dnotitia. "The HAN dataset is more than a research outcome, it represents a first step in allowing Korean culture to permeate global AI models, and will play a key role in ensuring data diversity and reducing bias across the AI ecosystem."

This achievement is part of the "Korean Cultural Video Understanding Dataset" project, supported by Korea's Ministry of Science and ICT (MSIT) and the National Information Society Agency (NIA). Building on this milestone, Dnotitia plans to further expand its efforts by developing multimodal AI datasets that reflect Korean cultural heritage and linguistic diversity. Their company aims not only to drive technical innovation but also to foster a more inclusive and

equitable AI ecosystem

Dnotitia Communications

Dnotitia, Inc.

dnpr@dnotitia.com

Visit us on social media:

[LinkedIn](#)

[Facebook](#)

[YouTube](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/843860270>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2025 Newsmatics Inc. All Right Reserved.