Domen Mongus
University of Maribor, Faculty of Electrical Engineering and Computer Science
domen.mongus@um.si

Mitko Nikov
University of Maribor, Faculty of Electrical Engineering and Computer Science

Mitja Žalik
University of Maribor, Faculty of Electrical Engineering and Computer Science

Matej Brumen
University of Maribor, Faculty of Electrical Engineering and Computer Science

Use of Data Spaces for Humanities and Cultural Heritage Research in the AID HCH Project

Although data spaces have become a recognised foundational technology for data sharing and the integration of distributed data sources in recent years, they still often get confused with traditional cloud systems and data centres. While the latter are mainly maintained as silos for storing and distributing data through application programming interfaces (APIs), data spaces merely hold the addresses of the APIs in a central catalogue, thus enabling access to multiple cloud systems concurrently through centralised services (Bacco et al., 2024).

In the research project 'AID HCH – Breakthrough in humanities and cultural heritage with artificial intelligence,' we address the development of a prototype data space to support the introduction of new artificial-intelligence concepts in the research of cultural heritage and the humanities. We primarily focus on developing a federated data catalogue and implementing services that will allow for simultaneously searching for data over multiple sources, fusing them, and training artificial-intelligence models that can be used for (i) studies of Slovene identity and the role of women, (ii) heritage narration, and (iii) preventive conservation.

The first stage of the development of the central data catalogue is focused on integrating spatio-temporal data that support all three targeted studies. This includes the registration of geographic maps, remote-sensing data (e.g. LiDAR), sensor data, and other documents that we geo-reference to selected spatial entities (e.g. newspapers to places), thereby creating a comprehensive picture of space and time. The observed spatial entities are then enriched with information obtained by processing individual data sources using trained artificial-intelligence models. The presented methodology has been integrated into a user-friendly node editor, as shown in Figure 1, which allows us to adapt the pipeline easily to different use cases (Lens, 2023).

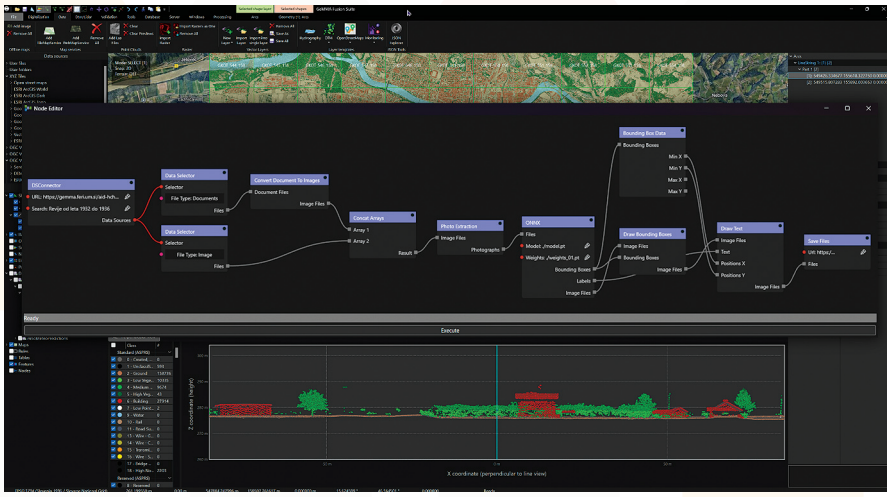

Figure 1 Graph-based pipeline for analysis of scanned documents and photographs built in a custom node editor

The validation of the presented methodology was done by investigating archived images and identifying samples that contain men, women or both, as presented in Figure 2. The preliminary results indicate that over 92% recall and precision can be achieved, thus showing the potential for upgrading the system into an actual semantic search engine.

3rd Annual Conference of the Slovenian Node of the European Research Infrastructure for Heritage Science
E-RIHS Slovenia
Koper, 18–19 November 2025

https://doi.org/10.26493/978-961-293-546-7.3

Figure 2 Detection of people in archived film photographs

Acknowledgements

We acknowledge financial support provided by the Slovenian Research and Innovation Agency (Research Project no. J7-60128: AID HCH – Breakthrough in humanities and cultural heritage with artificial intelligence, as the main funding source, with additional support from Research Programme Core Funding No. P2-0041).

References

Bacco, M., Kocian, A., Chessa, S., Crivello, A., & Barsocchi, P. (2024). What are data spaces? Systematic survey and future outlook. *Data in Brief, 57,* 110969.

Lens, S. (2023). *Procedural 3D modeling using geometry nodes in Blender: Discover the professional usage of geometry nodes and develop a creative approach to a node-based workflow*. Packt Publishing.