Proceedings of the Third International Workshop of Semantic Digital Humanities co-located with the Extended Semantic Web Conference 2026

Oleksandra Bruns[1,2], Arianna Graciotti[3], Bruno Sartini[4], Lise Stork[5] and Tabea Tietz[1,2]

[1]FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany
[2]Karlsruhe Institute of Technology (AIFB), Kaiserstr. 89, 76133 Karlsruhe, Germany
[3]Center for Language and Cognition, University of Groningen, 9712 EK Groningen, the Netherlands
[4]y,digital, Zeist, the Netherlands
[5]Informatics Institute, University of Amsterdam, LAB42, Science Park 9001098 XH Amsterdam, the Netherlands

Abstract

The International Workshop on Semantic Digital Humanities (SemDH2026) brought together the Semantic Web, Cultural Heritage, and Digital Humanities communities to discuss how Semantic Web and Artificial Intelligence technologies can support the representation and enrichment of cultural knowledge. The accepted papers addressed topics such as semantic infrastructures, information extraction from historical sources, and AI-supported cultural heritage applications, with recurring themes including interoperability, provenance, explainability, and trust in knowledge representation. The workshop also featured a keynote by Prof. Dr Francesca Tomasi and a community discussion on the role of Semantic Web for DH in the era of Large Language Models.

Keywords

Digital Humanities, Knowledge Graphs, Knowledge Representation, Large Language Models, Cultural Heritage, Ontologies, FAIR, Uncertainty

1. Organizing Committee

- Oleksandra Bruns, FIZ Karlsruhe and Karlsruhe Institute of Technology, Germany
- Arianna Graciotti, LILEC, University of Bologna, Italy
- Bruno Sartini, y.digital, The Netherlands
- Lise Stork, University of Amsterdam, The Netherlands
- Tabea Tietz, FIZ Karlsruhe and Karlsruhe Institute of Technology, Germany

2. Steering Committee

- Enrico Daga, Knowledge Media Institute (United Kingdom)
- Catherine Faron, Université Côte d'Azur (France)
- Laura Hollink, Centrum Wiskunde Informatica (The Netherlands)
- Eero Hyvonen, Aalto University (Finland)
- Mareike König, German Historical Institute Paris (France)
- Valentina Presutti, University of Bologna (Italy)
- Harald Sack, FIZ Karlsruhe and Karlsruhe Institute of Technology (Germany)

SemDH2026: Third International Workshop of Semantic Digital Humanities. Co-located with ESWC 2026, May 10, 2026, Dubrovnik, Croatia.

oleksandra.bruns@fiz-karlsruhe.de (O. Bruns); a.graciotti@rug.nl (A. Graciotti); bruno@y.digital (B. Sartini); b.sartini@lmu.de (L. Stork); tabea.tietz@fiz-karlsruhe.de (T. Tietz)
https://www.fiz-karlsruhe.de/de/bereiche/lebenslauf-und-publikationen-oleksandra-bruns (O. Bruns); https://www.fiz-karlsruhe.de/en/bereiche/lebenslauf-und-publikationen-tabea-tietz (T. Tietz)
0000-0002-8501-6700 (O. Bruns); 0009-0004-7918-809X (A. Graciotti); 0000-0002-9152-4402 (B. Sartini); 0000-0002-9152-4402 (L. Stork); 0000-0002-1648-1684 (T. Tietz)

- Stefan Schlobach, Vrije Universiteit Amsterdam (The Netherlands)
- Torsten Schrade, Academy of Sciences and Literature Mainz (Germany)
- Francesca Tomasi, University of Bologna (Italy)

3. Programme Committee

- Victor de Boer — Vrije Universiteit Amsterdam
- Stefano De Giorgis — Vrije Universiteit Amsterdam
- Nicolas Lazzari — University of Bologna
- Mathias Zinnen — Friedrich Alexander Universität Erlangen-Nürnberg
- Heikki Rantala — Aalto University
- Luiz Do Valle Miranda — Jagiellonian University
- Annastiina Ahola — Semantic Computing Research Group (SeCo), Aalto University
- Vandana Jha — Karlsruhe Institute of Technology
- Federico Pianzola — University of Groningen
- Catherine Faron — Université Côte d'Azur
- Rossana Damiano — Dipartimento di Informatica, Università di Torino
- Sarah Binta Alam Shoilee — Vrije Universiteit Amsterdam
- Sofia Baroncini — Leibniz Institut für Europäische Geschichte (IEG)
- Andre Valdestilhas — Vrije Universiteit Amsterdam
- Andreas Weber — University of Twente
- Olivier Bruneau — University of Lorraine - LHSP - Archives Poincaré
- Emmanuel Nauer — LORIA

4. Preface

The SemDH workshop serves as a space for bridging the gap between the Semantic Web and the Digital Humanities (DH) and Cultural Heritage (CH) communities. This year's edition of the workshop, SemDH2026, presented a variety of novel works that explored how Linked Open Data (LOD) addresses the challenges of representing, preserving, and enhancing cultural knowledge. There were 13 papers submitted for peer-review to this workshop. One paper was desk-rejected, and the remaining 12 were reviewed each by at least 3 members of the programme committee. After the reviews, 9 papers were accepted for this volume, 5 as full research papers, and 4 as short or position papers.

The papers explored a broad spectrum of topics at the intersection of Semantic Web technologies, CH, and DH. Several contributions focused on semantic infrastructures and knowledge graphs (KGs) for DH purposes, e.g., ontology design patterns for archaeological research, cross-domain CH discovery, knowledge graphs for performing arts and migration studies, and methods for linking historical persons to archival documents. Other papers investigated data extraction and enrichment approaches, such as information extraction from medieval Latin sources, and multi-modal narrative generation using KGs and Large Language Models (LLMs).

A common topic across many papers was the challenge of making complex CH and DH data more FAIR – findable, accessible, interoperable, and reusable – for both researchers and other users. Since users of the CH and DH KGs are mainly the representatives of GLAM institutions (galleries, libraries, archives and museums), many papers emphasised the importance of explainability, visualisation, and user-centred design, especially in complex contexts involving uncertainty and interpretation. Many papers presented LLM-based solutions for some of the tasks and highlighted both the benefits and the challenges that arise when integrating LLMs into Semantic Web and DH/CH workflows. This demonstrated how Semantic Web technologies and traditional Artificial Intelligence (AI) methods increasingly complement each other in order to achieve structured, meaningful and trustworthy representation of complex CH/DH knowledge while exploiting the ease-of-use and flexibility of LLMs.

This year's keynote by Prof. Dr. Francesca Tomasi directly addressed the role of Semantic Web technologies in the era of LLMs, arguing that ontologies, provenance models, and KGs remain essential for ensuring transparency, interpretability, and trustworthiness in AI-driven CH infrastructures. The community Question and Answer (Q&A) session, which followed the keynote, aimed at collecting workshop audience opinions and experiences on the same topic. The main insights from the keynote and the subsequent community discussion are summarised in Sections 6 and 7.

Taking into consideration both the initial reviews, the presentations and resulting discussions, the SemDH organizing committee decided to give the best paper award to *Cultural Heritage Survey: an Ontology Design Pattern* by Silvia Cappa et al. Finally, the SemDH Workshop was awarded the ESWC 2026 Dieter Fensel Visionary Contribution Award.

May 2026 Oleksandra Bruns, Arianna Graciotti, Bruno Sartini, Lise Stork and Tabea Tietz

5. Paper Presentations

Session I

- Linking Historical Persons to Archival Documents: Challenges and Approaches (*full paper*)
  Huan Chen, Gareth Jones, Declan O'Sullivan, Eamonn Kenny, Alex Randles, Neil Johnston and Rob Brennan
- Museum Monitoring: an environmental monitoring dataspace using The Things Network, Solid, and LDES (*full paper*)
  Arthur Vercruysse, Ben De Meester, Julián Rojas and Dieter Suls

Session II

- Navigating Babel: A Mid-Level Ontology for Cross-Domain Cultural Heritage Discovery (*full paper*)
  Mary Ann Tan, Genet Asefa Gesese and Harald Sack
- Generation of Multi-Modal Narratives of Cultural Objects from Knowledge Graphs and LLMs (*full paper*)
  Ishak Riali, Raphaël Gersen, Elizabeth Rodriguez Estrada, Gabriel Spautz Vieira and Martin Berger
- Cultural Heritage Survey: an Ontology Design Pattern (*full paper*)
  Silvia Cappa, Alessandra Chirivì, Hüseyin Erdoğan, Maria Chiara Frangipane, Matteo Greco, Giorgia Lodi, Andrea Giovanni Nuzzolese, Valentina Presutti and Cristiano Putzolu

Session III

- Interpretable Uncertainty in Colonial Collection Research: User-Informed Requirements and a Lightweight Modelling Pattern for Semantic Infrastructures (*short paper*)
  Deborah Ehlers and Philipp Uesbeck
- Schema-Driven Information Extraction from Medieval Latin Regesten: A Four-Way Evaluation of GLiNER2 for Ontology Population (*short paper*)
  Luana Moraes Costa, Bärbel Kröger and Christian Popp
- Making Semantic Data Accessible: A Human-Centered Approach to CIDOC CRM Visualization (*short paper*)
  Philipp Uesbeck and Deborah Ehlers
- Towards the Theatre Migrants Knowledge Graph (*short paper*)
  Jorit Hopp, Berenika Szymanski-Düll, Yan Lin and Daniel Hernández

6. Keynote: Do Semantic Web technologies still matter?

Francesca Tomasi

In the current landscape shaped by the rapid diffusion of Large Language Models (LLMs), the role of Semantic Web methodologies and technologies deserves renewed attention. While LLMs demonstrate remarkable capabilities in processing and generating natural language, they also highlight the importance of structured and semantically explicit knowledge infrastructures capable of ensuring transparency, reliability, and traceability of information. Within the CH domain, Semantic Web approaches provide a framework for enriching descriptive metadata through the reuse of shared conceptual models and ontologies. Models as "Digital hemeneutics" [1] allow institutions to represent not only the factual description of objects but also the processes through which knowledge about those objects has been produced, interpreted, and revised over time. Traditional archival and bibliographic descriptions, but also digital scholarly editions and full-text document collections, generally focus on recording objective information: authorship, dates, places, or institutional contexts. However, this descriptive layer often omits the interpretative dimension that underlies many of these assertions. The transition from traditional description to semantic models requires the explicit representation of contextual and interpretative information in order to produce reliable and trustworthy data.

In this perspective, RDF vocabularies should not only describe objects but also document the intellectual processes through which knowledge about those objects is constructed. This is particularly important in cases characterized by uncertainty or scholarly debate, such as disputed authorship, uncertain dating, or conflicting historical interpretations. Knowledge graphs offer a powerful solution for representing these interpretative layers. Instead of describing a resource through a single record, a KG models it as a network of entities, events, and assertions, each associated with its provenance and scholarly context. This narrative dimension is particularly relevant in the era of AI-driven knowledge systems. As LLMs increasingly rely on structured data sources to support reasoning and retrieval, the availability of transparent, provenance-aware KGs becomes essential. Enriching, both manually and automatically, metadata with interpretative and contextual information not only increases users' trust but also strengthens the epistemic reliability of CH data infrastructures. Ultimately, Semantic Web technologies provide the conceptual and technical framework necessary to represent the dynamic and interpretative nature of CH knowledge. By combining ontologies, provenance models, and KGs, archives and libraries can move beyond static descriptive records and toward richer representations that capture both the history of objects and the history of their interpretations.

7. Community Q&A: Do Semantic Web Technologies still matter?

The keynote was followed by an interactive community Q&A session aimed at collecting perspectives from workshop participants regarding the current role of Semantic Web technologies in DH and CH, particularly in relation to the increasing adoption of Large Language Models (LLMs).

The Q&A session consisted of six questions combining both open-ended prompts and questions with predefined answer variants. Following each question, participants were encouraged to elaborate on their responses and share their perspectives with the audience, allowing a question to serve as a starting point for broader community discussion. In particular, participants were asked:

- Which cultural heritage / DH sector do you primarily work in?
- What is the biggest obstacle to adopting Semantic Web technologies in DH?
- In one or two words, what is the biggest strength of Semantic Web technologies for DH?
- Has your work resulted in products (ontologies, KGs, tools, etc.) that were adopted by a CH institution?
- How often do you use Semantic Web technologies in your work?
- How often do you use LLMs in your work?

23 workshop participants contributed to the discussion and questionnaire. The participants represented a broad range of backgrounds, including art history, archives, archaeology, museums, libraries, musicology, ethnography, Human-Computer Interaction, knowledge organisation, and digital history, which directly reflects the interdisciplinary nature of the SemDH community.

One of the clearest findings was the strong agreement regarding the main strengths of Semantic Web technologies in DH and CH contexts. Participants most frequently highlighted interoperability and provenance, alongside related concepts such as explainability, traceability, explicit structure, and trustworthiness (see Figure 1).

Figure 1: Word cloud summarising participants' responses to the question "In one or two words, what is the biggest strength of Semantic Web technologies for DH?" Generated with Mentimeter.

At the same time, participants identified the complexity of Semantic Web technologies as the most significant obstacle to wider adoption within DH, see Figure 2. Other major barriers included lack of expertise, insufficient institutional support, sustainability concerns, and limited resources.

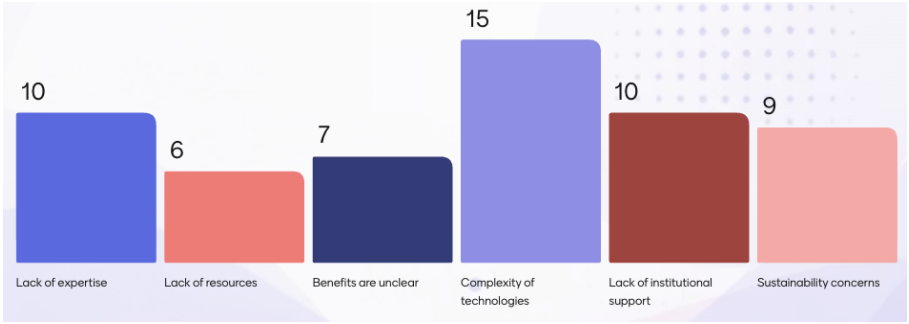

Figure 2: Participants' responses to the question "What is the biggest obstacle to adopting Semantic Web technologies in DH?" Generated with Mentimeter.

The results also revealed an interesting relationship between Semantic Web technologies and LLMs in everyday research practice. Participants reported using LLMs more frequently (14 participants use LLMs on a daily basis) than Semantic Web technologies (8 participants use Semantic Web technologies on a daily basis). The subsequent discussion explored the concrete ways in which participants currently employ LLMs in their work. Common examples included translation tasks, brainstorming and refining research ideas, drafting and summarising texts, and assisting with information search. These examples led to a broader reflection on the role of AI systems in creative processes. This sparked a discussion on the ethical use of LLMs in academic workflows. One participant explicitly raised concerns about whether extensive dependence on AI systems could negatively affect researchers' own critical thinking and creativity.

Finally, the responses concerning institutional adoption revealed a clear divide. While 7 participants

reported that their ontologies, KGs, or tools had already been integrated into institutional workflows within GLAM organisations, another 7 participants stated that their work remained at the level of theory, experimentation, or prototyping.

Overall, SemDH2026 once again demonstrated the importance of Semantic Web technologies within the DH and CH domains, while also highlighting the impact of AI and LLMs within the communities. The workshop brought together researchers from diverse disciplinary backgrounds, fostering lively discussions. During the workshop, including paper presentations, the keynote and the community discussion, a shared understanding emerged that Semantic Web technologies and AI methods should be viewed as complementary approaches for achieving broader FAIR and trustworthy representation of CH and DH knowledge.

References

[1] M. Daquino, V. Pasqual, F. Tomasi, Knowledge representation of digital hermeneutics of archival and literary sources, JLIS: Italian Journal of Library, Archives and Information Science= Rivista italiana di biblioteconomia, archivistica e scienza dell'informazione: 11, 3, 2020 (2020) 59–76.