Marburg, 12–13 December 2024

Artificial Intelligence in Archives and Collections:
Practices, Potentials, and Evidence Production in Dealing with Images and Multimodal Cultural Heritage

Leibniz
Value
of the Past

**Hybrid format: in-person and online,
Conference languages: German and English**

Dates and Conference Venue:

Thursday, December 12th, 12:30 – 19:30,
Friday December 13th, 9:00 – 15:45
Herder Institute for Historical Research on East
Central Europe – Institute of the Leibniz Association,
Gisonenweg 5 – 7, 35037 Marburg

The Research Lab 1.3. "Digital Heuristics and Digital History" of the Leibniz Research Alliance "Value of the Past" is organising a conference in Marburg on artificial intelligence in heritage institutions, such as archives and collections, and how these new technologies are transforming archival institutional practices. Topics will primarily focus on – but will not be limited to – visual sources, such as photography and graphic collections or those with mixed image-text sources and multimodal information processing.

The conference will provide a forum for researchers and practitioners from the humanities, archives and collections to connect with researchers and engineers in the fields of artificial intelligence, computer science and the digital humanities, to discuss new findings, and to exchange experiences. The event seeks to promote an interdisciplinary, cross-sectoral dialogue between research, development and practice.

We cordially invite stakeholders from all areas in the field. Attendance without contribution to the programme is welcome. Speakers from Austria, China, France, Germany, USA, UK and Sweden will talk about topics including: Exploring and Analysing Collections from Textual and Multimodal Contexts, Computer Vision: Semantic Segmentation, Classification, Analysing and

Understanding Images, The Effects of Computational Methods on Image Analytical Research and Visual Studies in Cultural Heritage, Opportunities and Challenges for the Automated Indexing and Cataloguing of Visual Sources in Archives and Collections.

Key Topic: Dealing with Images and Multimodal Cultural Heritage

In the years to come, curatorial and archival processes in memory and heritage institutions (including key aspects of cataloguing such as description, classification and categorization) will be increasingly supported by automated systems and artificial intelligence. These practices attribute value to sources and link archival materials and collection objects to societal narratives. Collections and archives thus form an essential basis for memory-related discourses and shape our view of the past.

New technologies have now reached the stage where they are potentially suitable for the requirements of cultural heritage institutions. There is potential promise in the partially automated indexing and cataloguing of historical sources, and particularly of digital images, the semantics and meaning of which have until recently only been accessible to the human eye and not

the machine. Images can now be automatically described in semantic terms and therefore made findable. AI methods – and most recently multimodal AI processing – are opening up new possibilities for automatic text and layout recognition, automated image annotation, and the analysis of visual sources and their contextualisation (e.g. text-image combinations or audio-visual sources).

At the same time, however, there is a lack of knowledge about how evidence – and facts – are generated and how AI processes affect the attribution of authenticity to archival documents and photos. At this moment, the humanities lack semantically high-quality and subject-appropriate training data sets. Similarly, there is hardly any agreement about what constitutes an acceptable outcome of computational classification processes, or what benchmarks should be used to evaluate the results. We therefore need to develop best practices (that may tap into explainable Artificial Intelligence (XAI) methods), benchmarks and goals. To increase the value that the information and knowledge held in our archives and collections have for future research, we need to deepen our understanding of processes and algorithms.

The conference will bring together curatorial and archiving knowledge and new AI-based methods, and will also provide a forum for ethical reflections on the use of AI in academic and archival practices. We will consider how automated processing and AI methods require detailed epistemic reflection and methodological-technical control to ensure that no false or tainted evidence is generated. Thus, the conference will discuss the effects of these new technologies on the production of evidence, thereby contributing to the crucial question of how the digital transformation is changing knowledge creation in the humanities and what this means for scholarship in historical disciplines.

Conference participation: Registration

For registration and further information please consult the conference website: **https://eveeno.com/441936742**

The detailed conference programme will be published here and updated regularly.

Most Papers will be presented in English.
Discussions can be in English and German.

Participation in the conference is free of charge.
Kindly refer to the conference website for information on possible costs for meals and the social programme.

Deadline for Applications:
please register as early as possible, since it will help us in organizing the event. Hotels are scarce in Marburg. Please reserve a room in good time of you want to take part in person.

Stipendia – Support for travel costs

We can assist selected participants with a stipendium towards travel costs and accommodation in Marburg. Please send a letter of explaining your situation of up to 500 words and a CV with your application. The letter should tell us a little about your experience in the topic, what you expect to gain from the event, and what you would like to contribute to the discussion.

Contact for Programme and Stipendium

archivesai@herder-institut.de

Organizers:

- Leibniz-Forschungsverbund „Wert der Vergangenheit", Lab 1.3. Digitale Heuristik und Historik
- Leibniz-Zentrum für Archäologie (LEIZA), Mainz
- Herder-Institut für historische Ostmitteleuropaforschung – Institut der Leibniz-Gemeinschaft, Marburg
- NFDI4Memory, Task Area Data Quality

Committee:

Elke Bauer, Simon Donig, Annette Frey, Dominik Kimmel

PROGRAMME (all presentations can be attended online)

Thursday, 12th December 2024

12.00 pm   Warm up

12.30 pm   **Welcome & Introduction**

Peter Haslinger (Herder-Institut für historische Ostmitteleuropaforschung – Institut der Leibniz-Gemeinschaft)

Achim Saupe (Leibniz-Forschungsverbund Wert der Vergangenheit)

Annette Frey (Leibniz-Zentrum für Archäologie)

Simon Donig (NFDI4Memory)

Dominik Kimmel (Leibniz-Zentrum für Archäologie)

Keynote 1

12.45 pm   Tayler Arnold, University of Richmond (online)
Explainable and Auditable Search and Discovery of Visual Cultural Heritage Collections

Panel 1   **Exploring and Analysing Collections from Textual and Multimodal Contexts**

1.30 pm   Günther Mühlberger, Universität Innsbruck
AI as an employee. How the digital transformation is (not only) changing archives

2.00 pm   Frank Puppe, Universität Würzburg
Pipeline for Digital Indexing of Archaeological Record Cards

2.30 pm   **Coffee Break**

2.45 pm   Mahsa Vafaie, FIZ Karlsruhe – Leibniz Institute for Information Infrastructure
Separation of machine-printed and handwritten text in archival documents

3.15 pm   Markus Huff, Leibniz-Institut für Wissensmedien, Tübingen
ArchiveGPT: Psychological and technological perspectives on the AI-supported archiving of image material

3.45 pm   Elisabeth Mödden, Deutsche Nationalbibliothek, Frankfurt a. M.
Automatic Indexing with Large Scale Vocabulary – Automatic Subject Indexing as Extreme Multi Label Learning Problem (Presentation in German)

4.15 pm   **Coffee break**

Panel 2   **Computer Vision: Semantic Segmentation, Analysing and Understanding Images**

4.45 pm   Ralph Ewert, TIB – Leibniz Information Centre for Science and Technology / Leibniz University Hannover
Unlocking Cultural Heritage: Computer Vision for Art and History Archives

5.15 pm   Karsten Tolle, Goethe University Frankfurt am Main
Potpourri of Computer Vision in Cultural Heritage

5.45 pm   N.N.
State of the Art Methods in Computer Vision

6.15 pm   **Evening Reception**

Friday, 13th December

9.00 am   Warm up
Continuation of Panel 2: Computer Vision: Semantic Segmentation, Analysing and Understanding Images

9.30 am   Angelica I. Aviles-Rivero, Tsinghua University, Beijing (online)
Dusting Off the Unlabelled Data: Graph Semi Supervised Learning for Large-Scale Datasets

10.00 am   Erik Radisch, Leibniz Institut für Länderkunde, Leipzig
A new Approach to Semi-Automated Annotations with Segment-Anything (Meta AI)

10.30 am   Yury Korolev, University of Bath
Image filtering based on total variation spectral decompositions: an overview of the method and an application in medieval paper analysis

11.00 am   **Coffee Break**

Panel 3   **The Effects of Computational Methods on Image Analytical Research and Visual Studies in Cultural Heritage**

11.15 am   Carola-Bibiane Schönlieb, University of Cambridge (online)
Unveiling the invisible – mathematical imaging for cultural heritage

11.45 am   Christopher Kermorvant, TEKLIA, Paris
How to use controlled vocabularies to describe early Japanese photographs using deep learning models

12.15 pm   Peter Bell, Philipps Universität Marburg
Ways of Seeing in AI and art history

12.45 pm   **Lunch Break**

Keynote 2

1.30 pm   James Evans, The University of Chicago (online)
The Geometry of Culture: Analyzing meaning through embeddings of text and images

Panel 4   **Opportunities and Challenges for the Automated Indexing and Cataloguing of Visual Sources in Archives and Collections**

2.15 pm   Nicole Graf, ETH Zürich (online)
Between ritual and relief – when the computer squints: analysis of AI-based indexing in the Image Archive of the ETH Library

2.45 pm   Larissa von Bychelberg, Uppsala University
Research at the Intersection of Archives, AI, and Authenticity. Can Artificial Intelligence contribute to an inclusive and sustainable assessment of archival records' authenticity?

3.15 pm   **Plenary Discussion**

3.45 pm   **End of the conference**

Programme subject to change