

Workshop of EHRI-CZ, GWZO and the Malach Center for Visual History: New Readings of Holocaust Testimony

Venue: Institute of Czech Literature of the Czech Academy of Sciences (Na Florenci 1420/3, Prague 1)

Date: 20– 21 October 2025

Programme:

Monday, 20 October 2025:

9:30 – 10:00 Registration of participants

10:00 – 10:20 Introduction

- Magdalena Sedlická (EHRI-CZ, Masaryk Institute and Archives of the Czech Academy of Sciences)
- Karin Roginer Hofmeister (Malach Center for Visual History)
- Michal Frankl (Leibniz Institute for the History and Culture of Eastern Europe)

10:20 – 11:00 Testimonies into Data: Histories, Questions and Challenges

Chair: Jiří Kocián (Malach Center for Visual History)

- Michal Frankl (Leibniz Institute for the History and Culture of Eastern Europe): “Datafying” Early Holocaust Testimony
- Jakub Mlynář (HES-SO University of Applied Sciences and Arts Western Switzerland): Metadata and Antedata: Lived Experience as Computable Objects

11:00 - 11:30 Coffee Break

11:30 - 13:00 New Testimonies: Quick Introductions

Chair: Karin Roginer Hofmeister (Malach Center for Visual History)

- Laura Miñano (University of Valencia): Challenging the Archive with Silence: Liminal Voices in Holocaust Survivors’ Oral Histories
- Stefania Zezza (University of Rome Tor Vergata): Revisiting David Boder’s Interviews: Linguistic Markers of Trauma and Deculturation in Early Holocaust Testimonies
- Radana Rutová - Michaela Dostálová (Terezín Memorial): Vzpomínky jako zdroj poznání každodennosti v Terezíně (Memories as a Source for Understanding Everyday Life in Terezín)
- Anat Bratman-Elhalel (Ghetto Fighters' House Museum): The Haim Gouri Testimonies Project
- Helena Sadílková - Milada Závodská (Charles University): Autobiografie romských přeživších holokaustu vzniklé v době státního socialismu v Československu (Autobiographies by Romani Holocaust Survivors written in Czechoslovakia during the State Socialism)
- Joanna Zofia Spyra (University of Huddersfield): Witnessing in Space: Mapping Survivor Testimonies in the Digital Reconstruction of Trawniki

- Katarzyna Kwiatkowska – Moskalewicz (Polish Academy of Sciences/Gabriel Narutowicz Institute of Political Thought): Memoirs of Jewish Members of the Communist Resistance as Holocaust Testimony
- Ivana Koutníková (Charles University/Paideia - The European Institute for Jewish Studies in Sweden): Forgotten Voices: Czechoslovak Testimonies of Victims of the Nazi Regime Collected in 1945 in Sweden
- Kate Marrison (University of Sussex): Transcript(s) of Interview with a Survivor as Part of the LDML Database

13:00 - 14:00 Lunch

14:00 - 14:25 Beyond Typical Testimony: Quick Introductions

Chair: Magdalena Sedlická (EHRI-CZ, Masaryk Institute and Archives of the Czech Academy of Sciences)

- Kateřina Portmann (Technical University of Liberec): Deníky Otto Goltze (Diary of Otto Goltz)
- Steffen Müller: Affidavit of Natan Gewürz. From His Proceeding for Compensation for "Damages to Body or Health" under the Federal Compensation Act
- Michal Frankl (Leibniz Institute for the History and Culture of Eastern Europe): Descent into Violence. Testimonies of Jews Expelled from Nazi Germany to Czechoslovakia

14:25 - 15:25 Testimony: Gaps and Promises, Challenges and Methods - Discussion

Chair: Michal Frankl (Leibniz Institute for the History and Culture of Eastern Europe)

15:25 - 15:40 Coffee Break

15:40 - 17:00 Digital Tools and Methods: Introductions and Discussion about Expectations

Chair: Martin Wynne (University of Oxford, CLARIN)

- Maria Dermentzi (Université Paris-Saclay/Toolbox 21): EHRI Annotator
EHRI Annotator is a web-based tool designed to facilitate the creation of EHRI Online Editions by streamlining the pre-annotation of Holocaust-related documents. It uses the EHRI-NER model, a domain-specific Named Entity Recognition (NER) model trained to identify people, locations, camps, ghettos, organizations, and dates across multiple languages. EHRI Annotator goes beyond entity recognition by linking detected entities to unique IDs in the EHRI Portal vocabularies and authority sets, as well as GeoNames for geographical locations, ensuring standardized references and interoperability.
- Yael Fried - Elin Eriksson (National Historical Museums): Interconnected Collections – Situating Swedish Holocaust Materials in an International Context through Standardized Vocabularies
In this project we are developing standardized vocabularies for Holocaust materials in Swedish and English, which will be made available as linked open data. The aim is to make Holocaust testimonies and other items more accessible and searchable, and to make such materials from Sweden more visible in an international context as they convey a part of the Holocaust survivor history that is not very well known outside of Sweden.

- Itamar Trainin (The Hebrew University of Jerusalem): The Shape of Testimony: Examining Testimony Interview Styles of Two Holocaust Archives through Computational Analysis
Guide to applying LLMs for the large-scale analysis of Holocaust testimonies, based on methodologies developed in our lab.
- Christopher Brückner, Pavel Pecina, Jiří Kocián, Karin Roginer Hofmeister (Charles University, Memorise project): “MEMORISE Named Entity Recognition”
We developed a Named Entity Recognition system to extract names of people, organizations, locations, camps, and ghettos from Holocaust-related texts. By training Charles University's state-of-the-art model NameTag 3 on ~70k hours of VHA interviews and annotated EHRI Online Editions, it can process archival documents in 17 languages.
- Pavel Ircing - Martin Bulín - Jan Švec (University of West Bohemia in Pilsen): “Digital Humanities with AI: From Data to Insight” *This workshop introduces cutting-edge AI technologies that can support research in digital humanities. Participants will explore how automatic speech recognition (ASR), large language models (LLMs) and retrieval-augmented generation (RAG) can be used to interact with and enrich their own data. They will learn how the combination of full-text and semantic search enables the discovery of patterns and hidden connections across large corpora.*

Tuesday, 21 October 2025:

9:00 - 9:30 Martin Wynne (University of Oxford, CLARIN): Corpus and Computational Linguistics in Holocaust Testimony Research

How are techniques from data-driven linguistics and natural language processing being used in new directions in research with Holocaust testimonies? This talk will explore and reflect on a number of the key areas, and also on the key role that the CLARIN and EHRI research infrastructures can play - not only supporting and promoting interdisciplinary digital research, but also ensuring that the tools, datasets are more easily discovered, analysed, compared and sustained for future research scenarios.

9:30 -12:30 Hands-on Sessions

Presenters will offer sessions sharing their tools and interactively working with participants and their datasets. Participants will be divided into groups and will be able to take part in two of the five following options:

1) Maria Dermentzi (Université Paris-Saclay/Toolbox 21): EHRI Annotator

- *This session will present the EHRI Annotator, a web-based tool for automating entity recognition and linking in multilingual Holocaust-related materials. Participants will have the opportunity to test the tool with their own documents and provide feedback to support its further development.*

2) Yael Fried, Elin Eriksson (National Historical Museums): Vocabularies and Holocaust Testimonies

- *A workshop and discussion about how to structure a vocabulary and how to use pre-existing terms to annotate and index Holocaust testimonies. We will discuss missing terms, hierarchies, sources and connecting vocabularies.*

3) Itamar Trainin (The Hebrew University of Jerusalem): Using LLMs for Distant Reading of Holocaust Testimonies

- *A practical guide to applying LLMs for the large-scale analysis of Holocaust testimonies, based on methodologies developed in our lab. Participants will gain hands-on experience with a complete end-to-end pipeline: from preparing and structuring data for LLM processing, to refining research questions that minimize interpretive errors, and designing multi-layered, LLM-based analytical tasks that operationalize those questions.*

4) Christopher Brückner, Pavel Pecina, Jiří Kocián, Karin Roginer Hofmeister (Charles University, Memorise project): “MEMORISE Named Entity Recognition”

- *We developed a Named Entity Recognition system to extract names of people, organizations, locations, camps, and ghettos from Holocaust-related texts. By training Charles University's state-of-the-art model NameTag 3 on ~70k hours of VHA interviews and annotated EHRI Online Editions, it can process archival documents in 17 languages.*

5) Pavel Ircing - Martin Bulín - Jan Švec (University of West Bohemia in Pilsen): “Digital Humanities with AI: From Data to Insight”

- *This workshop introduces cutting-edge AI technologies that can support research in digital humanities. Participants will explore how automatic speech recognition (ASR), large language models (LLMs) and retrieval-augmented generation (RAG) can be used to interact with and enrich their own data. They will learn how the combination of full-text and semantic search enables the discovery of patterns and hidden connections across large corpora.*

12:30 - 13:30 Lunch

13:30 - 14:30 Conclusions and Final Discussion

Chairs: Jiří Kocián (Malach Center for Visual History) + Maria Dermentzi (Université Paris-Saclay/Toolbox 21)