

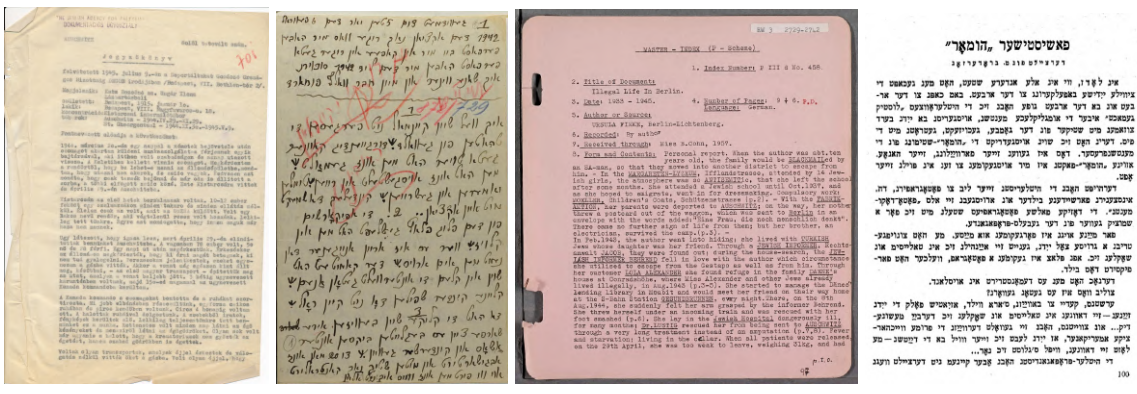


# Unlocking Holocaust Testimony

## EHRI-CLARIN Datathon

Budapest, February 26-27, 2025  
Venue: Eötvös Loránd University, Faculty of Social Sciences  
1117 Budapest, Pázmány Péter sétány 1/a

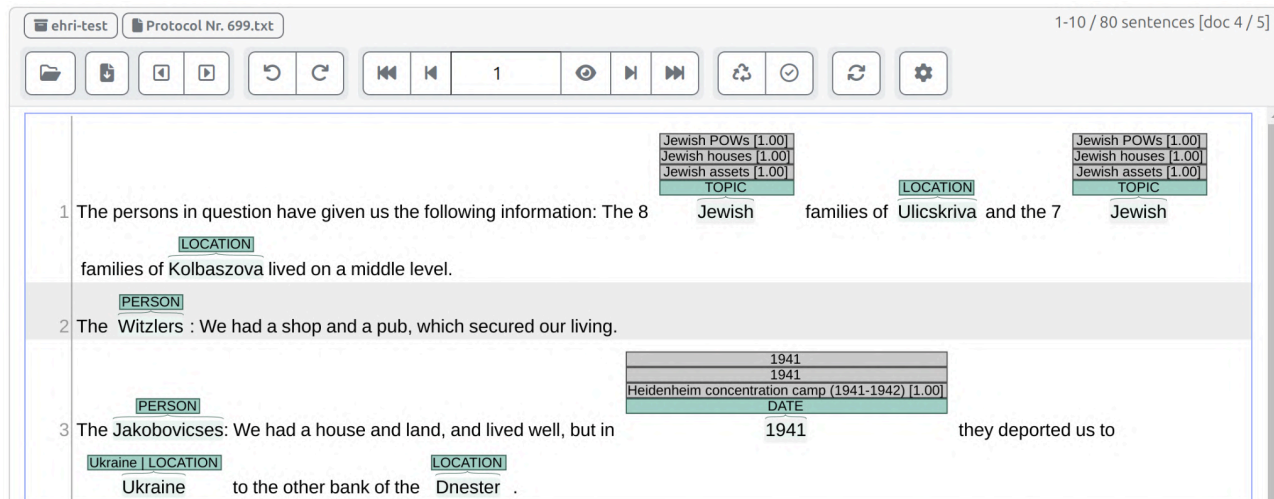
*Are you interested in testimonies of Holocaust survivors? Do you want to contribute to Holocaust research? Would you like to try cutting edge digital tools and methods?*



The hands-on workshop facilitates the building of an annotated, standardised and verified corpus of Holocaust testimonies. In the aftermath of the Holocaust, survivors gave eye-witness accounts in many formats and contexts, including the early documentation campaigns immediately after the liberation and much later national and transnational oral history projects. Differing in length and style, language and media, the tens of thousands of testimonies still haven't been fully exploited and compared in the research.

The "datathon" is a part of a sustained cooperation between the [European Holocaust Research Infrastructure](#) (EHRI) and the [Common Language Resources and Technology Infrastructure](#) (CLARIN) which aims to make the application of computational linguistics on Holocaust documentation easier and more accessible to a variety of researchers. It will engage both experts and citizen scientists in annotating Holocaust testimonies.

The concept of a “datathon” is derived from “hackathon”, a hands-on collaborative session to advance coding towards a shared goal. In this case, participants will work with text data, hence “datathon”, and use innovative tools to identify and annotate important entities, such as names of people and places, within Holocaust testimonies in Hungarian, English and German.



We welcome interested participants from within different disciplines and non-experts. They will be trained to link each name of a person or place they identify in testimonies to their unique identifiers within controlled vocabularies developed and used by EHRI. By doing this, participants will play a crucial role in helping organisers create a more powerful, interconnected resource that enables researchers to study, compare, and draw meaningful insights from Holocaust testimonies. Additionally, the annotated dataset produced during the datathon will support EHRI and CLARIN-affiliated digital humanists, computational linguists, and computational social scientists in developing and refining natural language processing tools, improving both information retrieval and visualisation capabilities.

By joining the datathon, you will:

- Learn more about Holocaust testimony.
- Acquire new skills in working with text corpora.
- Learn about digital humanities and natural language processing (NLP) methods.
- Learn how to use INCEpTION, an open source semantic annotation platform.
- Explore the tools, services and resources offered by EHRI and CLARIN to support research.

To participate, you will need:

- Active knowledge of English; knowledge of Hungarian and other languages is welcome.
- Willingness to learn to work with data and digital tools. Prior experience in annotating texts and programming is not necessary. Knowledge of Holocaust history is also not required. Detailed instructions will be provided.
- Your own laptop which you can use for annotating using an online platform.

If you are interested in joining, contact [barna.ildiko@tatk.elte.hu](mailto:barna.ildiko@tatk.elte.hu)

Organised in cooperation with:



Research Center  
for Computational  
Social Science

30  
Years



Leibniz Institute for  
the History and Culture  
of Eastern Europe



Eötvös Loránd Tudományegyetem  
Társadalomtudományi Kar



Magyar  
Zsidó  
Múzeum  
és Levéltár