European Holocaust Research Infrastructure
H2020-INFRAIA-2019-1
GA no. 871111

Deliverable 10.3

Geographic layer over the EHRI portal

Michal Frankl
MÚA

Mike Bryant
KCL

Michala Lônčíková, Aneta Plzáková
MÚA

Wolfgang Schellenbacher, Éva Kovács, Marianne Windsperger
VWI

Hana Kubátová
HDC

Michael Levy, Rebecca Dillmeier
USHMM

Chiara Renzo
CDEC

Start: September 2020 [M1]
Due: August 2023 [M36]
Actual: August 2023 [M36]

EHRI is funded by the European Union

Document Information

Project URL	www.ehri-project.eu
Document URL	https://www.ehri-project.eu/deliverables-ehri-3-2020-2024
Deliverable	D10.3 Geographic layer over the EHRI portal
Work Package	WP10
Lead Beneficiary	3 - MUA
Relevant Milestones	MS3
Dissemination level	Public
Contact Person	Michal Frankl, frankl@mua.cas.cz
Abstract (for dissemination)	The deliverable summarises the development of the EHRI Geospatial Repository which provides access to Holocaust-related geospatial data, allows researchers to build maps and conduct spatial analysis. To institutions, projects and individual researchers, it offers the possibility to standardise and deposit datasets. The Repository provides new functionality for EHRI to add maps to the EHRI Portal and to other EHRI services.
Management Summary	(required if the deliverable exceeds more than 25 pages) [Max. 500 words]

Table of Contents

1 Introduction

The focus on geospatial data and methods in the framework of the EHRI-3 project reacts to the introduction of spatial approaches and their impact in the field of Holocaust Studies. To extend the array of EHRI's services and in addition to the standardised archival descriptions made available in the EHRI Portal, WP10 prepared and launched a new repository which makes it possible to deposit, search, view, download and reuse geospatial datasets related to the history of the Holocaust.

Even though the *spatial turn* showed a profound impact on the humanities in general, it was particularly enriching in the field of Holocaust Studies. The inter-disciplinary group of researchers working as the Holocaust Geographies Collaborative proved exemplary and influential in showcasing the cooperation of geographers and historians. Many researchers have embraced the spatial history of the Holocaust as a beneficial perspective allowing to study the persecution and genocide on different scales, from the human body to a continent. Understanding space as social, constructed and relational, the new approach allows us to look differently at the roles and subjective perceptions of different types of actors during the genocide.

2 Spatial data in EHRI

In the past, EHRI collected a lot of information that links collections to specific places or regions, including archival descriptions published in the EHRI Portal. Using Wikidata to improve and standardise the information, EHRI created multilingual datasets of ghettos and camps which were published in the Portal. Made available as vocabularies, the records were used to improve the cataloguing of other entries, such as archival descriptions, in the Portal. However, these resources can be only viewed in the user interface of the Portal or downloaded as a vocabulary (in different RDF formats, such as RDF/XML or Turtle), but can not be queried by the standard geospatial methods and not easily be displayed on a map.

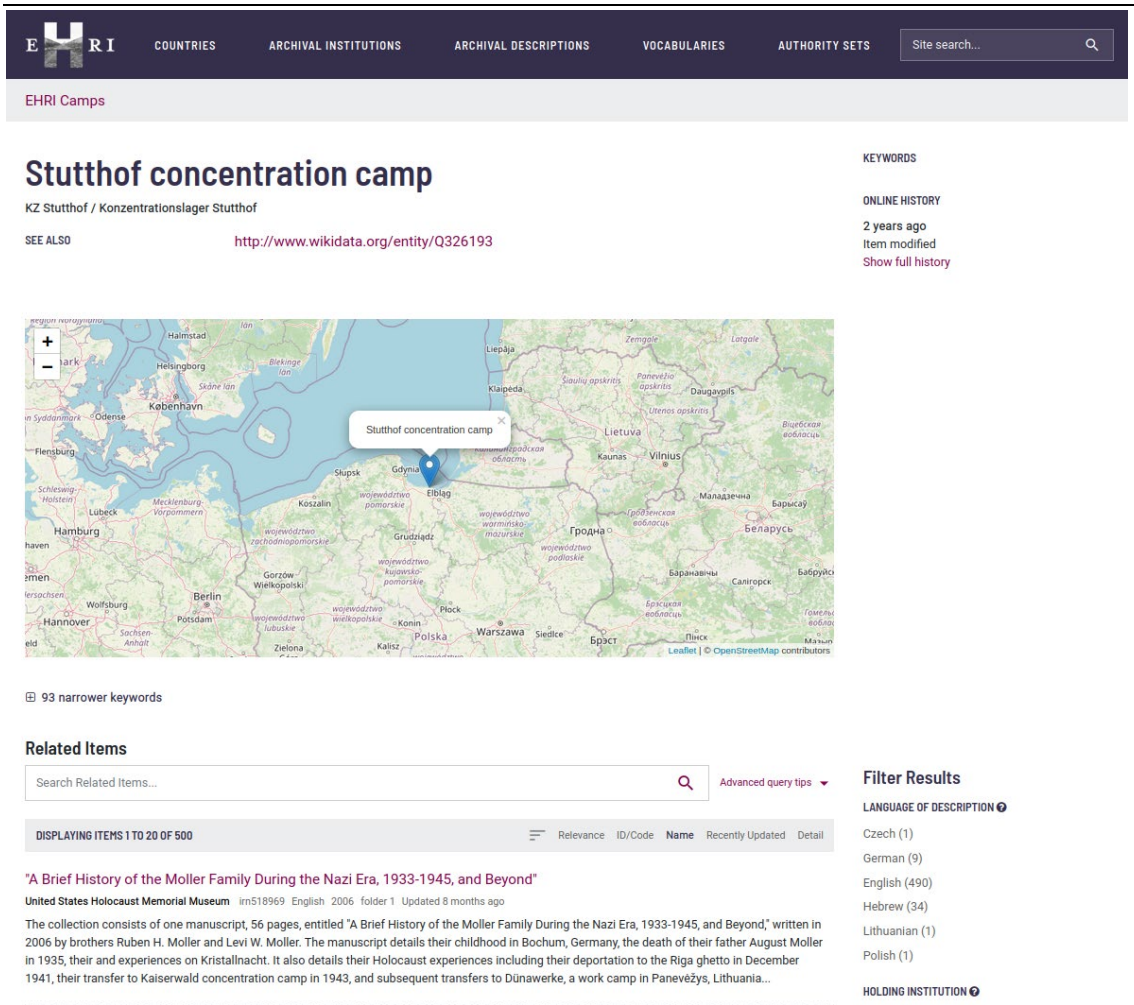

Figure 1: The record for the Stutthof concentration camp in the EHRI Portal with linked archival descriptions

In the EHRI Document Blog, which was launched in 2016, researchers from EHRI partner institutions and from the wider community often experimented with interactive maps. To support these efforts, the blog offers an out-of-the-box option to create rich interactive maps based on archival material, oral history interviews or other types of data. The blog uses the Omeka plugin Neatline, designed especially for small-data, curated approaches to achieve this digital storytelling functionality. The experience demonstrated both the interest in the research community and the need for EHRI staff to systematically support authors in cleaning their data and designing maps.

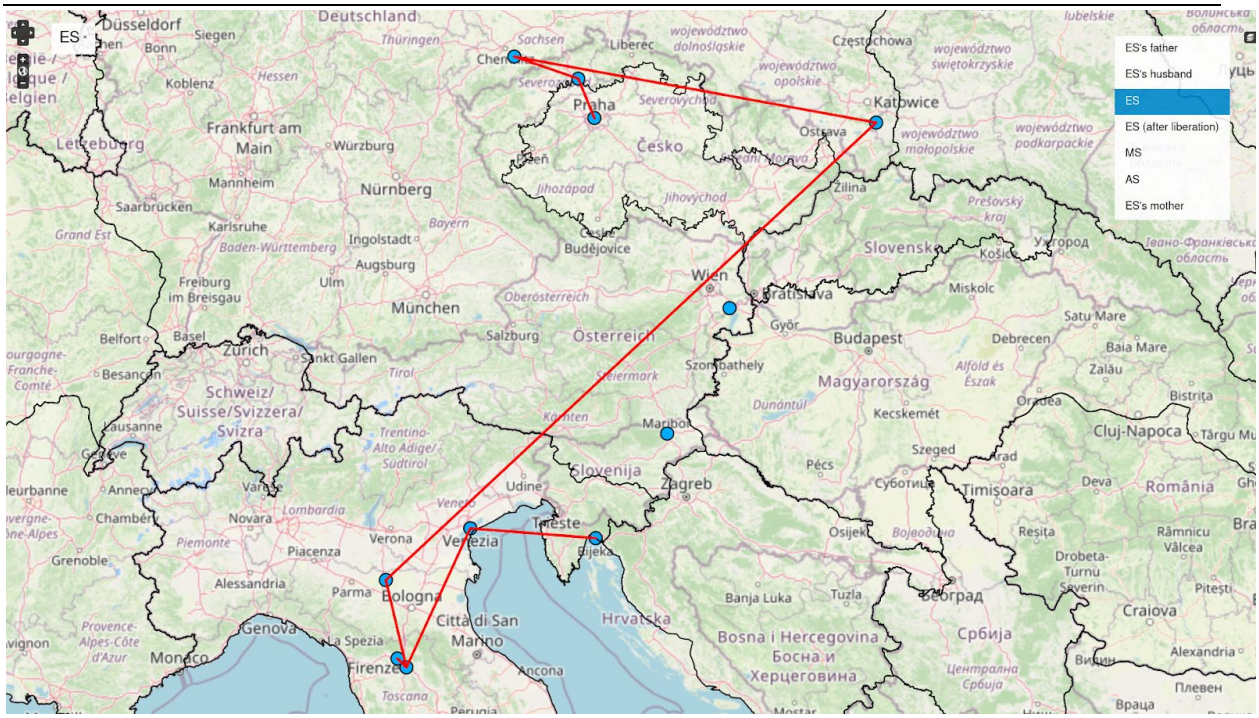

Figure 2: Example of interactive mapping in the EHRI Document Blog

Likewise, the online editions developed and/or supported by EHRI contain a large volume of geographic data. The documents encoded in TEI (Text Encoding Initiative standard) include annotated places, regions and other geographic features which are linked to the vocabularies of EHRI Camps and Ghettos, to Geonames as well as to other authoritative resources. Similar to the Blog, the editions use the Neatline plugin to display interactive maps automatically generated from the TEI source.

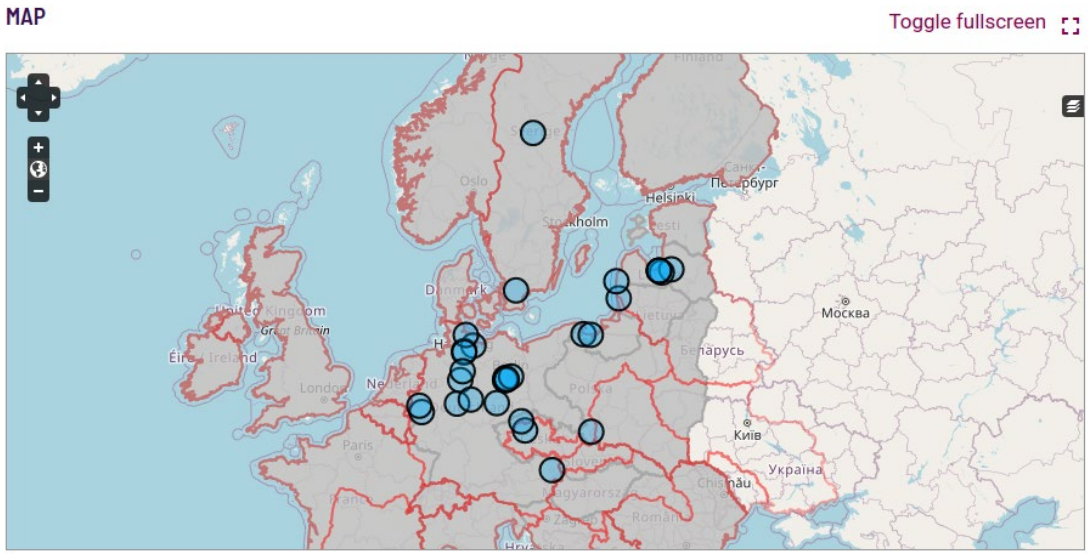

Figure 3: Example of an interactive map automatically generated from the TEI of a Holocaust testimony, https://early-testimony.ehri-project.eu/document/EHRI-ET-YV3549213

The experience with mapping in the Blog and in the online editions also made clear how difficult it was to find open datasets of relevant geographic data, such as historical borders, camps and ghettos which could be reused in a standardised way. To address these needs, in the seven years from launching the blog, its editorial team adapted, corrected and produced different new datasets. This experience, however, pointed to the need to provide access to such data on a more systematic basis.

In April 2019 at the Vienna Wiesenthal Institute for Holocaust Studies, EHRI organised a workshop titled *It happened here! Digital and shared: Holocaust history in public space* which brought together projects and organisations developing mobile (web) applications making available information about the Holocaust directly in the spaces where the persecution unfolded. It examined not only the technical aspects of these data projects, but also how these efforts contributed to a different forms of construction of space and new public history approaches rooted in everyday places and spaces. The many projects focusing on Jewish houses and places of persecution further demonstrated the grass-root activities in creating geospatial data related to the Holocaust, the need to share such data and to exchange experience and provide support.

3 Preparatory work

In order to build on this experience and to find the best solution to such challenges, WP10 first systematically investigated the needs of the research communities served by EHRI and examined standards and available software options.

3.1 Focus groups

To better understand the needs in the relevant user communities, on March 19 and 25, 2021 EHRI conducted two structured interviews with focus groups. The participants included researchers using geospatial data and/or methods for their own projects as well as representatives of documentation projects. Many were also involved in projects that document names and fates of persons persecuted during the Holocaust. WP10 took care to balance the focus groups in terms of gender, geography and academic position. The meetings were conducted over Zoom with two moderators from within EHRI facilitating the discussion.

For its internal evaluation, WP10 created anonymised and shortened transcripts of the recorded focus groups. The suggestions from the two groups included especially:

- Providing access to data which could be used to build maps, disambiguate place information (especially in multilingual settings) and to conduct further spatial analysis.
- Support and training in the complex and confusing field of digital geographic methods.

3.2 Further input

Following the request of WP10, the user survey conducted in the framework of the EHRI Preparatory Phase project (WP5) at the end of 2020 included a question examining user expectations related to the research in Holocaust geographies ("What geospatial EHRI services would you consider most useful to you?"). Across all user groups, the demand for open spatial data and for the possibility to display the content of the EHRI Portal on the map figured prominently. In addition, users expressed interest in training and tools which would allow them to work with data provided in this way.

WP10 also consulted with the Holocaust Geographies Collaborative to make sure that the new EHRI services meet its needs. These discussions also clearly demonstrated the user demand for a standard-compliant and open repository of geospatial data. Emphasis was put on the need not only to make available datasets with basic metadata, but also to include detailed information about their sources, history and possible gaps in coverage and knowledge.

4 EHRI Geospatial Repository

While EHRI originally considered adding geographic layers over the EHRI Portal, after mapping needs and possible solutions, WP10 opted for a more robust solution in the form of a dedicated EHRI Geospatial Repository which will provide data not only for the EHRI Portal, but to a number of services for EHRI, archives, projects and researchers. Dedicated solution can ensure compatibility with established geospatial standards and services and make it easier to be used by diverse types of users and services.

The Repository aims to build on the incredible productivity of the wider user communities which produce a large amount of - often overlapping - geospatial data. In designing the Repository, it also reacted to the different structure of such data and different ways it is created, in comparison to descriptions of archival holdings which are collected in the Portal. We also aim to avoid duplicating efforts when different institutions or projects gather similar types of spatial information, for instance, when cataloging archival collections or tracking the movements of Holocaust victims.

The EHRI Geospatial Repository which was officially launched during the EHRI-3 General Partner Meeting in Jerusalem on May 30, 2023 offers researchers, projects and organisations to:

- deposit and share geodata in a standard compliant, reusable way
- search and explore datasets based on their metadata
- use, download and combine datasets for mapping, data enrichment or spatial analysis

The repository is freely accessible at https://geodata.ehri-project.eu/.

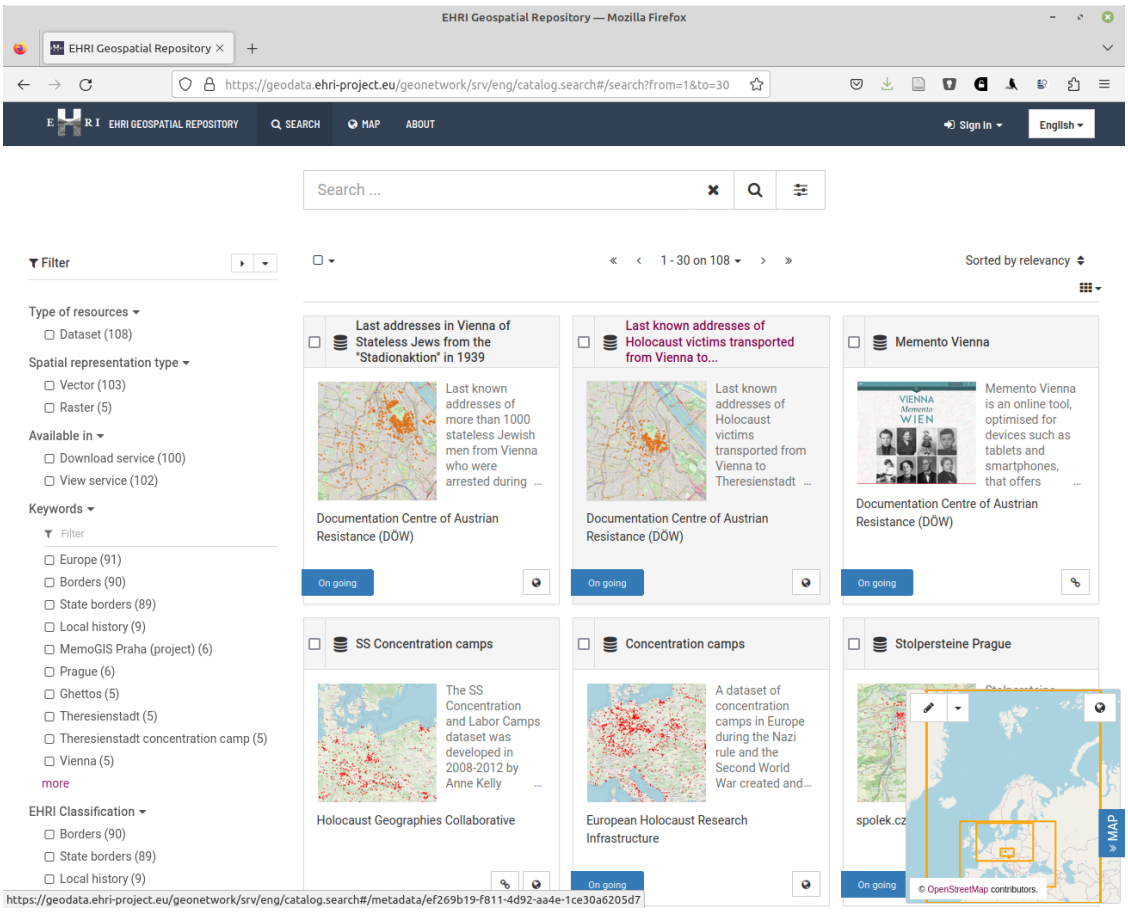

Figure 4: EHRI Geospatial Repository, search page

The Geospatial Repository aims to provide access to any geospatial data which can be relevant for the research about the Holocaust. This includes, but is not limited to the EHRI taxonomy used for cataloguing data in Holocaust Geographies which is developed here. Examples include historical borders, places of persecution such as ghettos and camps, data related to local history of the Holocaust and to history of refugees during the Holocaust or sites of remembrance. The data can be provided both in vector data (consisting of points, lines and polygons) and raster data (made up of pixels; images).

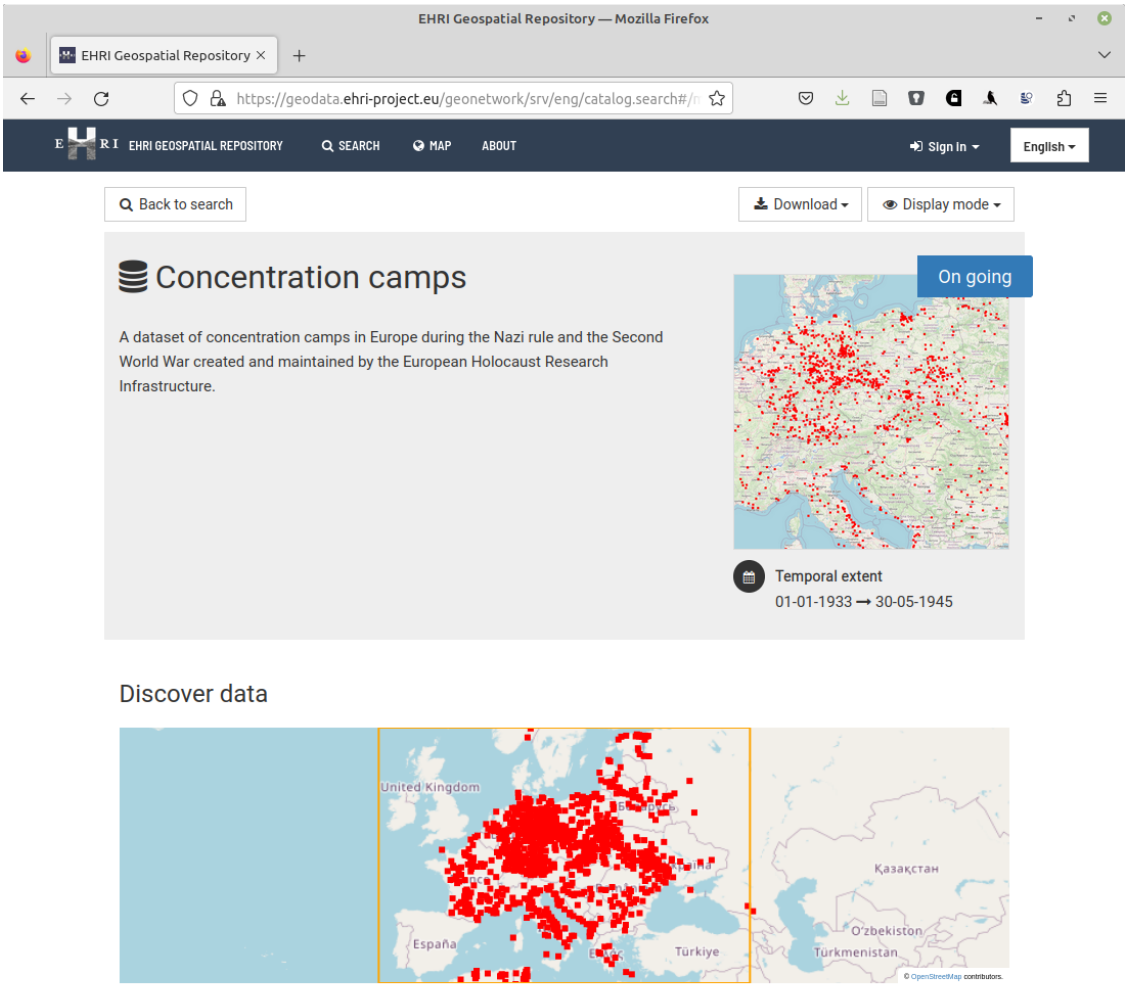

Figure 5: Example of a dataset, EHRI Concentration Camps

4.1 Technical implementation

Following a thorough exploration and testing of available open source software solutions, WP10 decided to use Geonetwork, a software designed to build catalogues of spatially referenced resources. Apart from its general functionality and design, the decisive features included the emphasis on geospatial metadata standards (especially ISO 19115/19139) and the possibility to provide access (via a linked Geoserver) to standard geospatial data services such as Web Map Service (WMS) and Web Feature Service (WFS). Geonetworks also allows to store original datasets in any format, making it possible to archive older versions of data or documentation.

Geonetwork makes it possible to identify datasets based on metadata search or selecting a region via a map. It allows to display descriptive metadata and download the full datasets as well as to combine datasets as layers on an interactive map.

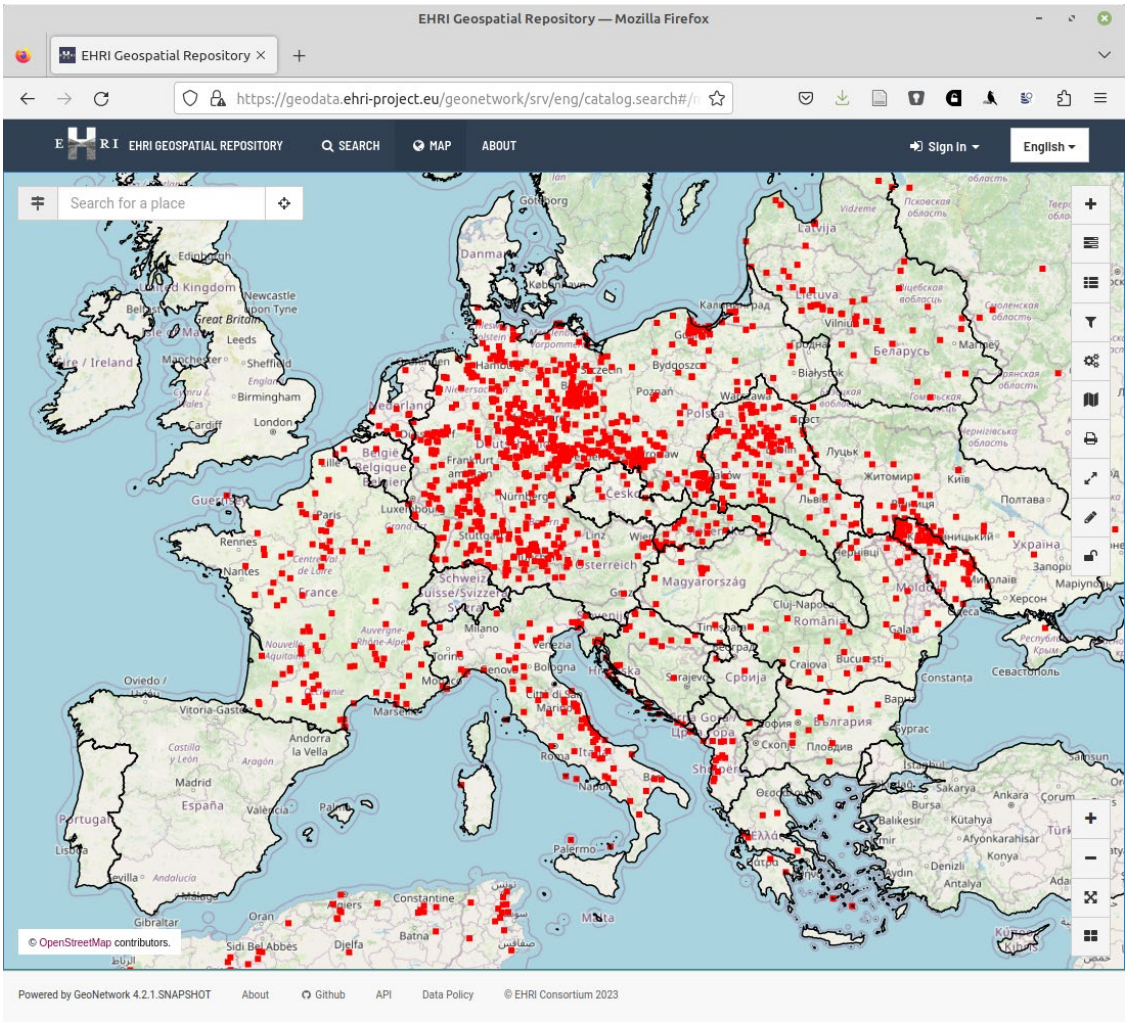

Figure 6: Example of an interactive map combining two layers, the EHRI dataset of concentration camps and borders as of the end of 1942

To further promote open standards, the data will be saved in the EHRI Geoserver repository as Geopackages, the file format promoted for geospatial information by the Open Geospatial Consortium. To make the preparation of data easier for users unaccustomed to working with often complex geospatial tools, WP10 built a simple CSV to Geopackage converter which is particularly useful for tabular data. For the administrators of the Repository, WP10 also built a tool to upload the resulting Geopackages to the EHRI Geoserver.

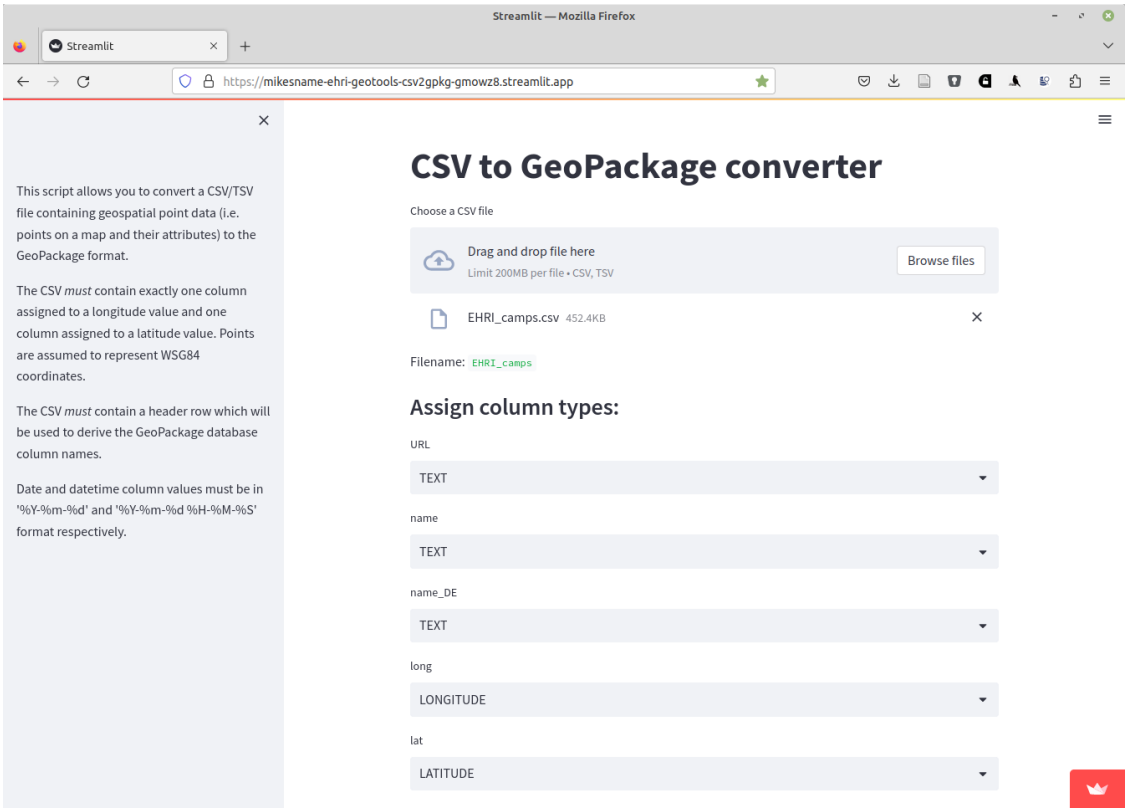

Figure 7: EHRI CSV to GeoPackage convertor

4.2 Data workflow

Data providers are required to submit a minimum of standardised metadata describing the dataset. This will facilitate discovery and provide users with an understanding of the (often complex) data. This also includes the provenance and history of the dataset, its sources and structure, as well as potential uncertainties and gaps which researchers should understand while using the data.

The datasets can be uploaded and saved in Geonetwork in a variety of formats (next to Geopackage, data can be provided in the first step as shapefiles, CSV, GeoTIFF, etc.). EHRI's staff will control the data, standardise it if needed and internally store it as a GeoPackage. In a next step, the full geospatial dataset will be pushed into the EHRI Geoserver and linked to the Geospatial Repository which serves as a store of descriptive metadata.

Because adding information directly through the Geonetwork interface can be tricky, due to the complexity of geospatial standards and their unfamiliarity to most users, data providers can submit their information about the data via a form. They can also contact the WP10 team via email (at geodata@ehri-project.eu). To facilitate the process, the EHRI team created documentation describing the repository, its usage and the process of depositing data. For general usage, it relies on the documentation for Geonetwork.

5 Next steps: dissemination, expansion and user support

To test and publish the EHRI Geospatial Repository, WP10 used a small number of about one hundred datasets from the Holocaust Geographies Collaborative and from EHRI's partner organisations. After the launch, EHRI will systematically approach organisations, projects and individual researchers who can provide further data. WP10 already collected a list of potential data providers.

The link to the repository has been added to the EHRI project website (in navigation and as a news story) and included in the EHRI Newsletter published in July 2023. EHRI will further disseminate the knowledge through presentations, how-tos showcasing possible usage (for instance connecting GIS software to the repository and effectively build maps) and articles in the EHRI Document Blog.

Using the services of the Geospatial Repository, WP10 will also integrate the data to its other platforms, such as the Document Blog and editions. The Neatline plugin used in both platforms can integrate WMS layers provided by the Repository (specifically, its Geoserver component). This will contribute to a smoother publication process, offer quicker response and provide more reliable service. In the future, EHRI will use the Repository to visualise the content of the Portal via maps and to experiment with the extraction of geospatial data from textual descriptions of archival material.

WP10 also started to work on a training concept which includes not only how-tos and contributions to the EHRI Document Blog, but also hands-on training and cooperation with academic institutions.

6 Conclusion

The EHRI Geospatial Repository brings a new quality to spatial studies about the Holocaust, offering researchers access to a wide range of data and the possibility to build maps and conduct analysis of geospatial data. To institutions, projects as well as individual researchers, it provides the possibility to standardise and deposit relevant data. Following the launch of the repository, EHRI continues to collect data, provide support and offer training.