



**European Holocaust Research Infrastructure
H2020-INFRAIA-2014-2015
GA no. 654164**

Deliverable 14.4

Marketplace of Digital Historiography Methods and Tools

**Rachel Pistol
KCL**

**Michael Bryant
KCL**

**Michael Levy
USHMM**

**Start: Month year [M 25]
Due: Month year [M 46]
Actual: Month year [M53]**



EHRI is funded by the European Union

Document Information

Project URL	www.ehri-project.eu
Document URL	[www.....]
Deliverable	D14.4 Report on Marketplace of Digital Historiography Methods and Tools
Work Package	WP14
Lead Beneficiary	14 – KCL
Relevant Milestones	MS4
Dissemination level	Public
Contact Person	Rachel Pistol, King's College London, rachel.pistol@kcl.ac.uk
Abstract (for dissemination)	In conversation with Holocaust researchers through events organized by EHRI it is apparent that many researchers have an interest in learning more about digital tools and methods but many researchers are not sure how to start learning about them. This deliverable seeks to help researchers discover more about tools that are available and demonstrate how they can be used in easy to understand language for those unused to experimenting with digital technology. This has been achieved through the creation of a series of online, illustrated guides.
Management Summary	(required if the deliverable exceeds more than 25 pages) [Max. 500 words]

Table of Contents

1	Introduction	4
2	Digital Methods	4
2.1	Goals of the Digital Tool Guides	4
2.2	Digital Tools Guides	4
2.3	The Resource.....	5
2.4	Results	7

1 Introduction

In the framework of WP14, EHRI recognises there are a number of digital tools available that can assist Holocaust researchers and archivists in their work. EHRI has, therefore, created an online space and populated it with several digital tool guides, where interested parties can learn more about how to use some of these tools.

This deliverable provides background information about the goals and creation process of the digital tool guides. The guides themselves are available on the EHRI Digital Tools repository on GitHub: <https://github.com/EHRI/ehri-data-analysis-tools>

2 Digital Methods

2.1 Goals of the Digital Tool Guides

In conversation with researchers working in Holocaust and refugee studies, through events organized by EHRI, it became apparent that many researchers have an interest in learning more about digital tools and methods that they hear discussed at EHRI events, but that many researchers are not sure how to start learning about them. There can be many barriers to learning more about digital methods and some researchers reported that they felt intimidated by the prospect of learning about digital tools, or that they were embarrassed to ask what may seem like very basic questions when talking to a researcher more advanced in using digital tools. Researchers also reported not knowing where to look for more information on how to use digital tools or even find tools that would be useful to them in the first place. This deliverable seeks to help these researchers discover more about tools that are available and demonstrate how they can be used in easy to understand language for those unused to experimenting with digital technology.

2.2 Digital Tools Guides

These digital tool guides build on the work of deliverable *14.3 Open Analysis Toolset*, which focused on the creation of an eleven lesson interactive course to provide detailed access to a digital tool that will help Holocaust and refugee studies researchers to use their data.

EHRI has recognized a need not only for detailed courses that go into depth on the subject of digital tools, but also for providing very short, illustrated, introductory guides, that will capture the interest of audiences less experienced with the use of digital technology. In order to build on one of the central objectives of EHRI, to reach out to and support the research community and develop digital methods for Holocaust research, it is necessary to create resources for all types of digital user. To further support the growth of the EHRI network and EHRI Portal, it is desirable to engage new users and encourage them with the fact that it is possible to present their data in new ways by using digital tools. Many researchers who are outside of the digital humanities are unaware of new developments in the digital marketplace and of tools that require very little previous knowledge in order to transform the way their research is conducted. The guides created as part of this deliverable are intended to offer inspiration to

researchers to encourage them to experiment with using digital tools as part of their research, and to learn new skills.

The guides have been created for novice use and contain illustrations as well as step-by-step instructions on how to achieve results using a variety of digital tools. Once a researcher has engaged with a guide on digital tools that enables them to achieve something new with their data it is likely he or she will be encouraged to engage in further use of digital resources. These guides have, therefore, been created in order to showcase a variety of different techniques and methods that are relevant to researchers of the Holocaust. This is a dynamic resource that will continue to grow as more relevant tools are identified.

2.3 The Resource

The guides are available on the EHRI Digital Tools repository on GitHub:
<https://github.com/EHRI/ehri-data-analysis-tools>

The guides take a cross-section of some of the most useful digital tools that are currently available and that can particularly benefit Holocaust researchers. This resource will continue to grow as more guides are created will help with research.

For the initial release 8 dedicated guides have been developed, covering the following topics:

- Extracting CSV data from the EHRI Search API
- Entity Matching
- Using OpenCV for Face Detection
- Importing tables from websites into spreadsheets
- Windows Subsystem for Linux (WSL): A convenient way to run Linux tools on a Windows 10 system
- Using spatial data in Tableau
- Topic Modelling Experiments
- How to capture and reference a webpage in your research using Zotero

The key to each guide is breaking down how to use the digital tool in easy to understand language, and with screenshots demonstrating every step of the guide.

Below is a selection of screenshots from two of the guides:

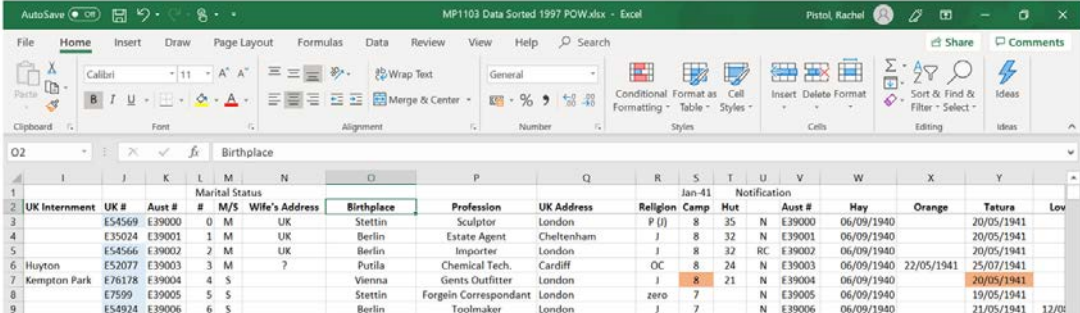
[GitHub, Inc. \(US\) https://github.com/EHRI/ehri-data-analysis-tools/tree/master/entity-mat](https://github.com/EHRI/ehri-data-analysis-tools/tree/master/entity-mat)

Entity Matching

At EHRI, we are constantly looking at Digital Tools that can assist in the research of Holocaust and refugee related topics. We hope to make these tools as accessible as possible so that researchers who have no experience with digital tools will consider trying new ways of using their data.

A number of datasets presented at EHRI related events contain place names that researchers wish to present on maps. In order to accurately display locations, latitude and longitude coordinates are essential. EHRI has created an entity matching tool to help researchers find coordinates for present day locations.

The dataset in this example relates to the birthplaces of a number of refugees who escaped from Nazi occupied Europe to Great Britain and were then deported to Australia. The data has been collected in an Excel spreadsheet and there is a column with the towns of the refugees' birthplace:

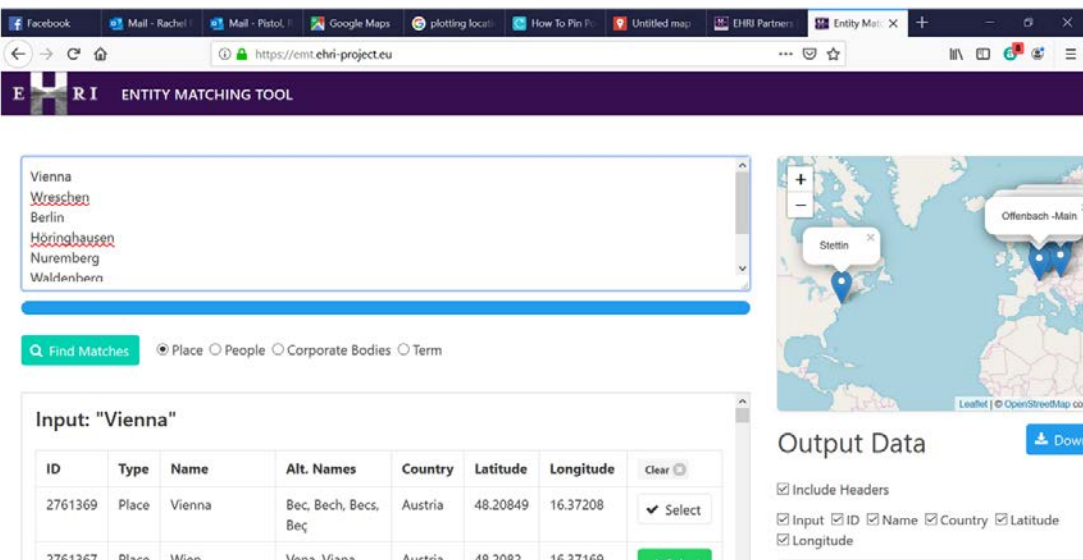


	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1																		
2	UK Internment	UK #	Aust #	#	M/S	Wife's Address	Birthplace	Profession	UK Address	Religion	Camp	Hut	Aust #	Hay	Orange	Tatura	Low	
3		E54569	E39000	0	M	UK	Stettin	Sculptor	London	P (J)	8	35	N	E39000	06/09/1940		20/05/1941	
4		E35024	E39001	1	M	UK	Berlin	Estate Agent	Cheltenham	J	8	32	N	E39001	06/09/1940		20/05/1941	
5		E54566	E39002	2	M	UK	Berlin	Importer	London	J	8	32	RC	E39002	06/09/1940		20/05/1941	
6	Huyton	E52077	E39003	3	M	?	Putila	Chemical Tech.	Cardiff	OC	8	24	N	E39003	06/09/1940	22/05/1941	25/07/1941	
7	Kempston Park	E76178	E39004	4	S		Vienna	Gents Outfitter	London	J	8	21	N	E39004	06/09/1940		20/05/1941	
8		E7599	E39005	5	S		Stettin	Foreign Correspondant	London	zero	7		N	E39005	06/09/1940		19/05/1941	
9		E54924	E39006	6	S		Berlin	Toolmaker	London	J	7		N	E39006	06/09/1940		21/05/1941	
10																	12/01	

Fig. 1. Screenshot from the Entity Matching Guide

For the purpose of this example, places are used, though it is also possible to match people, corporate bodies, and terms using this tool.

The cells listing the birthplaces are copied into the entity matching tool. The tool automatically turns the column of cells into a list of one entity per line. Click on the 'Find Matches' button.



Facebook Mail - Rachel Mail - Pistol Google Maps plotting local How To Pin P. Untitled map EHRI Partners Entity Mat

<https://emt.ehri-project.eu>

E H R I ENTITY MATCHING TOOL

Vienna
 Wreschen
 Berlin
 Horinghausen
 Nuremberg
 Waldenbarn

Place
 People
 Corporate Bodies
 Term

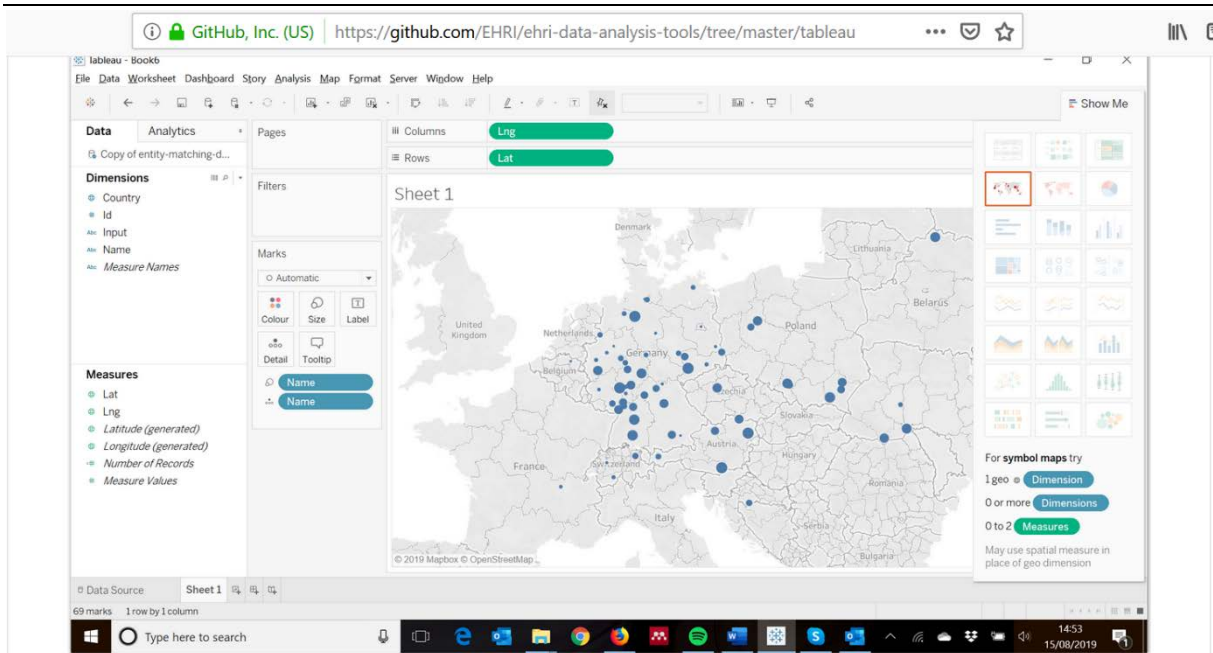
Input: "Vienna"

ID	Type	Name	Alt. Names	Country	Latitude	Longitude	Clear
2761369	Place	Vienna	Bec, Bech, Beccs, Beç	Austria	48.20849	16.37208	Select
2761367	Place	Wien	Vena, Viana	Austria	48.2082	16.37169	Select

Output Data

Include Headers
 Input
 ID
 Name
 Country
 Latitude
 Longitude

Fig. 2. Screenshot from the Entity Matching Tool guide



To change the size of the marker to be proportional to the number of records from the town or city, the 'Name' dimension for 'Size' needs to be changed into a measure. There are a variety of different functions that can be used in measures, and in this case we want to use the 'count' function as we wish to count the number of times a town or city appears in our data. Click on the 'Name' dimension in the 'Marks' pane that is linked to 'Size', then select 'Measure' and 'Count'.

Fig. 3. Screenshot from the Tableau guide

2.4 Results

The guides will be disseminated through the EHRI network. Researchers have already expressed excitement over having step-by-step illustrated instructions to follow that will help them experiment with new methods of using their data. EHRI looks forward to seeing this resource continue to develop.