

# ARBEITEN MIT HISTORISCHEN GEODATEN

KURZWORKSHOP  
UNIVERSITÄT TÜBINGEN  
JANUAR 2023

---

Dr. Monika Barget (Maastricht/Mainz)





# ÜBERBLICK



- 1) Was sind "Geodaten" und welche Rolle spielen sie in der historischen Forschung?
- 2) Wo findet man Geodaten und welche Bearbeitung ist erforderlich?
- 3) Mit welchen digitalen Werkzeugen kann man (historische) Geodaten bearbeiten?
- 4) Interaktives Arbeiten mit ausgewählten Werkzeugen:
  - a) Geocodierung mit Google Awesome
  - b) Geocodierung mit Python in Google Colab



1) Was sind "Geodaten" und welche Rolle spielen sie in der historischen Forschung?





# DEFINITION "GEODATEN"

---

"Unter Geodaten versteht man Informationen, die eine Zuordnung zu einer räumlichen Lage besitzen (Georeferenz). 90% aller Daten können einem Ort zugesprochen werden bzw. einen Raumbezug und sind somit Geodaten."

<https://www.geoportal.rlp.de/>

"Historische Geodaten bieten den raumbezogenen historischen Wissenschaften die gleichen Auswertungsmöglichkeiten wie aktuelle Geobasisdaten. Sobald bei historischen Daten datenschutzrechtliche Belange über natürliche Personen nicht mehr zu beachten sind, können auch Personendaten in Geoinformationssystemen visualisiert werden."

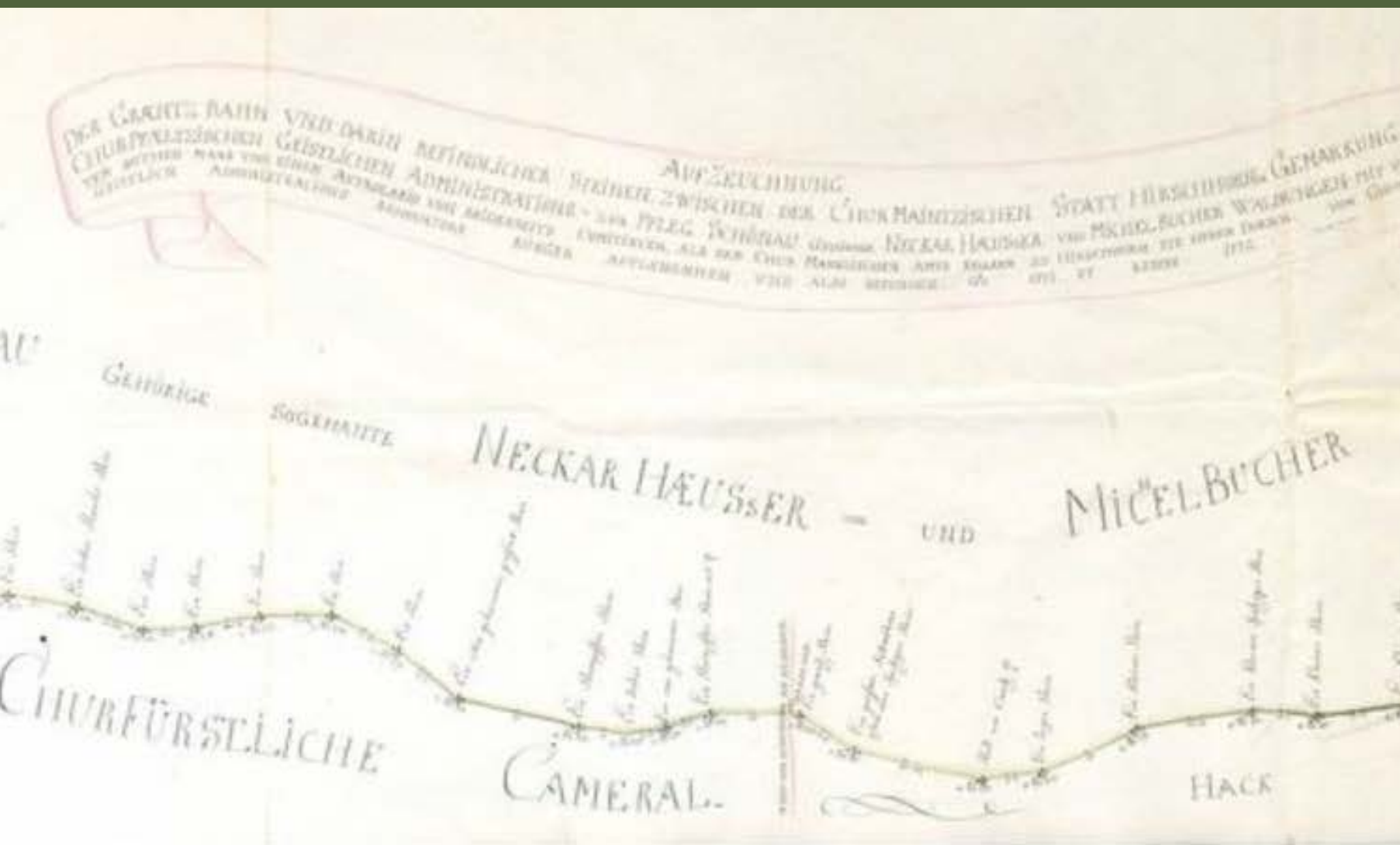
<https://www.leo-bw.de/themenmodul/sudwestdeutsche-archivalienkunde/archivaliengattungen/raumbezogene-abbildungen/geodaten>



# RELEVANZ IN DER HISTORISCHEN FORSCHUNG

## PRIMÄRE GEODATEN

- explizite Ortsnennungen in Ortsverzeichnissen
- Ortsnamen in Lexika (z.B. Zedlersches Lexikon)
- Reisebeschreibungen
- Ortsdaten in (handgezeichneten oder gedruckten) Karten
- Ortsangaben in Briefköpfen, auf Telegrammen etc.



## SEKUNDÄRE GEODATEN

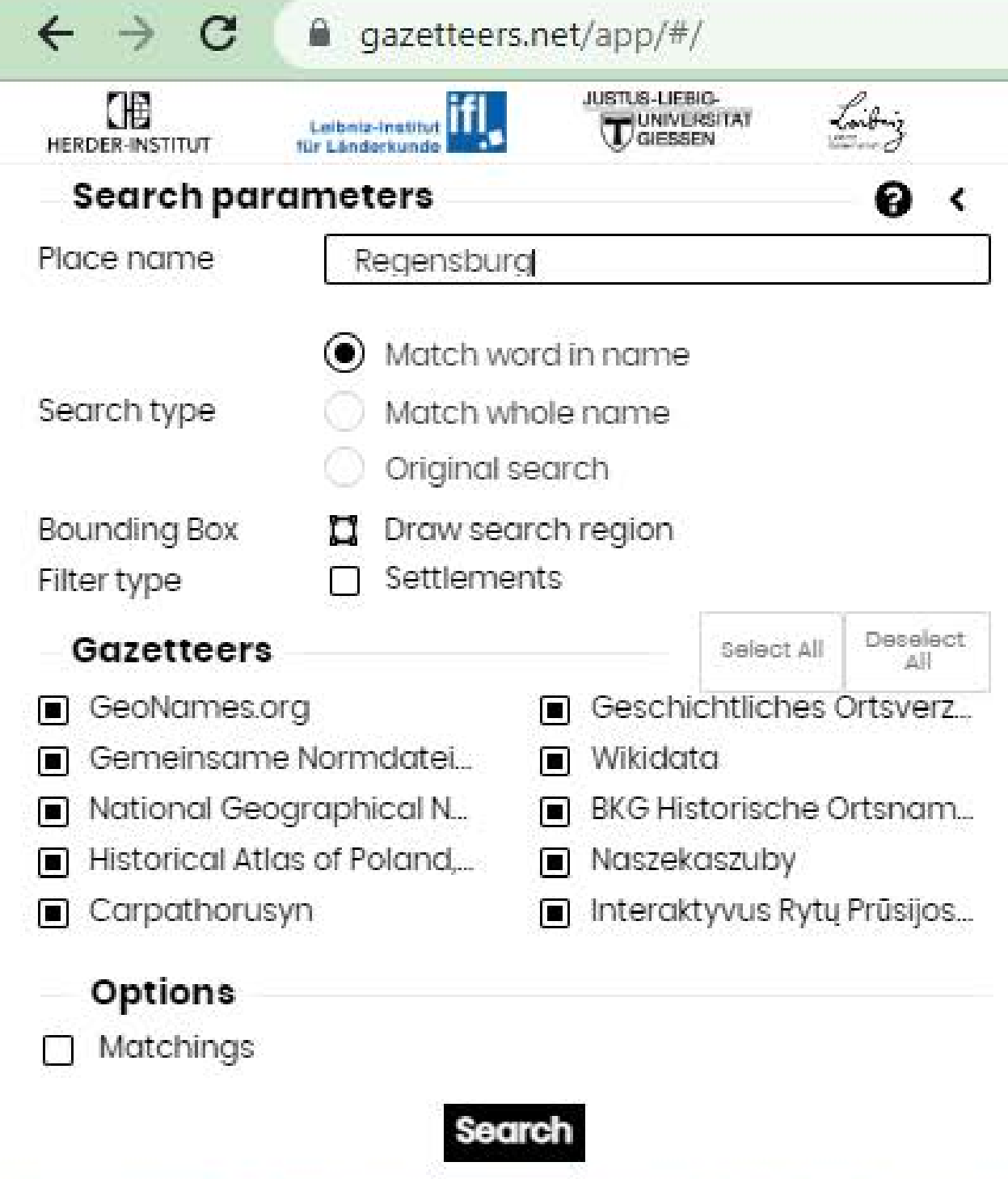
- rekonstruierte Ortsdaten aus verfügbaren Metadaten (oft basierend auf Ergebnissen der Provenienzforschung)
- aus Biographien abgeleitete Ortsdaten
- aus Institutioneninformationen abgeleitete Ortsdaten
- aus archäologischen Funden abgeleitete Ortsdaten



# BEISPIEL "DIGIKAR"

- 3-jähriges Kooperationsprojekt des Leibniz-Instituts für Europäische Geschichte Mainz (IEG), des Leibniz-Instituts für Länderkunde Leipzig (IfL), des Leibniz-Instituts für Ost- und Südosteuropaforschung Regensburg (IOS), der Johannes Gutenberg-Universität Mainz (JGU) und der École des Hautes Études en Sciences Sociales Paris, Frankreich (EHESS)
- Alternative und experimentelle Erfassung, Modellierung und Visualisierung historischer Ortsdaten aus Kurmainz und Kursachen

<https://digikar.eu/>

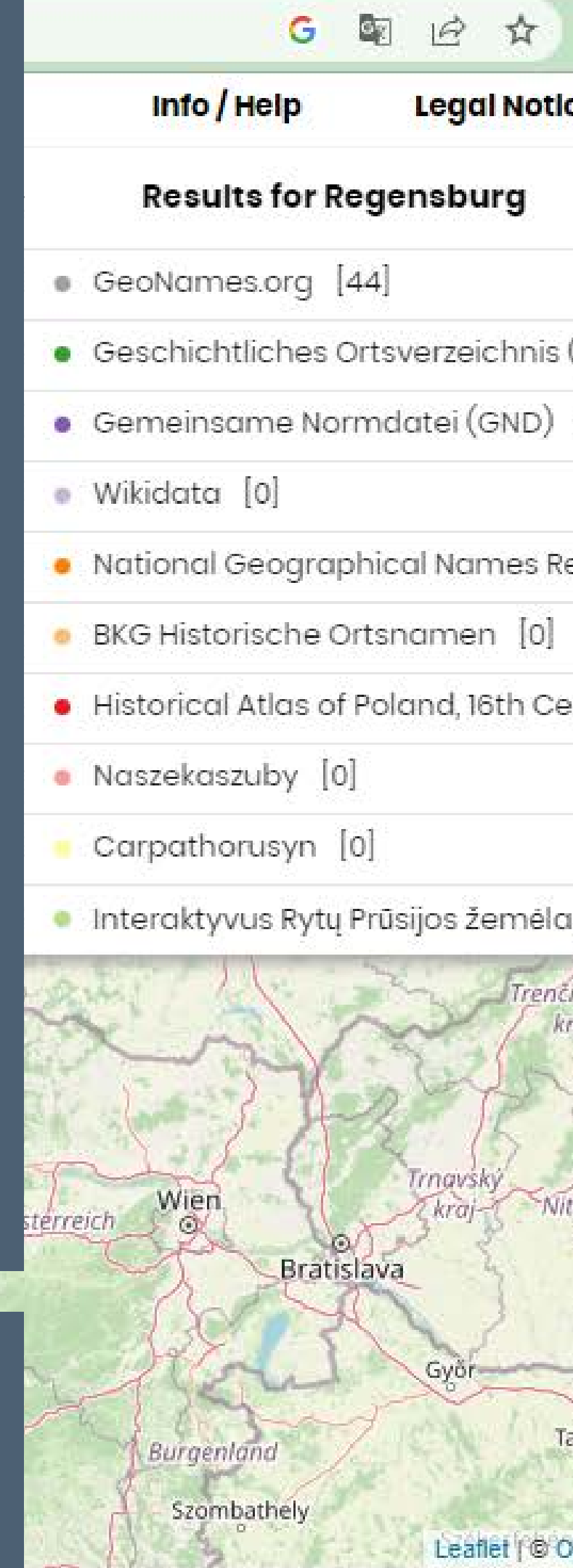


## BEISPIEL "GAZETTEERS NET"

- "The gazetteers.net web application is developed as part of the Gazetteer research project by the Herder Institute (HI), the Institute for Regional Geography (IfL) and the Justus Liebig University Giessen (JLU). The application is intended to support users in working with different digital gazetteers, and to help them explore their content and metadata structure."
- Suche nach Ortsnamen und Verknüpfung mit vorhandenen Einträgen aus nationalen und internationalen (Norm-)Datenbanken (z.B. GND)



<https://gazetteers.net/app/#/>







# BEISPIEL "WHG"

Amsterdam

Search Index 🔍

Search Database 🗄️

pre-filters

PLACE CATEGORIES [CHECK ALL](#) ?

- administrative entities (A)
- cities, towns, hamlets... (P)
- sites, buildings, complexes... (S)
- roads, routes, rail... (R)
- regions, landscape areas (L)
- terrestrial landforms (T)
- water bodies (H)

TEMPORAL (DEFAULT)

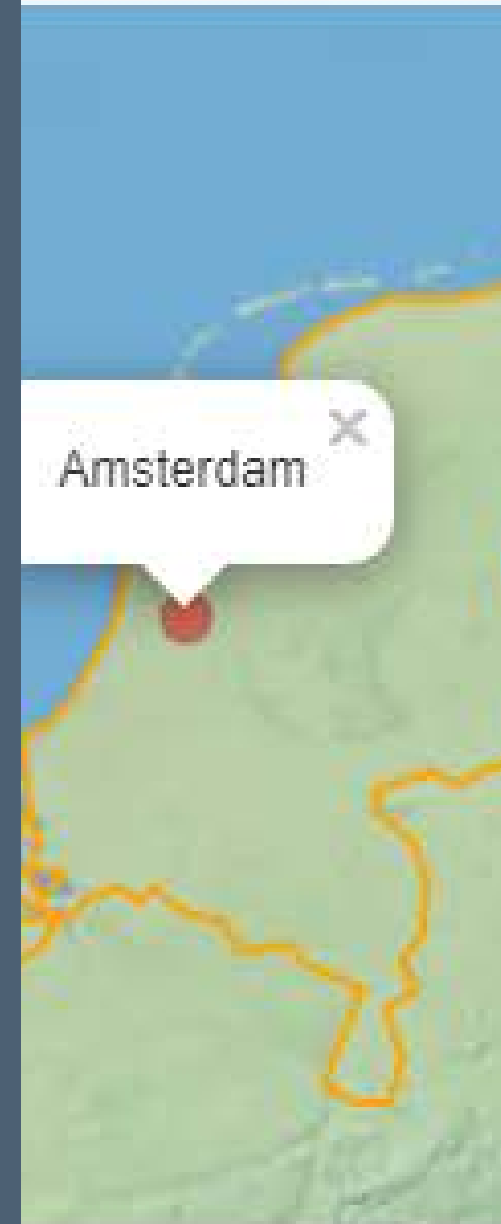
1700

SPATIAL

Region, country

[reset search](#)

- "World Historical Gazetteer (WHG) is providing a collection of content and services that permit world historians, their students, and the general public to do spatial and temporal reasoning and visualization in a data rich environment at global and trans-regional scales."
- Datenbeiträge in zwei Formaten (Linked Places and Linked Traces) möglich.
- Visualisierungs- und Analysetools
- vor allem: Verknüpfung der eigenen Daten mit bestehenden Datensätzen und Teilen von Daten für die weitere Forschung
- Schwerpunkt auf Erfassung der Zeitdimension



UNION INDEX SEARCH RESULTS (1)

List may include records with no geometry

Title ⚡	Linked ⚡	Countries ⚡	Type(s)
<a href="#">Amsterdam</a>	4	NL	inhabited p...

<https://whgazetteer.org/>





Time  
Machine

ABOUT US

BUILDING A TIME MACHINE

TIME

# BEISPIEL "TIME MACHINE"

- "Time Machine is aiming to join Europe's rich past with up-to-date digital technologies and infrastructures, creating a collective digital information system mapping the European economic, social, cultural and geographical evolution across times."
- Leitung: Time Machine Organisation (TMO) Wien
- Mitglieder: ca. 800 Kultureinrichtungen, Universitäten, private Forschungsinstitute und Tech-Unternehmen weltweit.
- Förderung: österreichische Regierung & Horizon2020

<https://www.timemachine.eu/>

ADDING A NEW DIMENSION TO THE PAST



2) Wo findet man Geodaten  
und welche Bearbeitung ist  
erforderlich?





# HERAUSFORDERUNGEN

---

- Geodaten liegen in verschiedenen Quellentypen von handschriftlichen Archivalien bis hin zu gedruckten Karten vor.
- Geodaten wurden in der Archivgeschichte nicht immer systematisch erfasst (z.B. in Metadaten).
- Vorhandene strukturierte Geodaten sind oft zeitabhängig (siehe Hinweise zu politischen Strukturen der 30er Jahre in den Metadaten der Mainzer Geburts- und Lehrbriefe).
- Die Extraktion und Modellierung von Geodaten aus historischen Quellen ist zeitaufwendig und erfordert oft schwierige Entscheidungen im Bereich der Normalisierung und Datenunsicherheit.
- **WICHTIG:** Ortsnamen per se genügen nicht!





# VOM ORTSNAMEN ZUR GEOINFORMATION

**NAME**



Identifikation als "Ortsname": z.B. Paris  
als Stadt vs. Paris als mythologische Figur



Zeitabhängige "Ortstypen"  
(ggf. Orientierung an  
bestehenden Ontologien):

z.B. ein Dorf, das sich zu einer  
Stadt entwickelt und damit  
auch neue Rechte erhält.

Zeitabhängige Geocodierung  
(meist manuelle Korrektur  
automatischer Geocodierung):

z.B. ein Kloster, das nach dem  
Dreißigjährigen Krieg seinen  
Standort verlagerte.



# FORMATE FÜR INTEROPERABLE GEODATEN

Geodaten können nicht einfach als Listen in WORD-Dateien gesammelt werden.

Auch die Erfassung in Tabellen ist meist nur der erste Schritt zur weiteren Aufbereitung. **Weit verbreitete Formate für Geodaten:**

## Shapefiles

(entwickelt von ESRI für ArcView):

- .shp (Geometriedaten)
- .dbf (Attributdaten)
- .shx (Index zur Verknüpfung)

## GeoJSON

(seit 2015 standardisiert durch die "Geographic JSON working group" der "Internet Engineering Task Force"):

- .geojson (oder .json, gemeinsame Erfassung von Geometrien und Attributen in einer Datei)



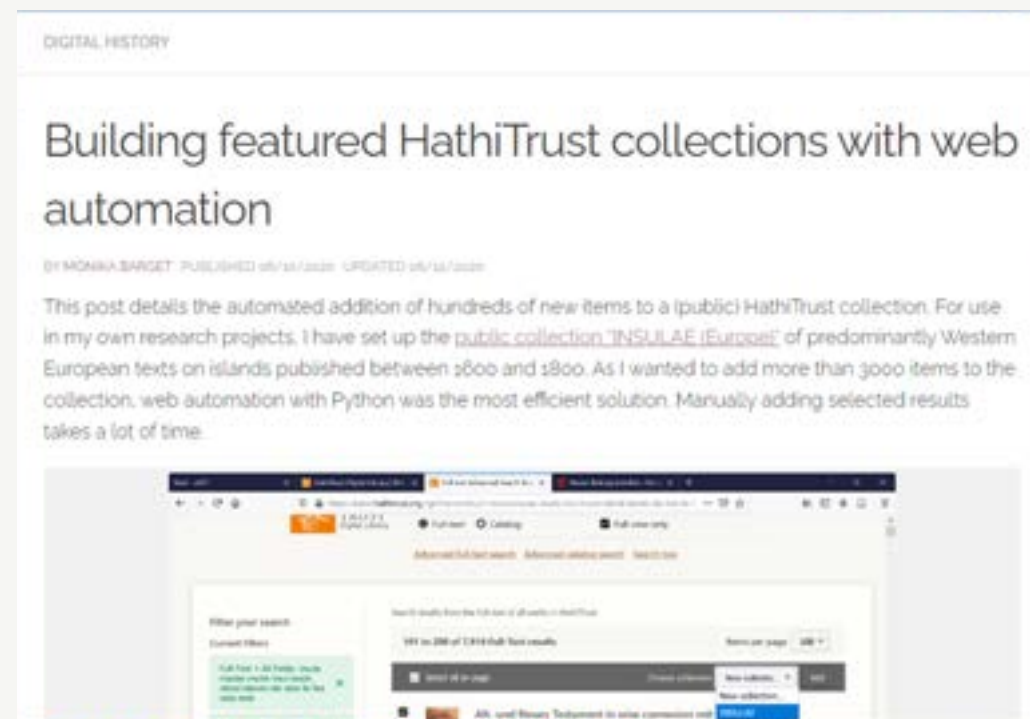
# EXTRAKTION VON GEODATEN

## Auswertung von APIs

```
<change>
  <date type="creation" n="1432309523" value="2015-05-22"/>
  <date type="lastmod" n="1587714958" value="2020-04-24"/>
</change>
</revisionDesc>
<teiHeader>
<text>
<body>
<div>
<listPerson>
  <person xml:id="cd515562-be2e-40a0-8427-ac3179667e0d">
    <persName>
      <roleName type="honorific" full="abb">Prof. Dr. med.</roleName>
      <forename>Friedrich</forename>
      <addName>Wilhelm</addName>
      <surname>Brussatis</surname>
      <reg>Brussatis, Friedrich</reg>
    </persName>
    <sex value="1">male</sex>
    <birth when="1919-04-01">
      <date>01.04.1919</date>
      <placeName>
    </birth>
    <death when="1989-01-26">
      <date>26.01.1989</date>
      <placeName>
    </death>
    <faith>0</faith>
  </listPerson>
  <listEvent>
    <event role="from" to="to">
      <label>
    </label>
    <note>
    </note>
  </listEvent>
</div>
</body>
</text>
</teiHeader>
</change>
```

Beispiel biographischer Daten aus  
<http://gutenberg-biographics.ub.uni-mainz.de/api/items/persons/>

## Webscraping (e.g. Metadata Harvesting)



Blogpost zur automatisierten Erstellung von Quellensammlungen in HathiTrust:  
<https://insulae.hypotheses.org/169>

## Eigene Datenstrukturierung (z.B. in EXCEL Tabellen)

	Factoid ID	pers ID	pers name	narrative	event type	after-ot	before-vent	start-vent	end-vent	datpers	titlers	functi	place	namst	name	rel	pers	ce	quotati
2	1113	Jahns_00-P-060	Aegidius Joseph Va	Geburt			1749-10-0	1749-10-09				Kind	Mons im Hennegau						
3	1114	Jahns_00-P-060	Aegidius Joseph Va	Taufe			0000-00-0	0000-00-00					katholisch						
4	1119	Jahns_00-P-060	Aegidius Joseph Va	Primäre Bildung	station		1759-00-0	1766-00-00					Wetzlar		Jesuitengymnasium Wetzlar				
5	1116	Jahns_00-P-060	Aegidius Joseph Va	Rezeption			1795-01-0	1795-01-02							Fränkische Reichsritterschaft, Ka				
6	1117	Jahns_00-P-060	Aegidius Joseph Va	Rezeption			1803-09-2	1803-09-20							Breisgauische Reichsritterschaft				
7	1120	Jahns_00-P-060	Aegidius Joseph Va	Immatrikulation			1766-10-0	1766-10-09					logices ca Würzburg						
8	1121	Jahns_00-P-060	Aegidius Joseph Va	Immatrikulation			1769-11-1	1769-11-17					Heidelberg						
9	1126	Jahns_00-P-060	Aegidius Joseph Va	Immatrikulation			1771-12-1	1771-12-10					IKG-Prak Wetzlar		Reichskammergericht				
10	1121	Jahns_00-P-060	Aegidius Joseph Va	Studium			1766-00-0	1766-00-00					stud. phil Würzburg						"ebd. 1766-1
11	1122	Jahns_00-P-060	Aegidius Joseph Va	Studium			1768-00-0	1769-00-00					stud. iur Würzburg						"1769-1769 ei
12	1124	Jahns_00-P-060	Aegidius Joseph Va	Studium			1769-00-0	1770-00-00					stud. iur. Heidelberg						
13	1125	Jahns_00-P-060	Aegidius Joseph Va	Praktikum			1771-00-0	1772-00-00					IKG-Prak Wetzlar		Reichskammergericht				"hörte 'colleg
14	1127	Jahns_00-P-060	Aegidius Joseph Va	Praktikum			1772-00-0	1775-00-0	C	Startdatum unsiche			Praktikan Regensb.		Reichstag				"1772 od. 177
15	1128	Jahns_00-P-060	Aegidius Joseph Va	Praktikum			1773-00-0	1775-00-0	C	Startdatum unsiche			Praktikan Regensb.		Reichstag				"1772 od. 177
16	1133	Jahns_00-P-060	Aegidius Joseph Va	Praktikum			1773-00-0	1775-00-00					Praktikan Regensb.		Reichstag				"Konzipit (Se
17	1129	Jahns_00-P-060	Aegidius Joseph Va	Promotion			1783-05-2	1783-05-21					Dr. iur.		Freiburg				"vermutl. Eht
18	1130	Jahns_00-P-060	Aegidius Joseph Va	Funktionsausübung			1767-00-0	1774-00-00					Donszell; Fritzlär		Petersstift				"possessio pr
19	1132	Jahns_00-P-060	Aegidius Joseph Va	Funktionsausübung			1773-00-0	1775-00-00					Sekretär Freiburg		kaiserialich königlich				"Konzipit (Se
20	1134	Jahns_00-P-060	Aegidius Joseph Va	Funktionsausübung			1773-00-0	1782-00-00					kaiserialich Freiburg						
21	1135	Jahns_00-P-060	Aegidius Joseph Va	Funktionsausübung			1775-00-0	1782-00-00					kaiserialich Freiburg						

Übertragung der Biographien  
frühneuzeitlicher Amtsträger aus Mainz in  
ein Tabellenformat (Factoid-Modell) zur  
weiteren semi-manuellen Bearbeitung

# BEISPIEL: EXTRAKTION VON ORTSDATEN AUS GEDRUCKTEN KARTEN MIT DEM OCR-WERKZEIT TRANSKRIBUS LITE

The screenshot displays the Transkribus Lite web application interface. At the top, the navigation bar includes the Transkribus Lite logo, links for Home, Jobs, and Info, and a user profile for m.barget@maastrichtuniversity.nl with options for Credits and EN. Below the navigation bar, the breadcrumb trail shows 'Collections > MAPS > Topographia1646\_maps > Page 6'. The main content area is split into two panels. The left panel shows a historical map with various place names and a red rectangular box highlighting a specific area. The right panel displays the OCR results for the highlighted area, showing a list of place names: Ludesdorff / Wied / Isenberg, Delem, Waldmanshausen, Hammerstein, Hollach, Stockhausen / LHON, Andernach / Enger, Hademer, Leune, and Montabour. The interface also includes a search bar with 'IN\_PROGRESS | m.barget@maastrichtuniversity.nl', a toolbar with icons for annotation, zoom, and navigation, and a vertical 'Feedback' button on the right side.

Transkribus  
lite

Home Jobs Info

m.barget@maastrichtuniversity.nl Credits EN

< 6 6 Go > Collections > MAPS  
Topographia1646\_maps > Page 6

IN\_PROGRESS | m.barget@maastrichtuniversity.nl

Annotation A+ A+ In Progress

Ludesdorff / Wied / Isenberg  
Delem

Waldmanshausen  
Hammerstein  
Hollach  
Stockhausen / LHON  
Andernach / Enger  
Hademer  
Leune  
Montabour

Feedback



3) Mit welchen digitalen Werkzeugen kann man (historische) Geodaten bearbeiten?



```

Tab-delimited-text_Palladio.txt - Editor
Datei Bearbeiten Format Ansicht Hilfe
Places "Coordinates"
Mainz "50.0012314, 8.2762513"
Meißen "51.1630871, 13.4704939"
Wiesbaden "50.0820384, 8.2416556"
Köln "50.938361, 6.959974"
Paris "48.8588897, 2.320041"
Bonn "50.735851, 7.10066"
Kopenhagen "55.6867243, 12.5700724"
Eltville "50.0559826, 8.0891474"
Maastricht "50.8579855, 5.6969882"
Lanaken "50.8892784, 5.6513208"
Heerlen "50.8775239, 5.9815066"
Aachen "50.776351, 6.083862"

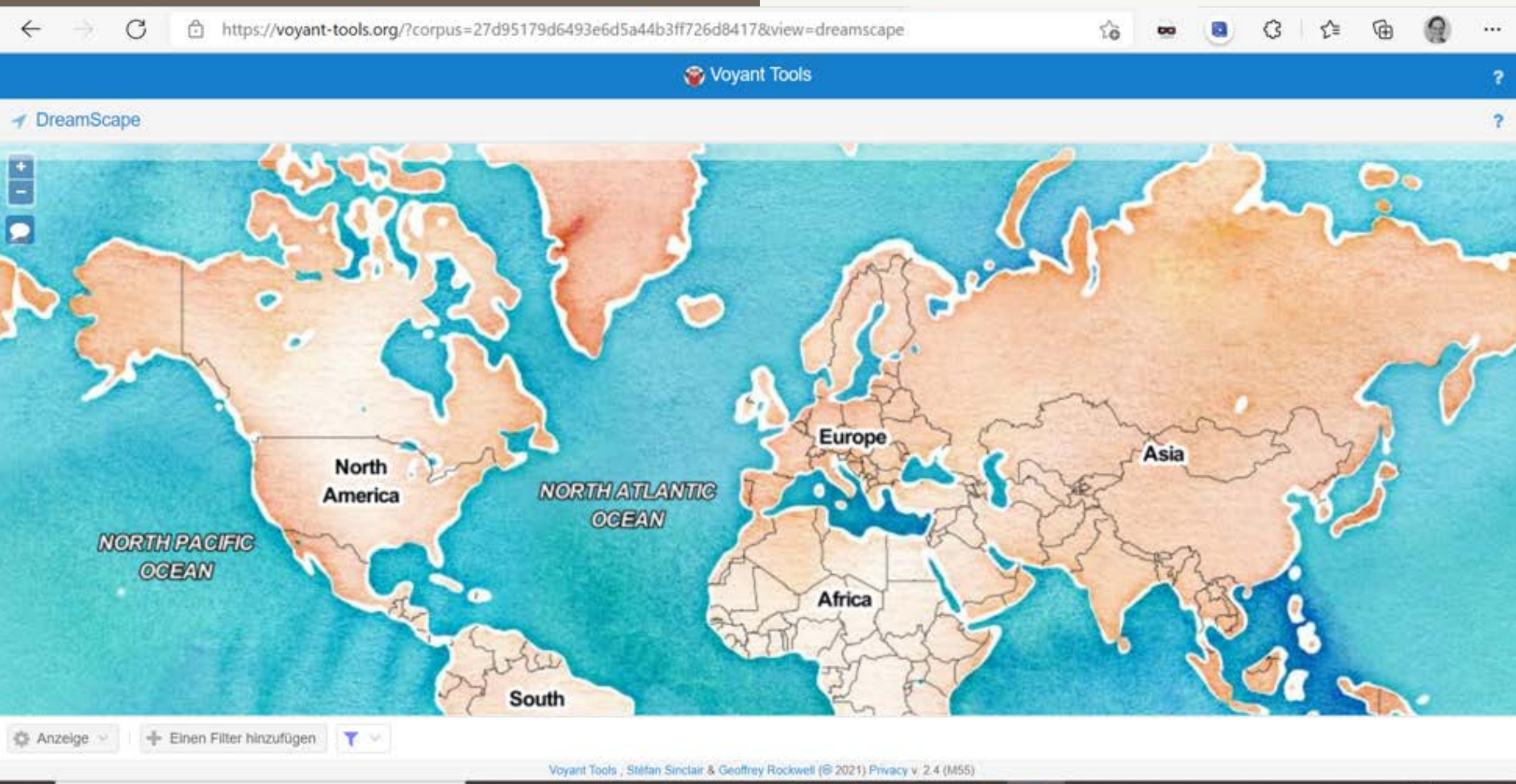
```

# PALLADIO (BROWSER)

Palladio (Stanford University) erlaubt die Erstellung einfacher Karten und Netzwerkdiagramme. Die Daten sollten idealerweise in einem "tab-delimited" Textformat eingegeben werden. Die Möglichkeiten der Visualisierung sind begrenzt. Ein hochwertiger Export statischer Karten ist nicht möglich.







Voyant Tools ist eine ständig wachsende Sammlung von Textanalysetools. Die Daten können als TXT, XML oder auch PDF (mit eingebettetem OCR) hochgeladen werden.

Das "Dreamscapes" Werkzeug verspricht die automatische Erkennung von Ortsnamen in Texten und ihre Darstellung auf einer Karte.

Die Ergebnisse sind sehr durchwachsen. Oft werden keine Orte erkannt.

# VOYANT DREAMSCAPES (BROWSER)





# GEOJSON.IO (BROWSER)

Semi-manuelle Bearbeitung  
von Geodaten im GeoJSON  
Format auf geojson.io

"geojson.io is a tool for editing GeoJSON data on the internet. It enables editing through a map interface, raw GeoJSON, and exporting and importing a large number of formats. geojson.io is a tool for editing GeoJSON data on the internet."

(GeoJSON Webseite am 12. Oktober 2020)

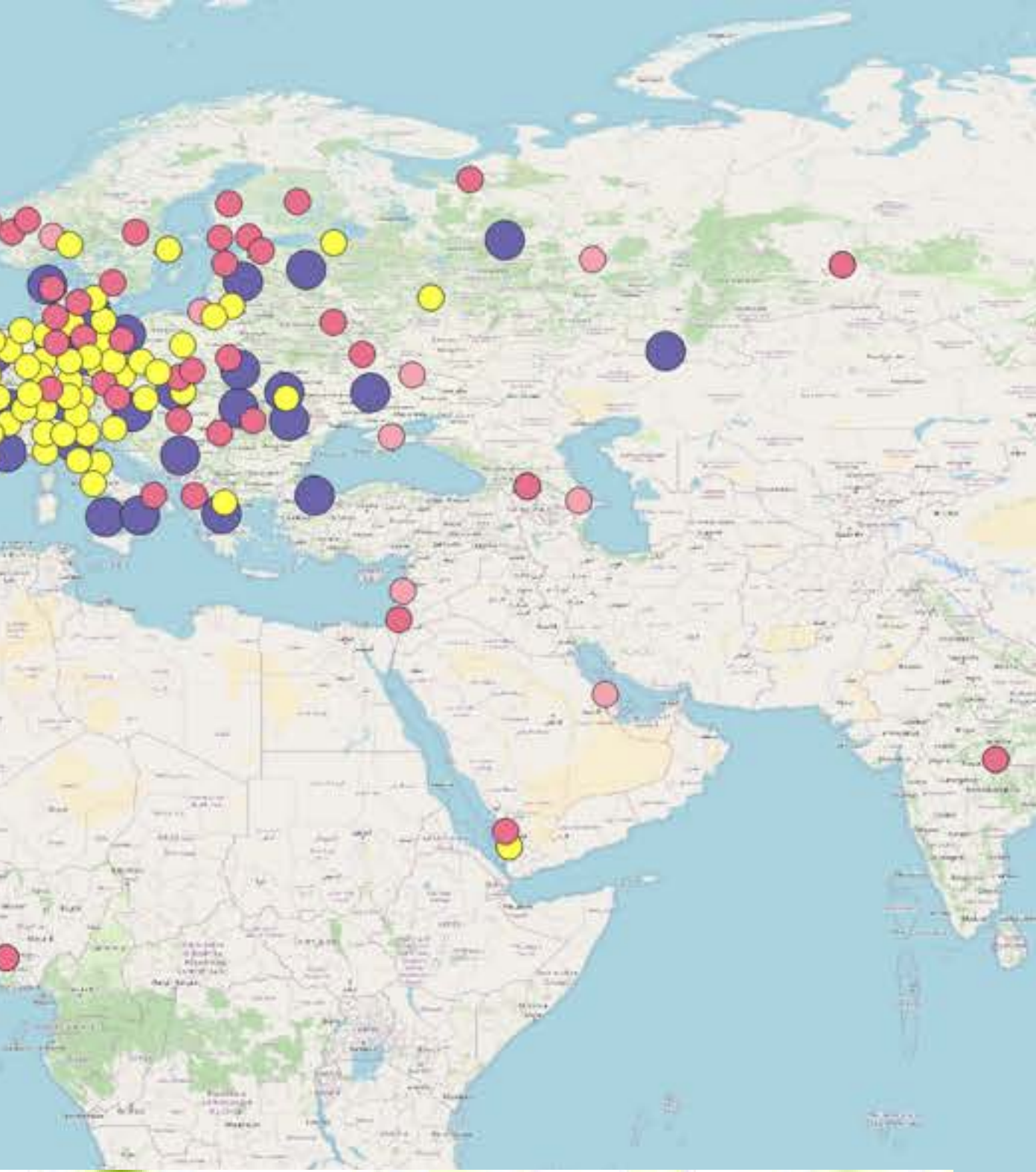
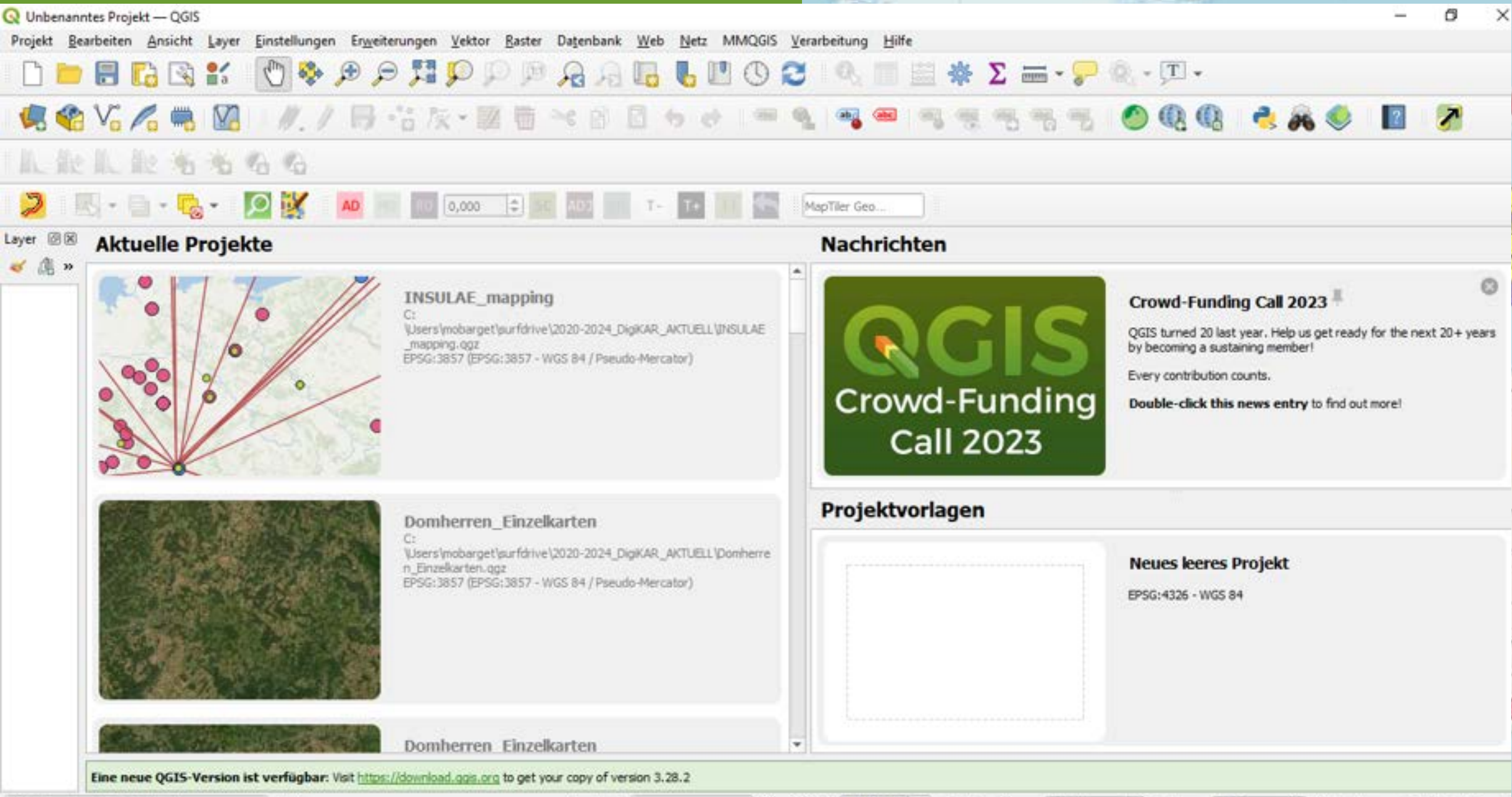
Tutorial:

[https://monikabarget.github.io/GeoHumTutorials/Tutorial\\_GeoJSON](https://monikabarget.github.io/GeoHumTutorials/Tutorial_GeoJSON)

The screenshot shows the geojson.io web application. The browser address bar displays the URL: `https://geojson.io/#map=11/42.6908/18.0114`. The map interface shows a geographical area with a red pin on Dubrovnik and a blue pin on Cavtat. The right sidebar shows the JSON editor with the following code:

```
1 {
2   "type": "FeatureCollection",
3   "features": [
4     {
5       "type": "Feature",
6       "properties": {
7         "marker-color": "#f30a6d",
8         "marker-size": "medium",
9         "marker-symbol": "",
10        "NAME": "Dubrovnik"
11      },
12      "geometry": {
13        "type": "Point",
14        "coordinates": [
15          18.0889892578125,
16          42.65416193033991
17        ]
18      }
19    },
20    {
21      "type": "Feature",
22      "properties": {
23        "marker-color": "#2440d9",
24        "marker-size": "medium",
25        "marker-symbol": ""
26      }
27    }
28  ]
29 }
```

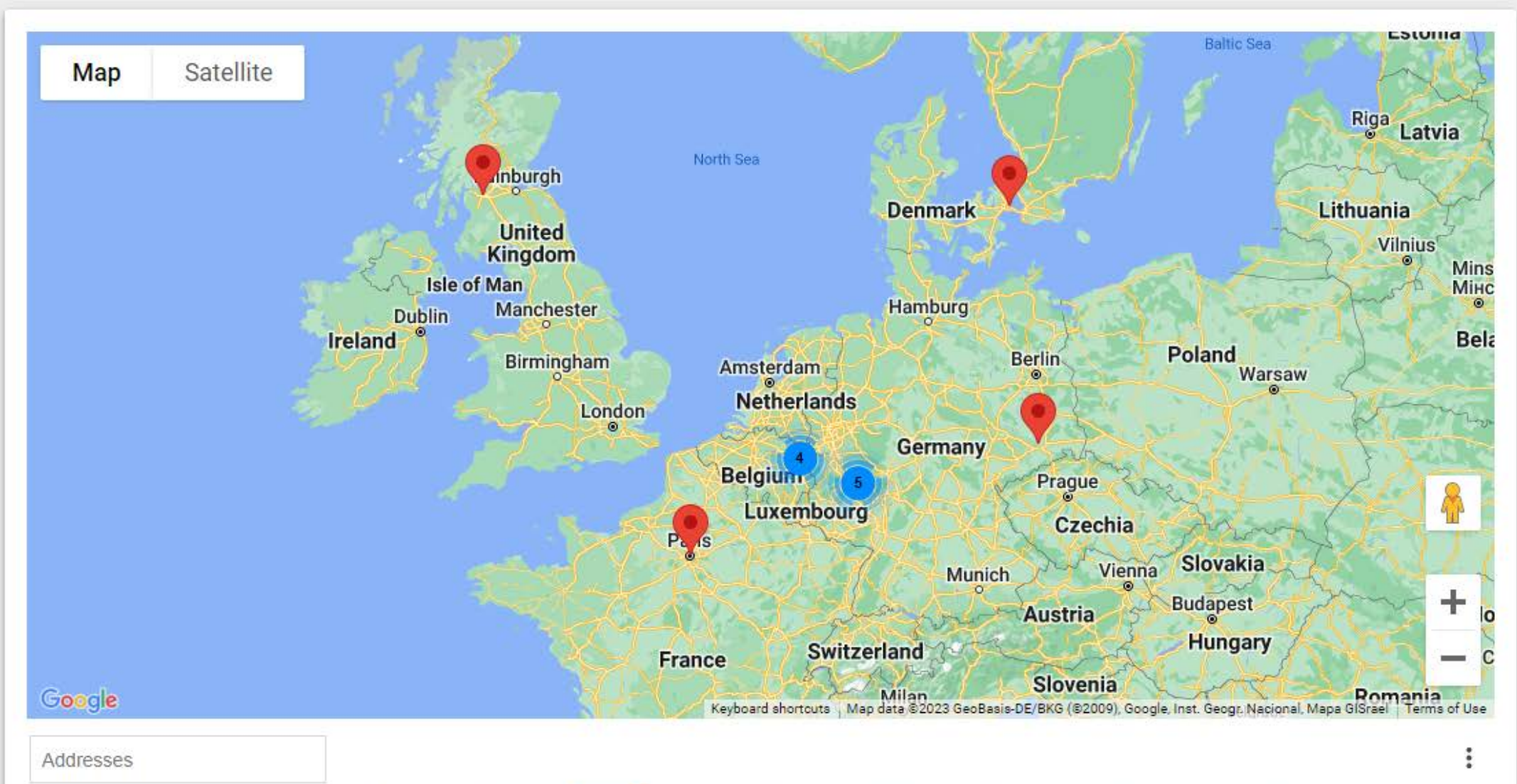




GIS SOFTWARE  
(E.G. QGIS)







# GOOGLE AWESOME



Geocode is a tool that helps you get latitudes & longitudes from addresses in a Google Sheet to display them on a map you can share.

By: Talarian

Listing updated: January 10, 2023





# GEOCODIERUNG & PLOTTING MIT SKRIPTEN (SH. COLAB)



```
github.com/MonikaBarget/GeoHumTutorials/blob/master/Colab_Geocoding/Geocode_Plot_Geonames_interactiveMAP.ipynb

Geocoder for Python supports the following ones:

  • (geocoding) retrieve GeoNames's geocoded data from a query string, and
  various filters
  • (details) retrieve all geonames data for a given geonames_id
  • (children) retrieve the hierarchy of a given geonames_id
  • (hierarchy) retrieve all children for a given geonames_id

Full documentation: https://geocoder.readthedocs.io/providers/GeoNames.html

The first step is to get COLAB working:

In [ ]: ## mount drive
from google.colab import drive
drive.mount("/content/drive")

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

A file path needs to be defined for storing input or output files linked with this script:

In [ ]: directory="/content/drive/My Drive/Colab_Geohumanities/"

Now we can install packages that are not part of Python's standard distribution but are necessary for geocoding and plotting maps. There will most likely be a dependency error for NumPy, but the script should still work.

In [ ]: ## install packages that are not part of Python's standard distribution

!pip install geocoder
!pip install basemap
!pip install ipyleaflet
!pip install geojson
```

4) Interaktives Arbeiten mit  
ausgewählten Werkzeugen





# TEST 1: GEOCODIERUNG MIT GOOGLE AWESOME

The screenshot shows a Google Sheets spreadsheet with a list of cities in column A. The 'Extensions' menu is open, and 'Geocode by Awesome Table' is selected. The sidebar on the right shows the configuration for the geocoding process.

	A	B
1	Addresses	
2	MapsTooltip - Hidden - StringFilter	
3	Mainz	
4	Meißen	
5	Wiesbaden	
6	Köln	
7	Paris	
8	Bonn	
9	Glasgow	
10	Kopenhagen	
11	Eltville	
12	Maastricht	
13	Lanaken	
14	Heerlen	
15	Aachen	
16		
17		
18		
19		
20		
21		
22		
23		

**Geocode**

Geocode gets latitudes and longitudes from **full addresses** and creates a customizable map.

**Current sheet**  
Tabelle1

**Address column**  
Addresses

Are your addresses in multiple columns?

Try wider results

**Map**

The map has been created. If you click on a marker in the map, the tooltip will simply display the address. But you can do much better and display any info you want in this tooltip.

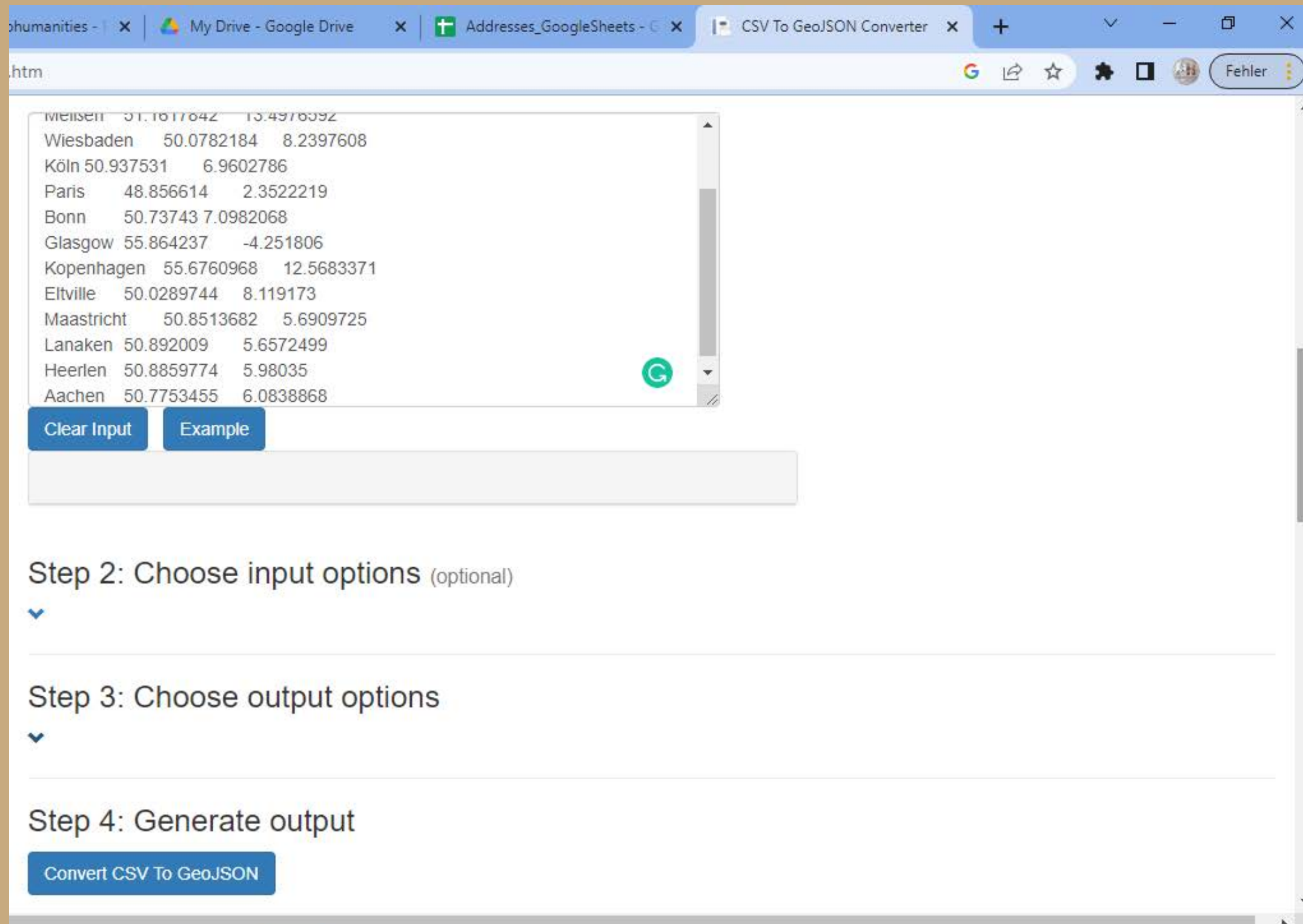
Customize Tooltip Open map

**Geocode!**

Proudly brought to you by [Awesome Table](#)

# CSV TO GEOJSON:

[HTTPS://WWW.CONVERTCSV.COM/CSV-TO-GEOJSON.HTM](https://www.convertcsv.com/csv-to-geojson.htm)



The screenshot shows a web browser window with the URL <https://www.convertcsv.com/csv-to-geojson.htm>. The page contains a text area with a list of cities and their coordinates, a 'Convert' button, and a 'Clear Input' button. Below the text area, there are four steps: Step 2: Choose input options (optional), Step 3: Choose output options, and Step 4: Generate output. A 'Convert CSV To GeoJSON' button is located at the bottom of the page.

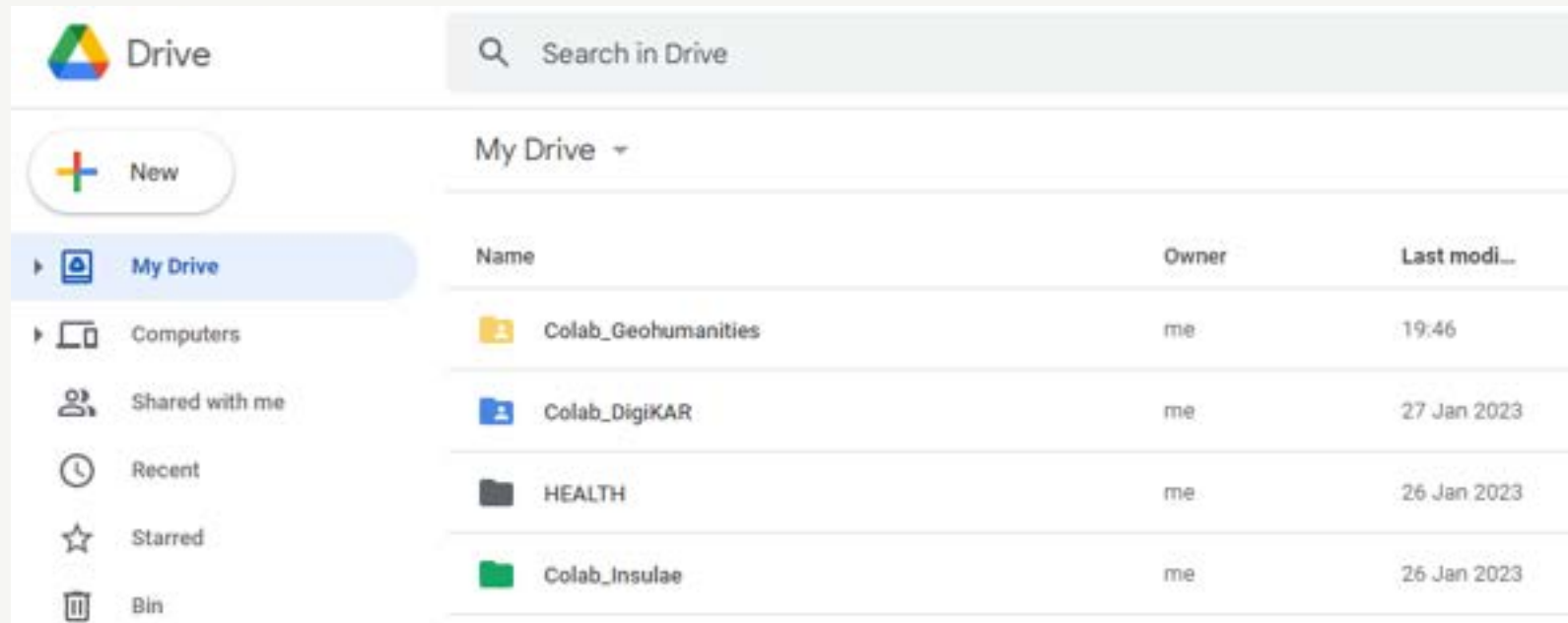
Meisen	51.1617642	13.4976392
Wiesbaden	50.0782184	8.2397608
Köln	50.937531	6.9602786
Paris	48.856614	2.3522219
Bonn	50.73743	7.0982068
Glasgow	55.864237	-4.251806
Kopenhagen	55.6760968	12.5683371
Eitville	50.0289744	8.119173
Maastricht	50.8513682	5.6909725
Lanaken	50.892009	5.6572499
Heerlen	50.8859774	5.98035
Aachen	50.7753455	6.0838868

Dieses Online Werkzeug erlaubt die Erstellung von GeoJSON Dateien aus Tabellen, die an sich nicht von GIS Software gelesen werden können.

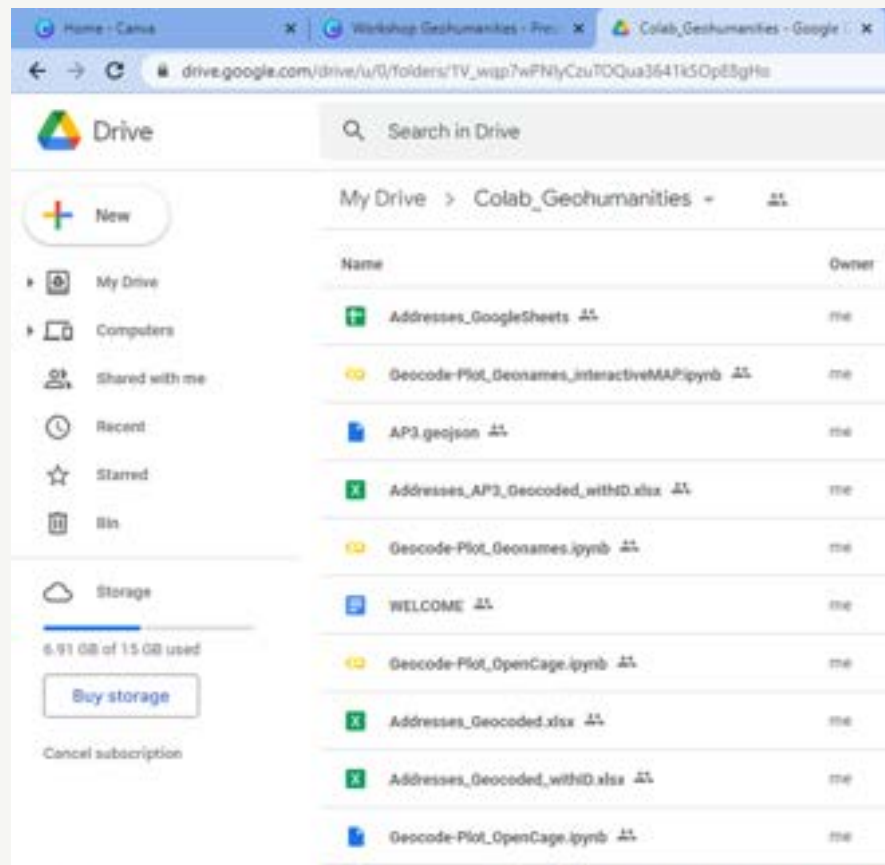
Allerdings ist die Erstellung im graphischen Interface aufwendig. Skripte erledigen diese Aufgabe schneller.



# TEST 2: GEOCODIERUNG MIT PYTHON VIA GEO APIS



Synchronisierten Ordner kopieren oder virtuellen Ordner herunterladen und als eigenen Ordner mit neuem Namen (!) zum Drive hinzufügen.



Aus eigenem Ordner Skripte im Jupyter Notebook Format öffnen.

Pfad zum eigenen Ordner im Skript ergänzen und Code ausführen:

```
DIRECTORY="/CONTENT/  
DRIVE/MY DRIVE/  
COLAB_GEOHUMANITIES/"
```



RUN



OK

The screenshot shows a Google Colab notebook interface. The browser address bar displays the URL: `colab.research.google.com/drive/13nCyb2p2KU2YZUBr7oN9jcylyX4HcpBR`. The notebook title is `Geocode-Plot_Geonames_interactiveMAP.ipynb`, with a star icon and the text "Last edited on 26 January". The notebook content includes:

- A section titled **GeoNames.** with the text: "GeoNames is mainly using REST APIs. It offers 40 different webservices."
- A section titled **Geocoder** for Python supports the following ones:
  - (geocoding) retrieve GeoNames's geocoded data from a query string, and various filters
  - (details) retrieve all geonames data for a given geonames\_id
  - (children) retrieve the hierarchy of a given geonames\_id
  - (hierarchy) retrieve all children for a given geonames\_id
- Full documentation: <https://geocoder.readthedocs.io/providers/GeoNames.html>
- The first step is to get COLAB working:
- A code cell with a play button icon containing:

```
## mount drive
from google.colab import drive
drive.mount("/content/drive")
```
- A message: "Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force\_remount=True)." with a copy icon.
- A text prompt: "A file path needs to be defined for storing input or output files linked with this script:"
- A code cell with a list icon containing:

```
[ ] directory="/content/drive/My Drive/Colab_Geohumanities/"
```



# THANKYOU

DigiKAR Webseite: <https://digikar.eu/>

(inklusive Links zu Vorträgen, Daten, Blogposts und anderen Veröffentlichungen)

Geohumanities Repository:

<https://monikabarget.github.io/GeoHumTutorials>

Forschungsblog zum Insel-Projekt (und anderen DH Themen):

<https://insulae.hypotheses.org/>

M

[m.barget@maastrichtuniversity.nl](mailto:m.barget@maastrichtuniversity.nl)

