



University of Pittsburgh

# The World Historical Gazetteer: Digital Public and World History

Susan Grunewald, Ph.D.  
Digital History Postdoctoral Associate  
World History Center  
University of Pittsburgh  
@SusanGrunewald1  
[scg52@pitt.edu](mailto:scg52@pitt.edu)





# WHAT IS A GAZETTEER?



University of Pittsburgh

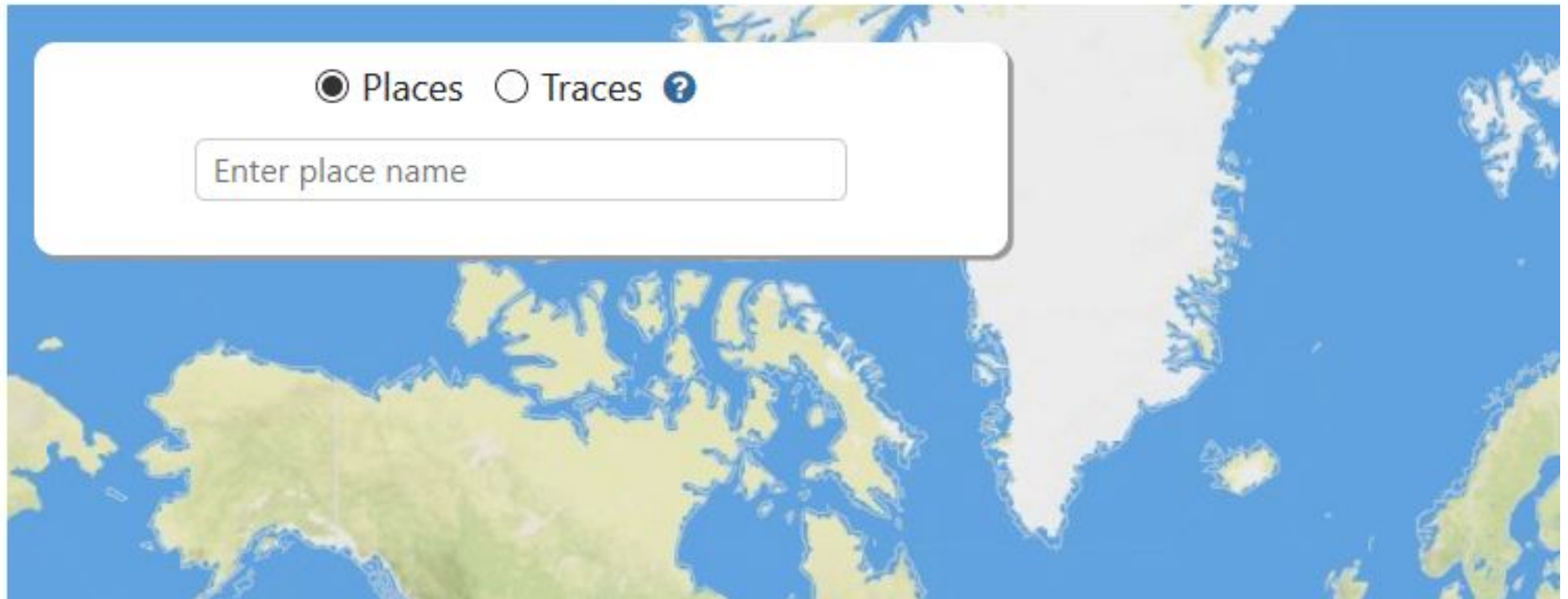
# OVERVIEW OF WHG PROJECT



# World-Historical Gazetteer

**BETA 0.2**

[walkthrough](#) | [feedback](#)





## Aggregating contributed **place** and **trace** data

### Place records

Records of references to place (toponyms, ethnonyms) from historical sources:

- texts of all kinds
- tabular records
- print gazetteers
- old maps



**Linked Places format**

### Trace annotations

Records of historical entities of any kind for which setting (location at time) is of interest, annotated with IDs for relevant places

- people
- events
- works



**Linked Traces annotation format**



● Places ○ Traces ?

Byzantium

SEARCH RESULTS (1)

**Byzantium** (nation) [TR] ●

var: *Istanbul, Constantinople, Istanbul*;  
,*istæn'bu:l/* or */i'stænbʊl/*; Turkish: *İstanbul* [*is'tanbuʃ* ] ,  
historically known as Constantinople and *Byzantium*...

100 km

Tiles © MapBox | CC-BY-NC 3.0 | Tiles and Data © 2013 AWMC



## Istanbul

### Attestations

Istanbul is a { place type }, { on the {...} }

**Istanbul** (pid:85196) dataset: [black](#)

Variants: **Istanbul** | **Byzantium**

Types: *settlement*

Links: [dbp:Istanbul](#) (closeMatch) ; [tgn:7002473](#) (closeMatch)

Temporal: 1947/1968 | 1945/1952 | 1914/1918 | 1939/1942 | 1942/1945

All t [names,geoms,types,whens]: [], [], [], [[1914, 1968]], [], [], [], [[-700, 1923]], [], [], [], [[-700, 642]]

**Constantinople** (pid:83140) dataset: [black](#)

Variants: **Constantinople** | **Byzantium**

Types: *settlement* | *battle*

Links: [dbp:Constantinople](#) (closeMatch) ; [tgn:7002473](#) (closeMatch)

Temporal: 500/500 | 1310/1360 | 1100/1300 | 1096/1270 | 1090/1090 | 600/800 | 800/1200 | 1100/1400 | 1840/1910 | 1815/1852 | 1400/1860 | -700/120 | 1325/1346 | 1200/1700 | 1860/1920 | 300/500 | 1870/1910 | 500/1500 | 622/627 | 750/809 | 762/800 | 400/650 | 800/1000 | 1204/1300 | 1200/1400 | 789/1100 | 1250/1300 | 1225/1350 | 1794/1815 | 1792/1815 | 1453/1571 | 1300/1500 | 1512/1639 | 1800/1913 | 1040/1100 | 1379/1405 | 313/642 | 400/800 | 1914/1918 | 1914/1918 | 1918/1923

All t [names,geoms,types,whens]: [], [], [], [[1914, 1968]], [], [], [], [[-700, 1923]], [], [], [], [[-700, 642]]

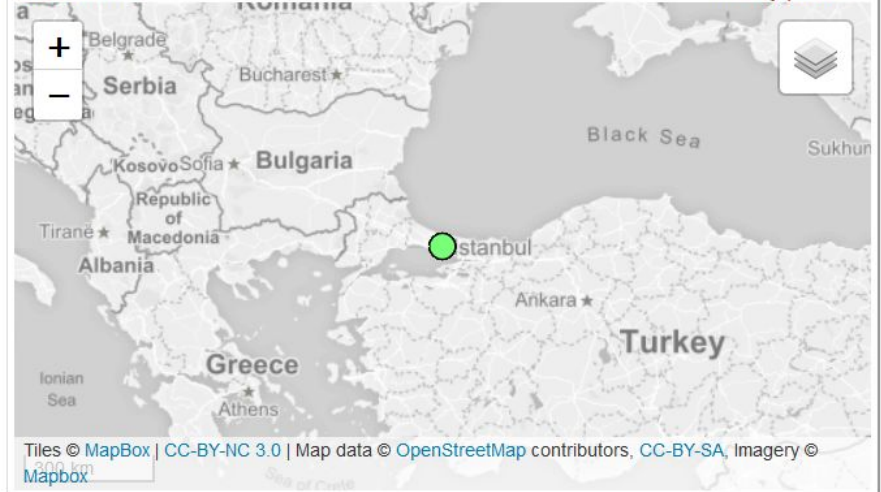
**Byzantium** (pid:82439) dataset: [black](#)

### TEMPORAL ATTESTATIONS

{ summing temporal visualization } *future*

### GEOGRAPHY

nearby places



### TRACES

- { an event trace }
- { a person trace }
- { a works trace }



○ Places ● Traces ?

**Gautama Buddha**

*PLACES (9)*

- [Lumbini](#) (birthplace)  
*earliest: -0563, latest: -0480*
- [Kusinagara](#) (deathplace)  
*earliest: -0483, latest: -0400*
- [Magadha](#) (taught)
- [Koshala](#) (taught)
- [Kapilavastu](#) (resided)
- [Rajagriha](#) (resided)
- [Bodh Gaya](#) (enlightened)  
*earliest: -0528, latest: -0445*
- [Sarnath](#) (resided)
- [Indo-Gangetic Plain](#) (resided)

500 km

Tiles © MapBox | CC-BY-NC 3.0 | Tiles and Data © 2013 AWMC





## Lumbini

### Attestations

Lumbini is a { place type }, { on the {...} }

**Lumbini** (pid:86438) dataset: [black](#)

Variants: **Lumbini**

Types: *settlement*

Links: [dbp:Lumbini](#) (closeMatch) ; [tgn:7032815](#) (closeMatch)

Temporal: -365/-181 | -1550/-200 | -900/600

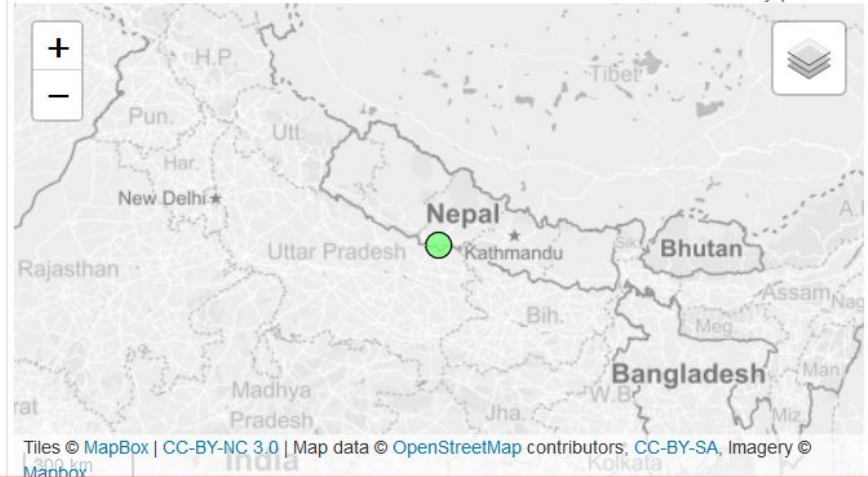
All t [names,geoms,types,whens]: [[], [], [], [[-1550, 600]]]

### TEMPORAL ATTESTATIONS

{ summing temporal visualization } future

### GEOGRAPHY

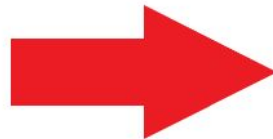
[nearby places](#)



Tiles © MapBox | CC-BY-NC 3.0 | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

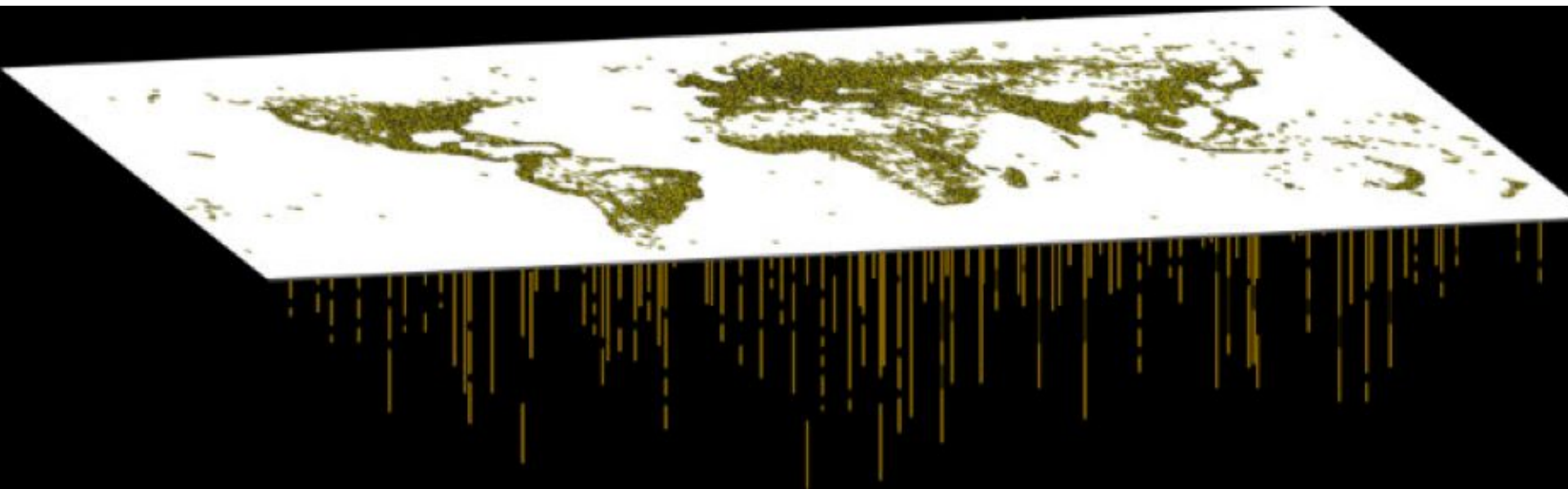
### TRACES

- **birthplace** ['Person']; [Gautama Buddha](#) ; when: c.-0563 - c.-0480
- **waypoint** ['Journey', 'Text']; [Xuanzang](#) : a Buddhist pilgrim on the Silk Road ; when: 0636





University of Pittsburgh





University of Pittsburgh

# TECHNICAL SPECIFICATIONS



cities w/pop > 500k; capitals

*links*

**WHG Core**



geonames 500

**TGN**

geonames 462 more

geonames

wikidata

ne mountains

loc

viaf

ne rivers

geonames, wikidata, viaf

ne lakes

wikidata

Getty TGN (less mountains, rivers, lakes)

**TGN**

Black atlas

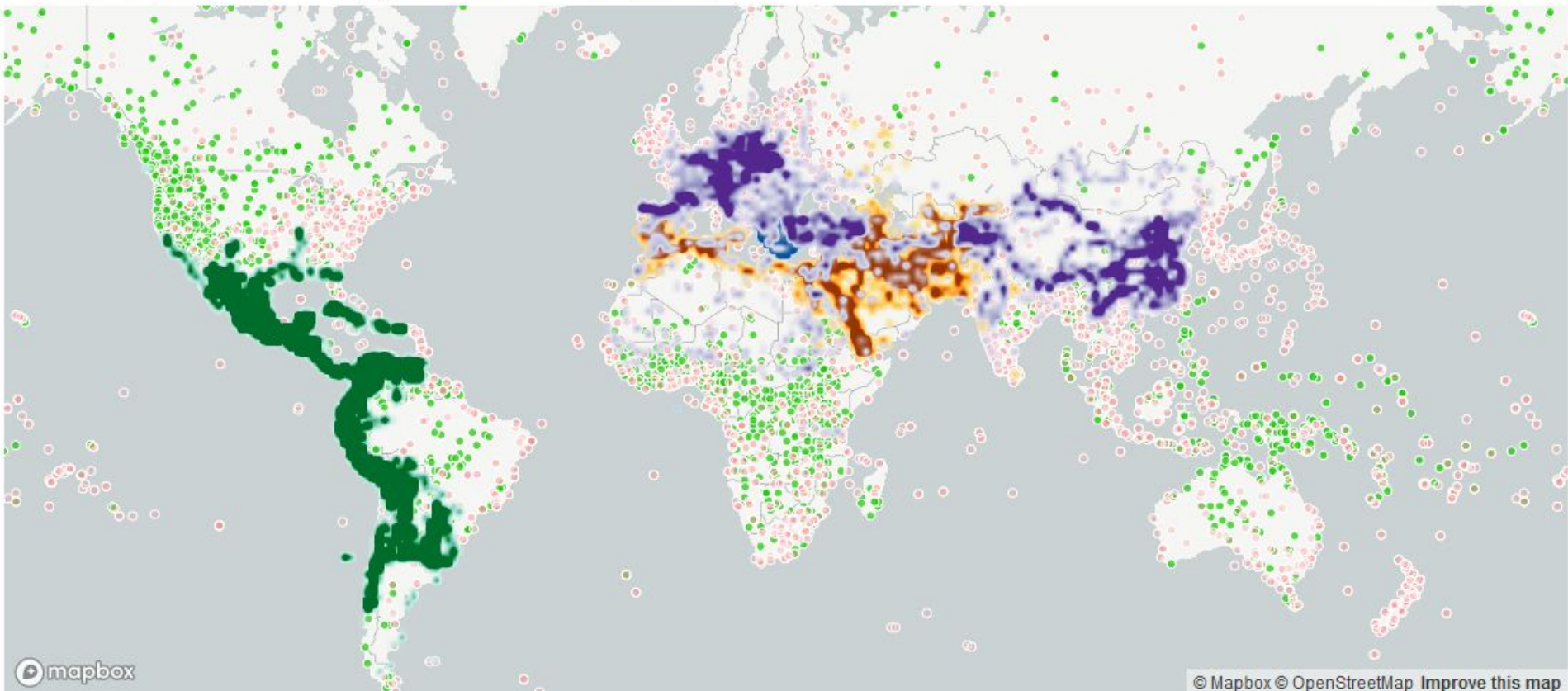
D-Place

**TGN, geonames, wikidata, loc, viaf**



## Dataset Maps (heat maps of datasets in queue)

Select dataset





## Linked Places Format

- Attempt to create standard for place data
- It is both valid JSON-LTD and GeoJSON
- Permits names, time, geometries, and relations



## Linked Places Format

- LP-TSV: tab-delimited text file
- Requires name
- Permits country codes, types
- Geometry



## Linked Traces Format

- Supports place name
- Supports temporal range
- Supports type of relation (waypoint, person, artifact, etc)





## Linked Traces TSV

- Similar formatting to Linked Places TSV
- Data must be standardized to our specifics to upload into system



University of Pittsburgh

**CONTRIBUTING**



**BETA 0.2**

[walkthrough](#) | [feedback](#)

[Maps](#)

[API](#)

[Tutorials](#)

[A](#)



[Register](#)

[login](#)

Places  Traces [?](#)

Enter place name





Places  Traces

Enter place name



3000 km



### Datasets



	id	name	label	uploaded	format	# rows	status	owner
	2	<a href="#">D-Place</a>	dplace (core)	11-Feb-2019, 07:00 (UTC)	direct	1428	reconciling	whgadmin
	1	<a href="#">DK Atlas of World History</a>	black (core)	10-Jan-2019, 07:00 (UTC)	direct	10643	reconciling	whgadmin

### Study Areas



id	title	type	ccodes
----	-------	------	--------

### Collections



User collections of place and trace records (*future*)



## Upload dataset

Name

Unique label  20 char max, no spaces

Description

Data type

Format  ⓘ

URI base

No file selected.

### Sample data

During this beta phase, some sample datasets are available from our GitHub repository. Download [this zip file](#), check the README.md file, and you can experiment with loading and reconciling them from your local computer.



## Dataset Portal

[Browse data](#) [Delete dataset](#)

**Name:** [Alcedo](#)

**Description:** Upload of Alcedo TSV.

**IDs** (label/number): [AlcedoTest/358](#)

**File:** user\_grunewas/alcedo\_200.tsv

**Date:** 30-Sep-2019, 18:09 (UTC)

**Total rows:** 201

**Data type:** place

**URI base:** <http://whgazetteer.org/api/places/>

**Format:** delimited (tab)

**Status:** format\_ok

**Columns:** ['id', 'title', 'title\_source', 'types', 'aat\_types', 'parent\_name', 'matches']

Counts ( initial, **added** )

**Names:** 0

**Links:** 0, **0**

**Geometries:** 0, **0**

### Reconciliation Tasks





## Initiate Reconciliation: *Alcedo* (*AlcedoTest*)

### Authority name source

- Getty TGN (tgn)
- Wikidata (wd) *1+ seconds per record*
- DBpedia (dbp)
- GeoNames (gn)

### Geographic constraint

Select pre-defined region ▾ **or...**

Select user-defined study area ▾ [create user area](#)

[Start](#) [Cancel](#)  accept geometries in matches

Results summary will show here...







## Initiate Reconciliation: *Alcedo* (*AlcedoTest*)

### Authority name source

- Getty TGN (tgn)
- Wikidata (wd) *1+ seconds per record*
- DBpedia (dbp)
- GeoNames (gn)

{ TODO: progress bar }

**Authority:** tgn

**Task ID:** 839f36a3-5385-4b04-8bf5-c5252fc15fa6

**Result summary:** {'count': 200, 'got\_hits': 58, 'total': 216, 'pass1': 0, 'pass2': 0, 'pass3': 216, 'pass1remains': -1, 'pass2remains': -1, 'pass3remains': -1, 'no\_hits': {'count': 142}, 'elapsed': '00:03'}

[Return to Dataset](#)

[Review Hits](#)



Reconciliation Review: [AlcedoTest](#) > tgn

PASS 3 ▾

task id: 839f36a3-5385-4b04-8bf5-c5252fc15fa6

Record 1 of 58 [next](#) [last](#) »

## Abades



WHG place id: 227041

Source id: 00010

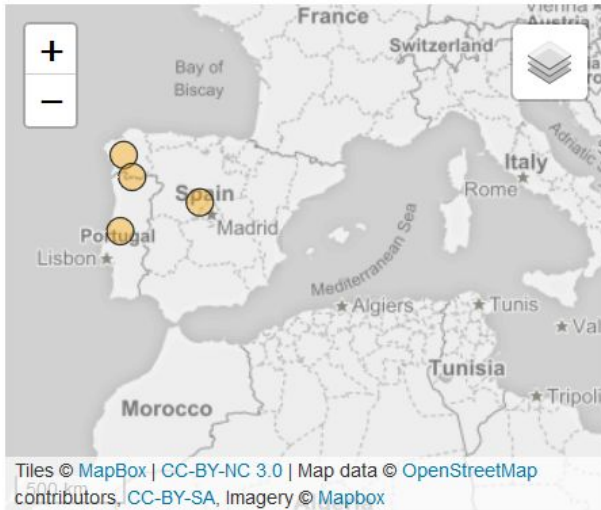
Name variants: Abades;

Relations: Popayan;

Place type(s): pueblo (towns);

Links: [indias:5600547](#)

Save



closeMatch  exactMatch  related  no match

Title/Preferred: **Abades**

TGN ID: 7348127

Variants: Abades;

Types: ['inhabited places (aat:300008347)']

Parents: Pontevedra > Galicia > España > Europe > World

closeMatch  exactMatch  related  no match

Title/Preferred: **Abades**

TGN ID: 7744661

Variants: Abades;

Types: ['inhabited places (aat:300008347)']

Parents: Santarém > Portugal > Europe > World

closeMatch  exactMatch  related  no match

Title/Preferred: **Abades**

TGN ID: 7348128

Variants: San Pelayo; Abades;

Types: ['inhabited places (aat:300008347)']

Parents: Ourense > Galicia > España > Europe > World

closeMatch  exactMatch  related  no match



**BETA 0.2**

[walkthrough](#) | [feedback](#)

[Maps](#)

[API](#)

[Tutorials](#)




[About WHG](#) ▾

[My Data](#)



[Logout](#)

### Datasets

	id	name	label	uploaded	format	# rows	status	owner
	358	<a href="#">Alcedo</a>	AlcedoTest	30-Sep-2019, 18:09 (UTC)	delimited	200	reconciling	grunewas
	2	<a href="#">D-Place</a>	dplace (core)	11-Feb-2019, 07:00 (UTC)	direct	1428	reconciling	whgadmin
	1	<a href="#">DK Atlas of World History</a>	black (core)	10-Jan-2019, 07:00 (UTC)	direct	10643	reconciling	whgadmin

### Study Areas

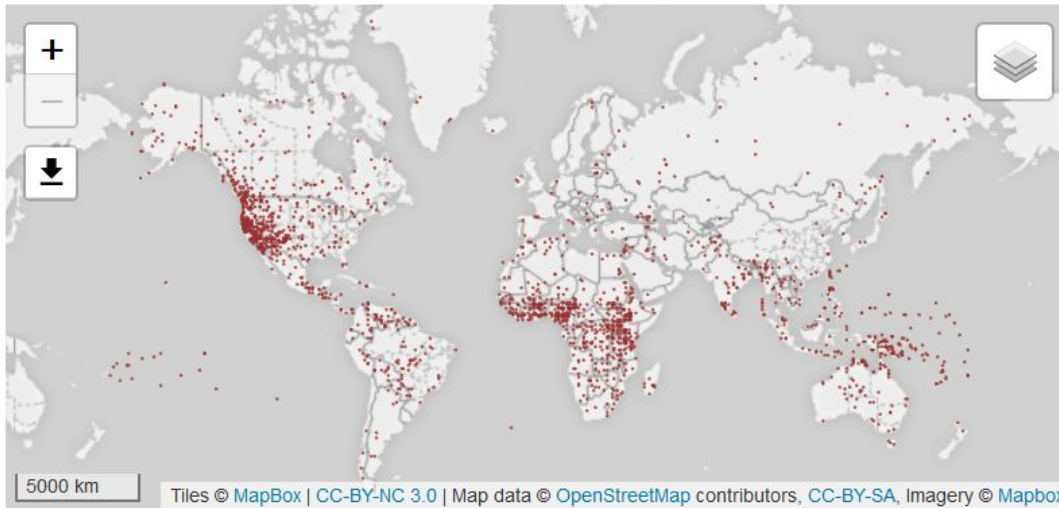
id	title	type	ccodes
----	-------	------	--------

### Collections

User collections of place and trace records (*future*)

**BETA 0.2**[walkthrough](#) | [feedback](#)

### D-Place ("dplace")

**Title:** !Kung**Variants:** Was Nyae; Kung Bushmen; !Kung (Was Nyae); !Kung;**Types:** (cultural group/aat:300387171)**Linked records:**closeMatch: [ [dbp:!Kung\\_people](#) ]

**Description:** The !Kung, also spelled !Xun, are a San people living in the Kalahari Desert in Namibia, Botswana and in Angola. They speak the !Kung language, noted for its extensive use of click consonants. The "!" in the name "!Kung" is a click that sounds something like a cork pulled from a bottle. However, th ...

Show 10 entries

Search: 

pid	src_id	title	ccodes
124653	xd1	!Kung	BW,NA
124654	xd10	Herero	NA
124655	xd100	Duruma	KE
124656	xd1000	Kuskowagmut	US
124657	xd1001	Syilx	CA,US
124658	xd1003	Round Lake Ojibwa	CA
124659	xd1004	North Albany Ojibwa	CA
124660	xd1005	Waswanipi Cree	CA
124661	xd1006	Wegamon Ojibwa	CA
124662	xd1007	Mountain Dene	CA

Showing 1 to 10 of 1,428



## What You Will Get

- Search for places
- Search for data
- Map historical places
- Get geometries for places from names
- Publishing credit for uploads (much later)



## Status

- Late fall/early winter 2019: Beta Release

<http://whgazetteer.org/>

@WHGazetteer



**QUESTIONS?**



University of Pittsburgh

**TRY YOURSELF**





<http://dev.whgazetteer.org/>

- <https://tinyurl.com/yxcft6cb> -  
**WHG Github**
- <https://tinyurl.com/yyslln4n> -  
**Sample Data**



University of Pittsburgh

*Thank you!*

Susan Grunewald, Ph.D.  
Digital History Postdoctoral Associate  
World History Center  
University of Pittsburgh  
@SusanGrunewald1  
[scg52@pitt.edu](mailto:scg52@pitt.edu)

