World-Historical Gazetteer

The University of Pittsburgh World History Center

University of Pittsburgh

National Endowment for the Humanities

LINKED OPEN DATA

Version 1 was launched in July, 2020

NEH-funded project at the University of Pittsburgh's World History Center (WHC) (2017-2020)

Ruth Mostern (WHC Director)
Principal Investigator

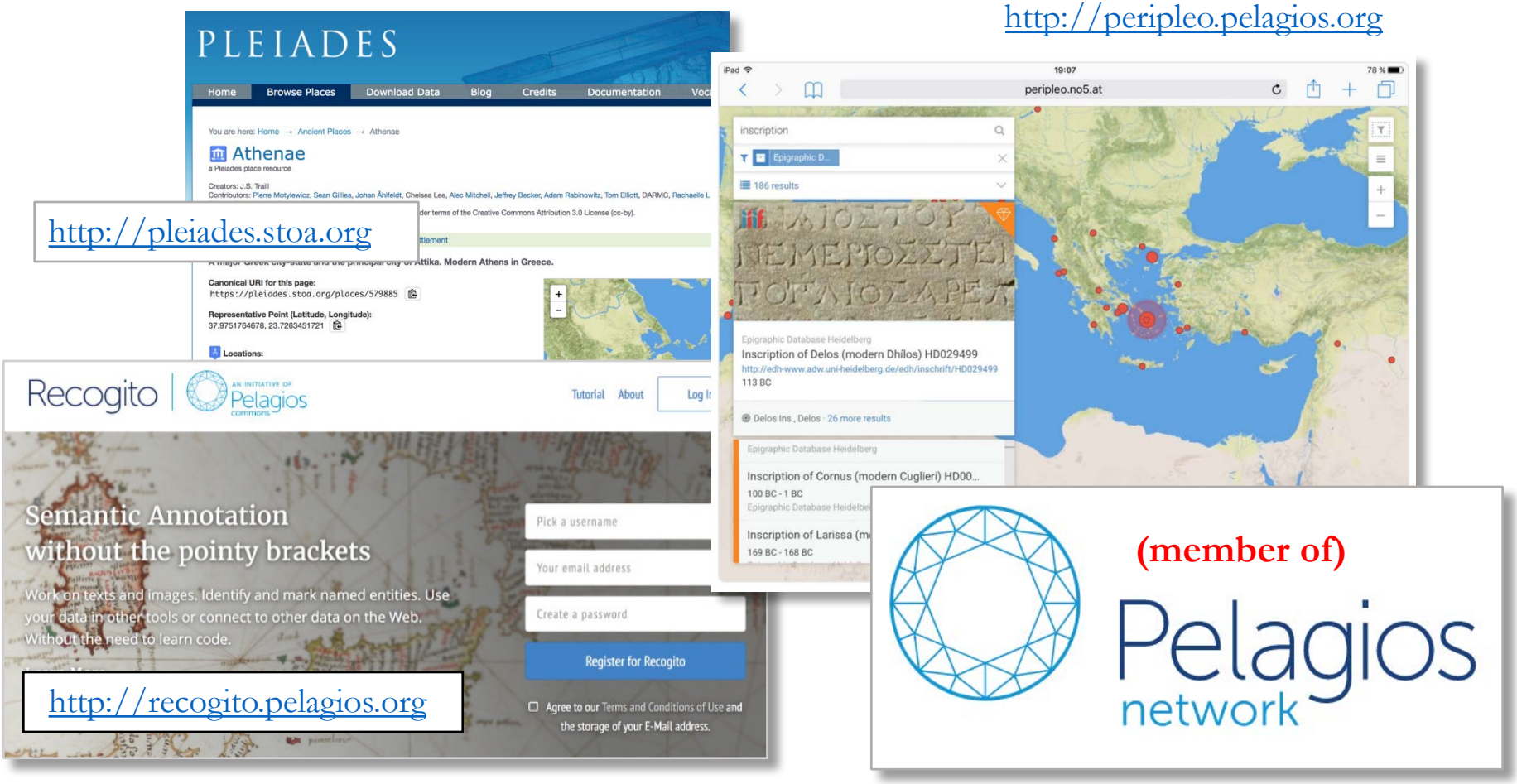
Karl Grossner
Technical Director & Lead Developer

Susan Grunewald
WHC post-doctoral fellow

Patrick Manning (WHC Founder)
Project consultant

http://whgazetteer.org

Following on from Pelagios' **Peripleo**, which was seeded by the **Pleiades** gazetteer



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

I don't go into what we mean by **traces** further here, but there is further explanation in the WHG site guide.

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

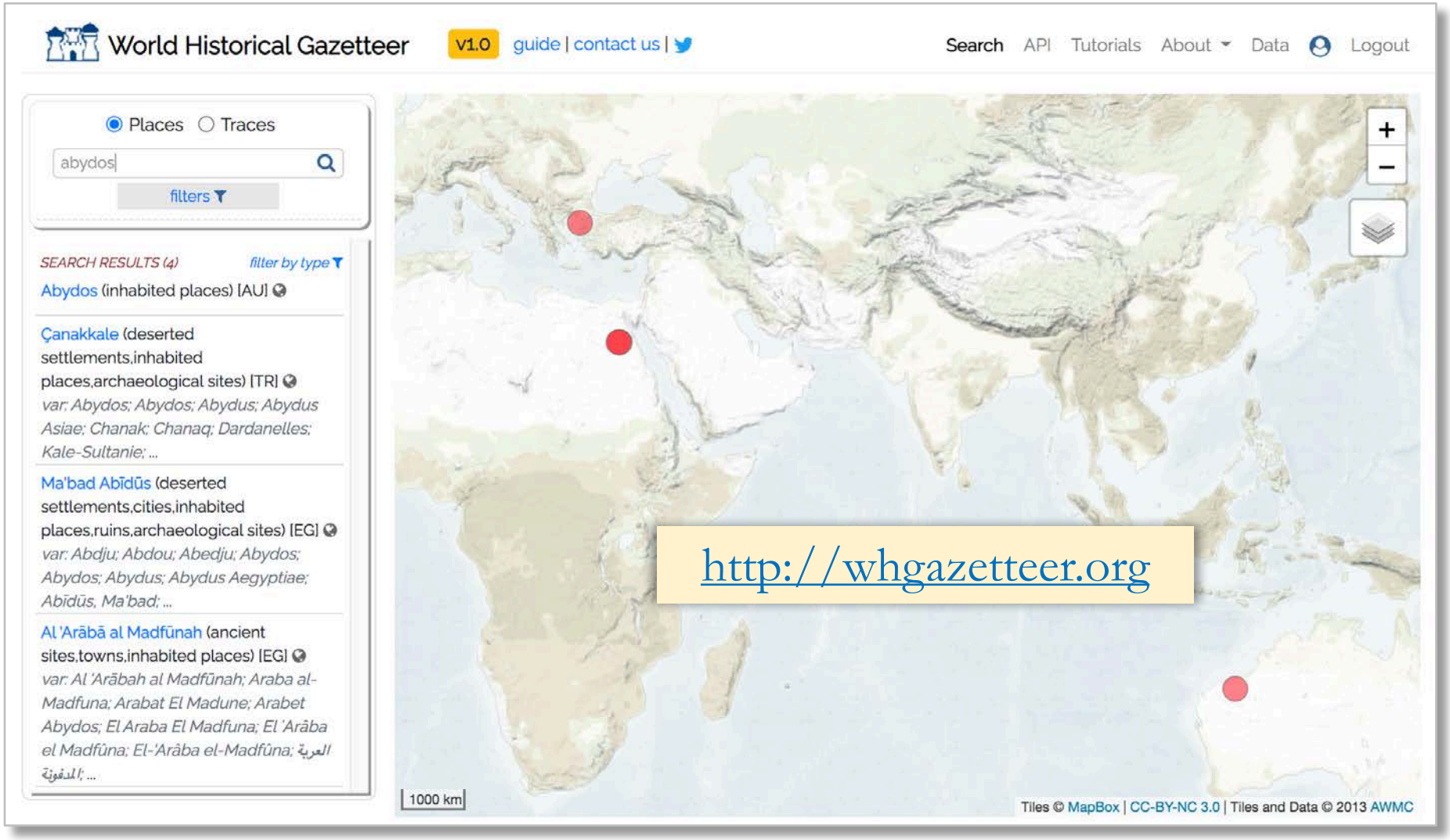
Trace annotations

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

- people
- events
- works

Linked Traces annotation format

settlements, administrative areas, regions, natural features, peoples, routes



filtered search, API, private workspace, reconciliation services, contributions

China Historical GIS

Euratlas

HGIS de las Indias

al-Thurayyā (Islam)

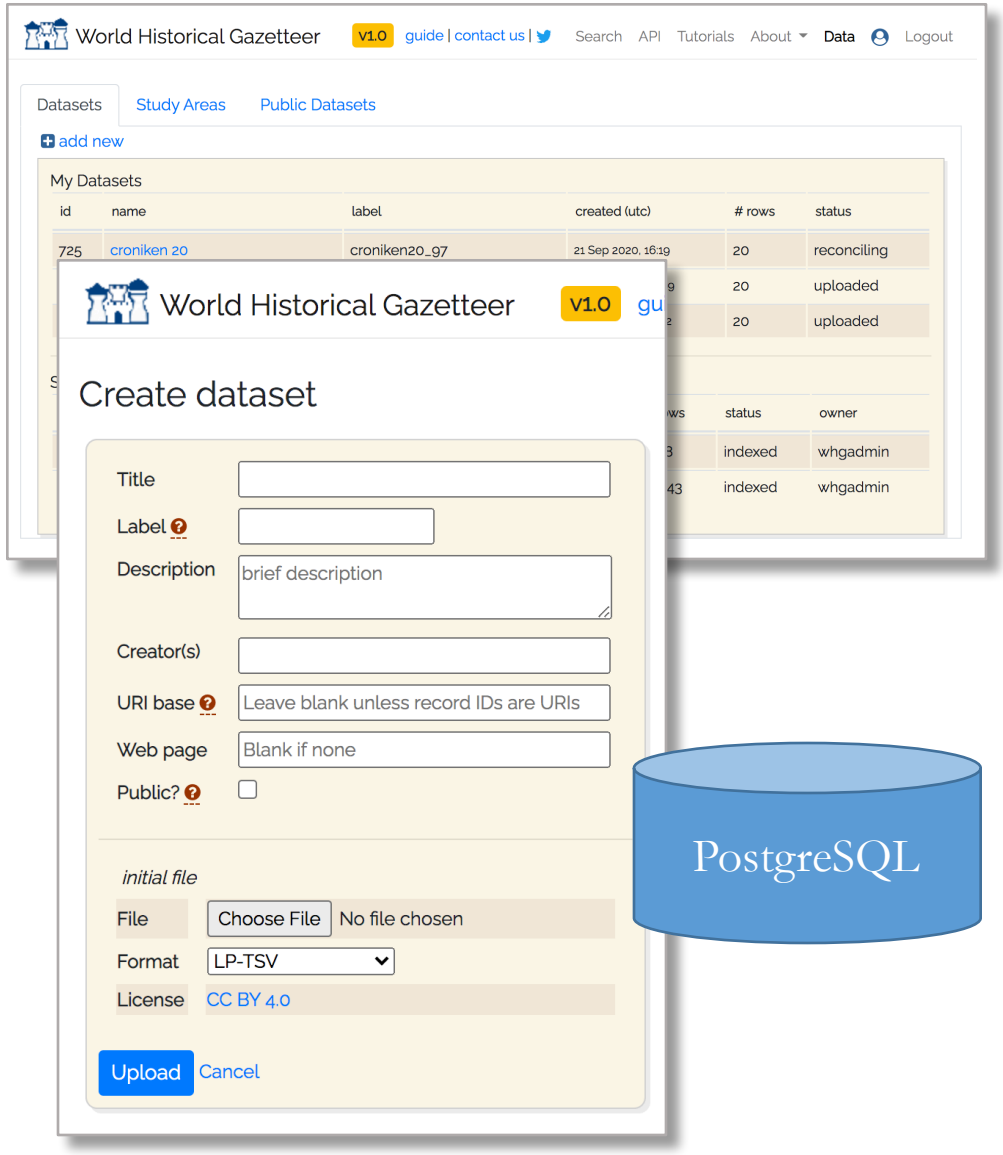
D-Place

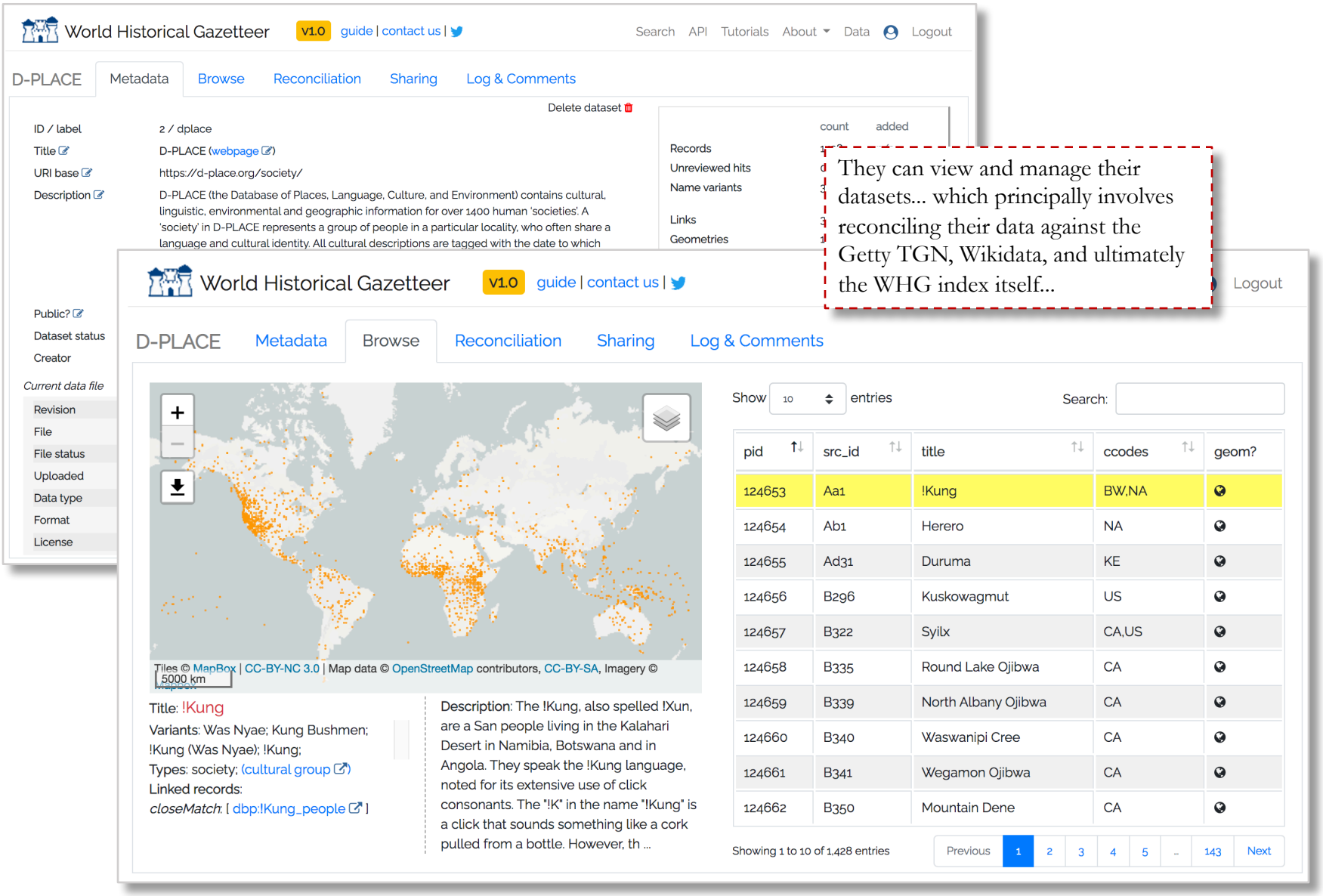
Enslaved

Old World Trade (OWTRAD)

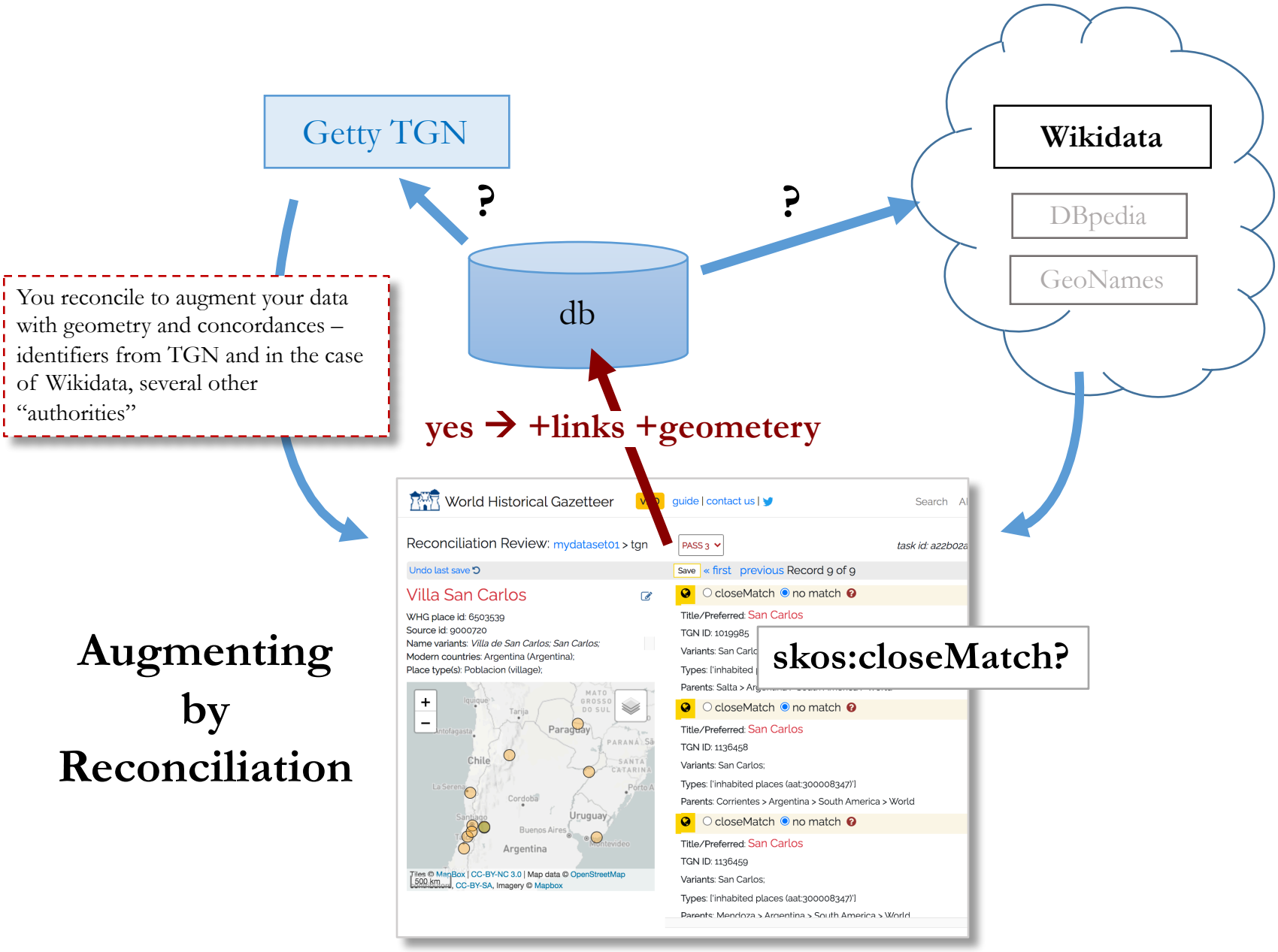
Pleiades

Registered users with place data (individuals or project members) can upload datasets to a private "workspace," where they are stored in a relational database.



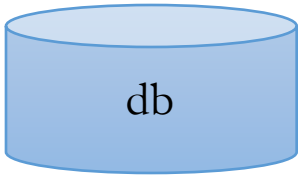


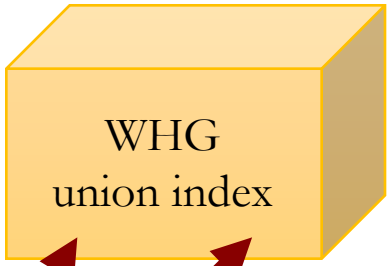
They can view and manage their datasets... which principally involves reconciling their data against the Getty TGN, Wikidata, and ultimately the WHG index itself...



Augmenting by Reconciliation

The accessioning step involves adding data to the WHG index (now rich with as much geometry and as many authority ID matches as possible)



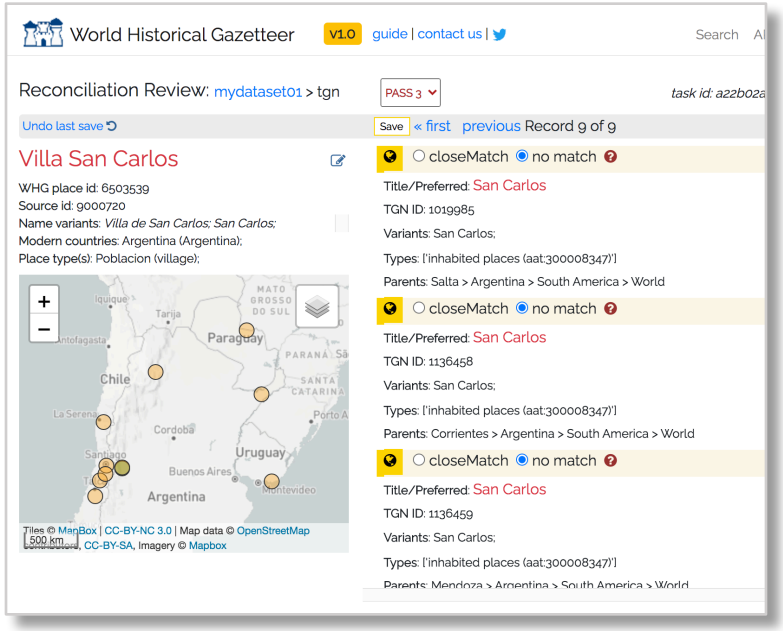


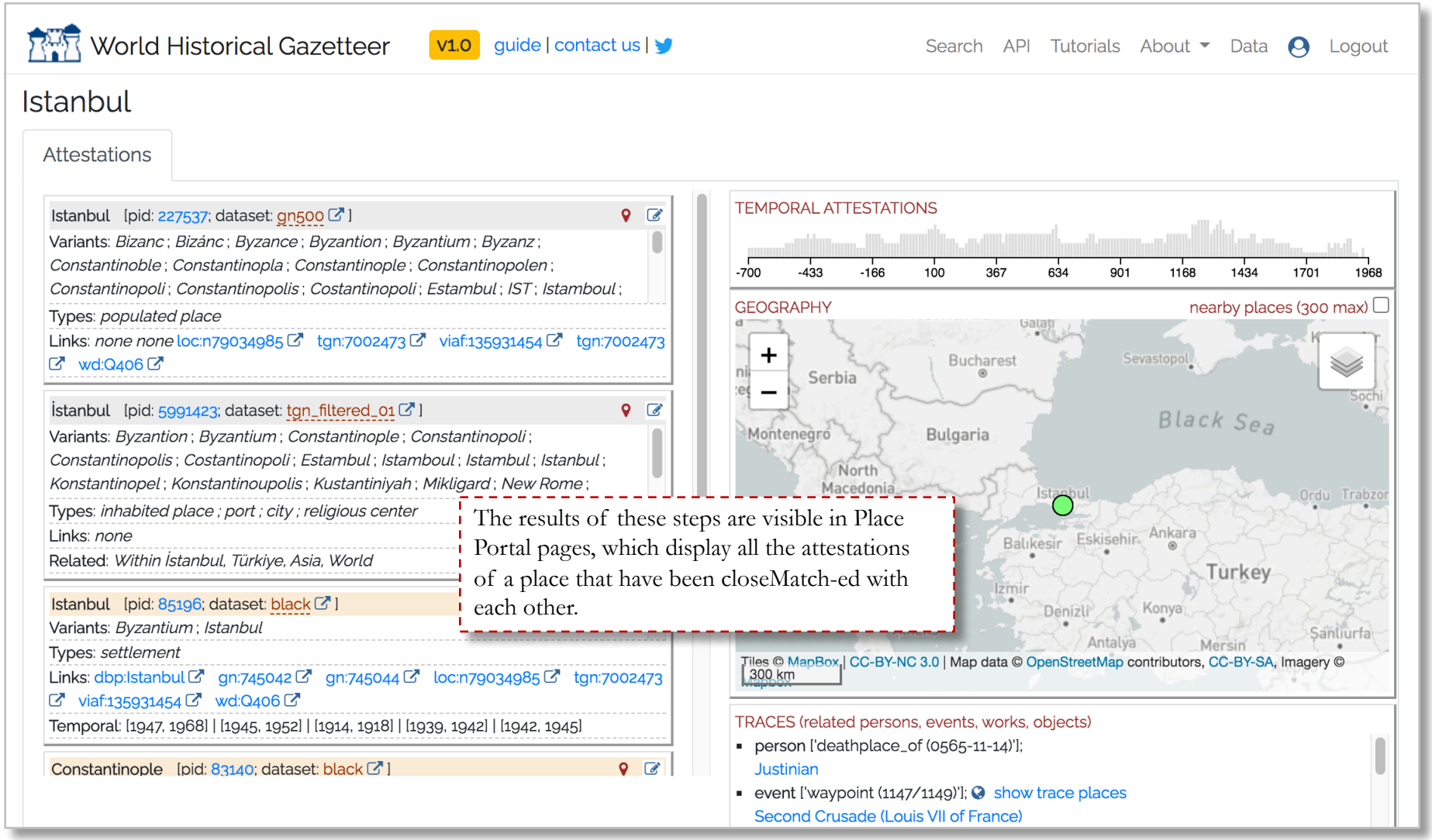
yes → child

no → new record

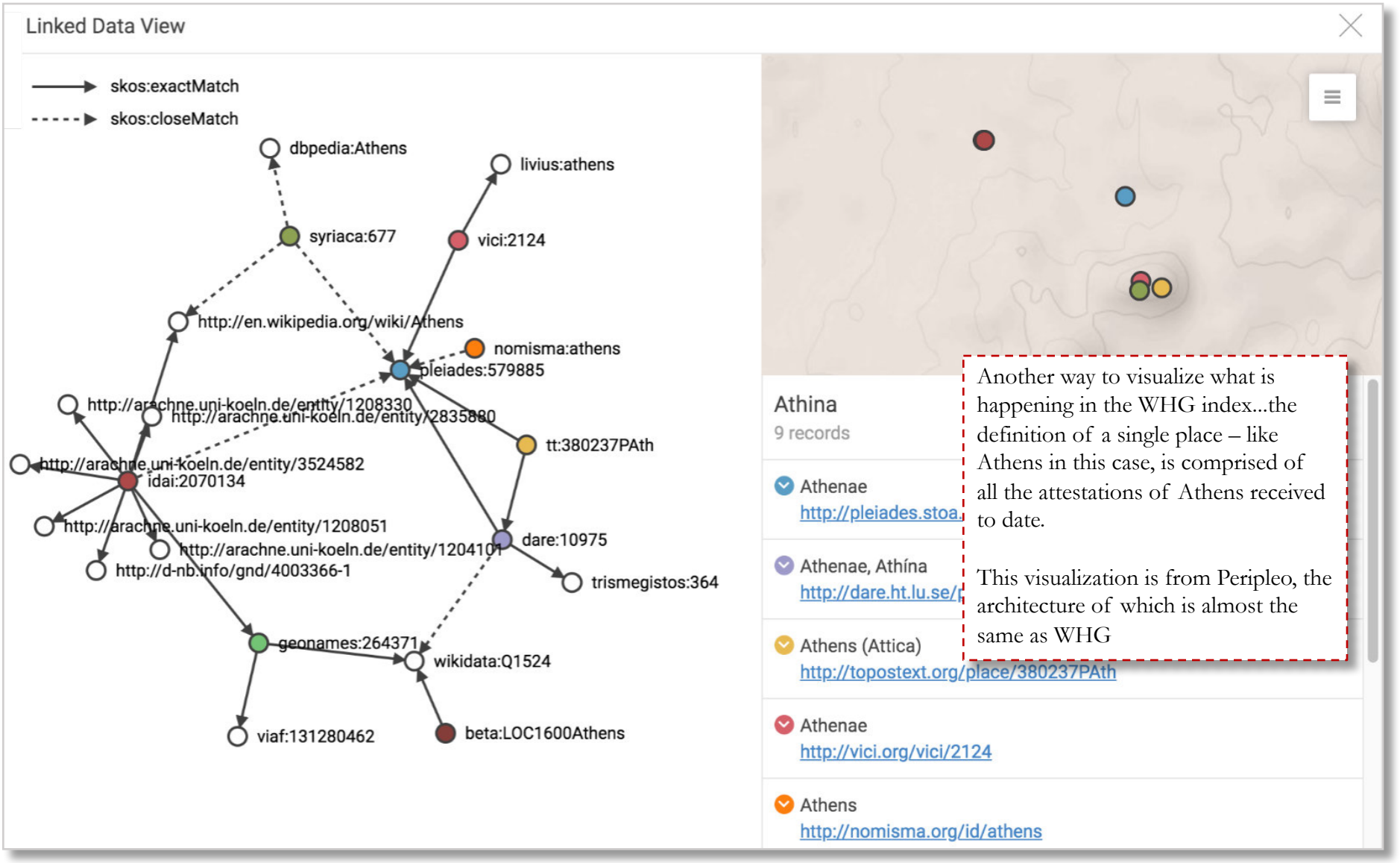
skos:closeMatch?

Accessioning to WHG index

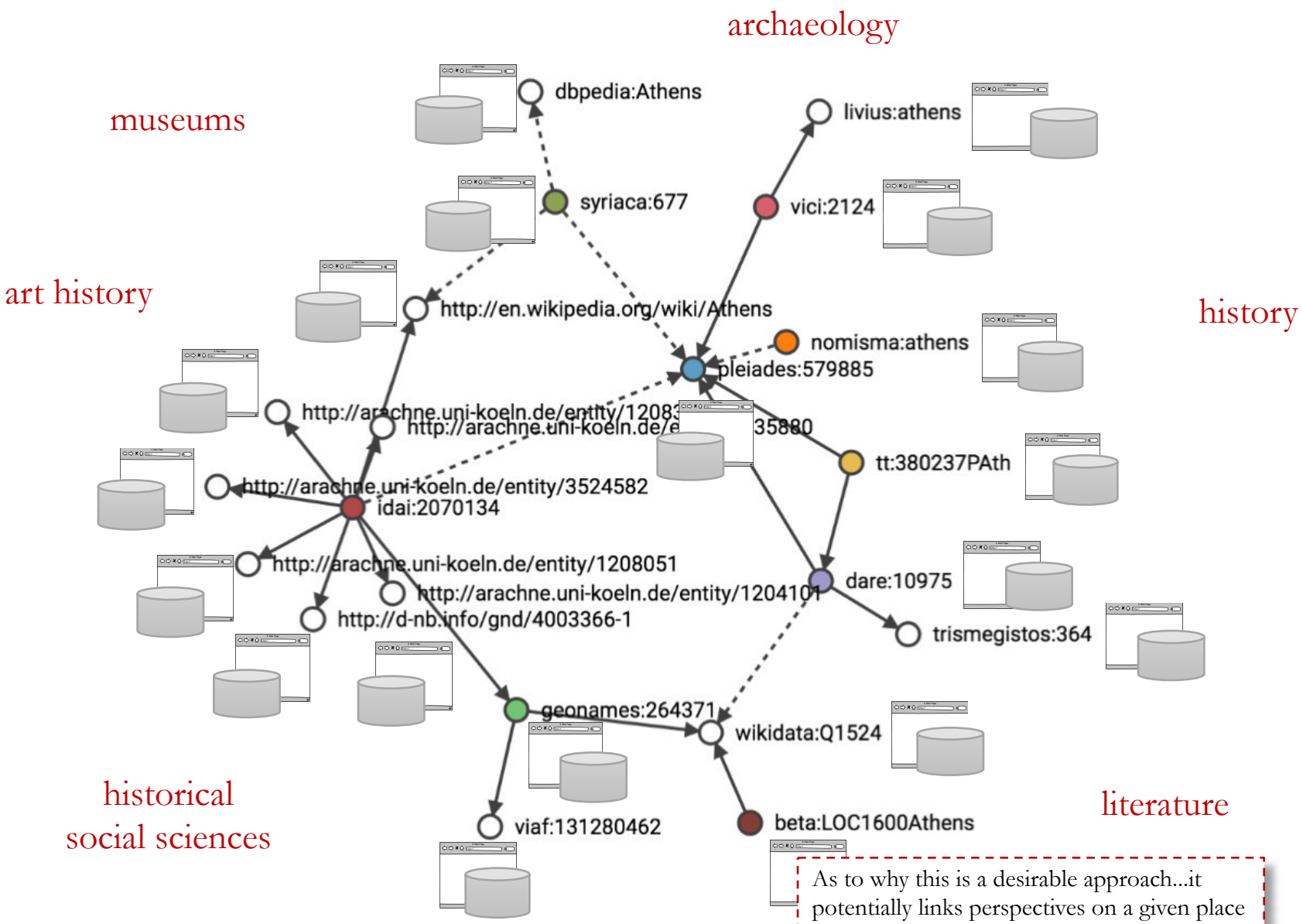




The results of these steps are visible in Place Portal pages, which display all the attestations of a place that have been closeMatch-ed with each other.



http://peripleo.pelagios.org



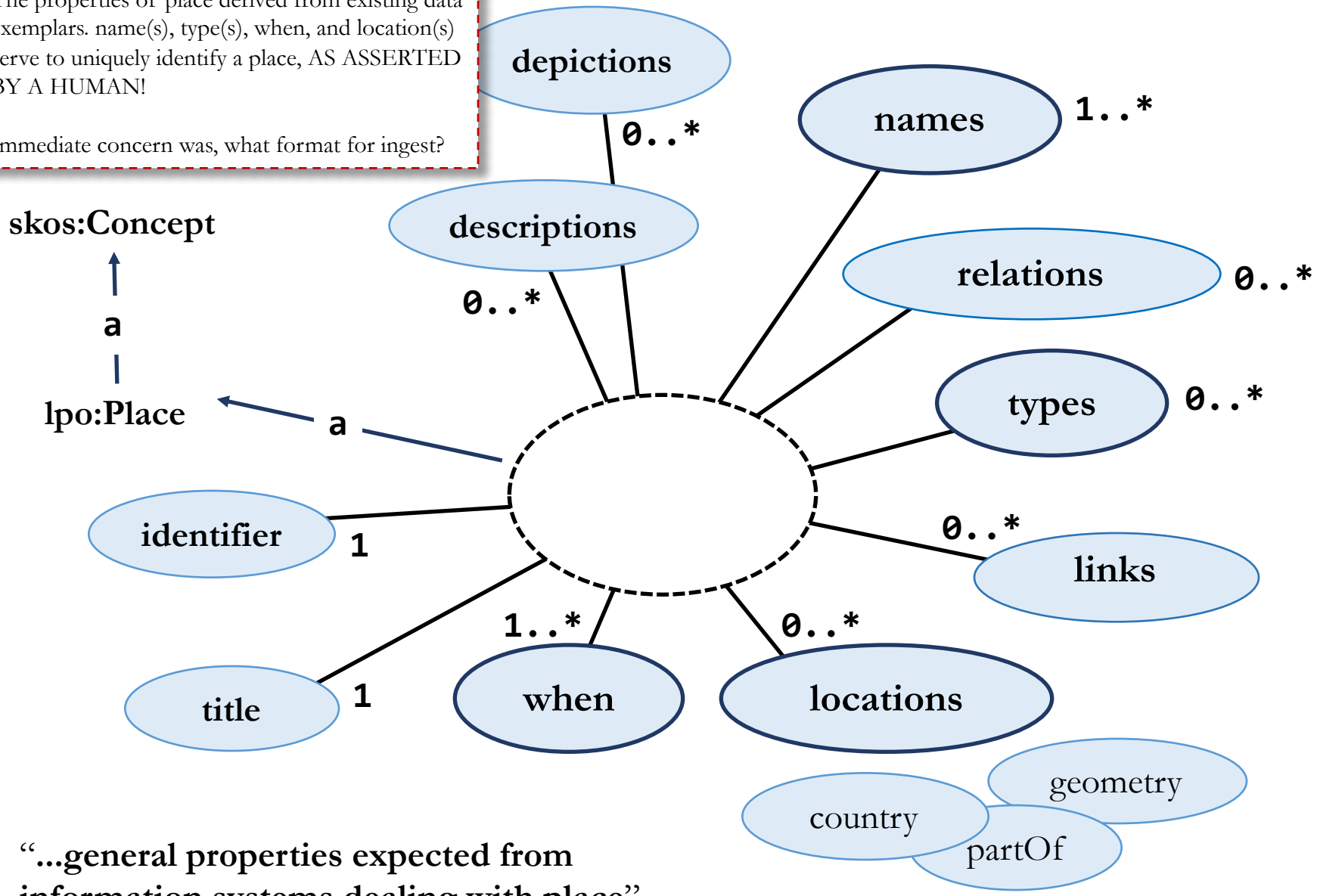
As to why this is a desirable approach...it potentially links perspectives on a given place from many fields

Linked Places

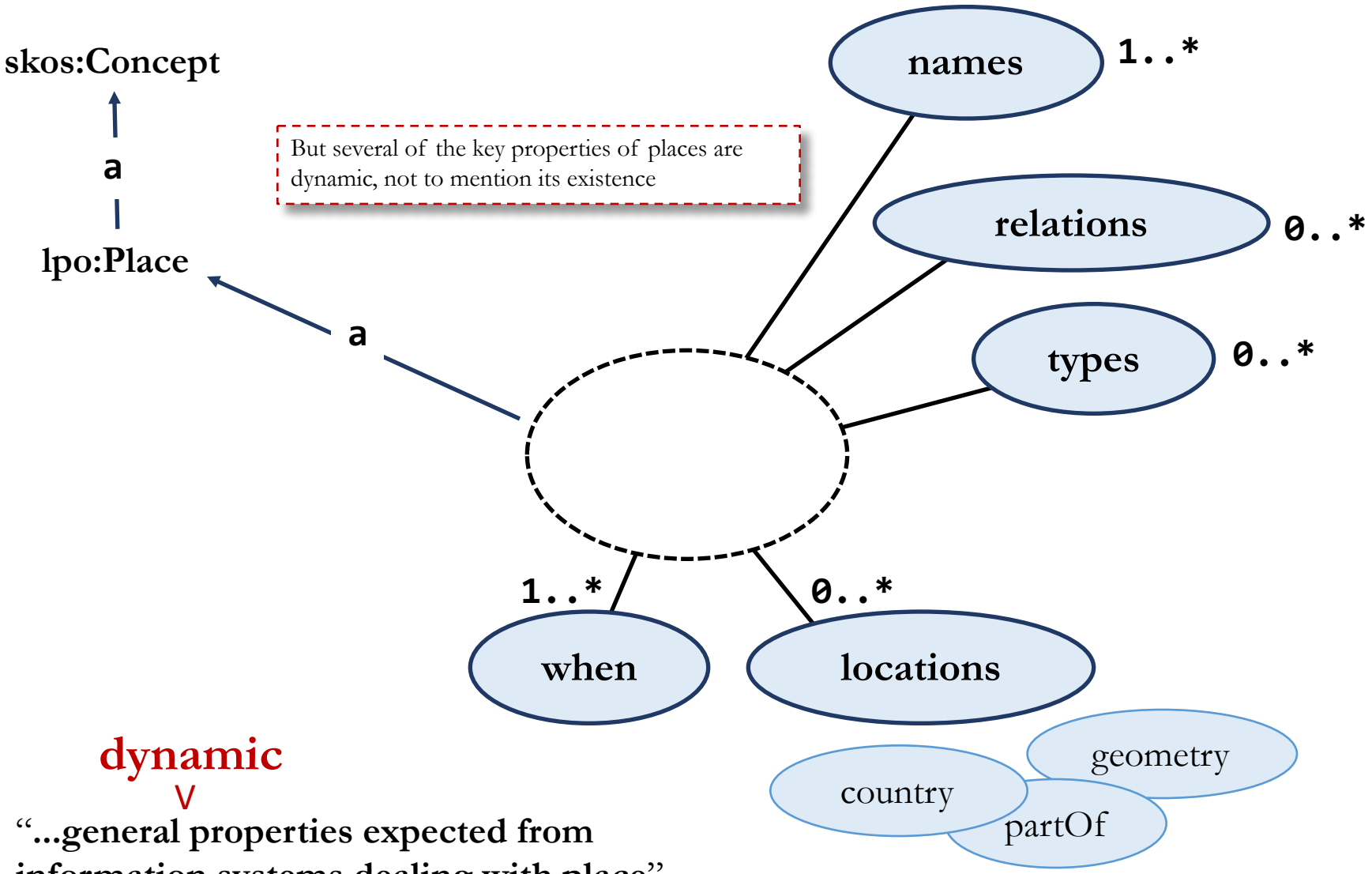
A model and a format

The properties of place derived from existing data exemplars. name(s), type(s), when, and location(s) serve to uniquely identify a place, AS ASSERTED BY A HUMAN!

Immediate concern was, what format for ingest?



"...general properties expected from information systems dealing with place"



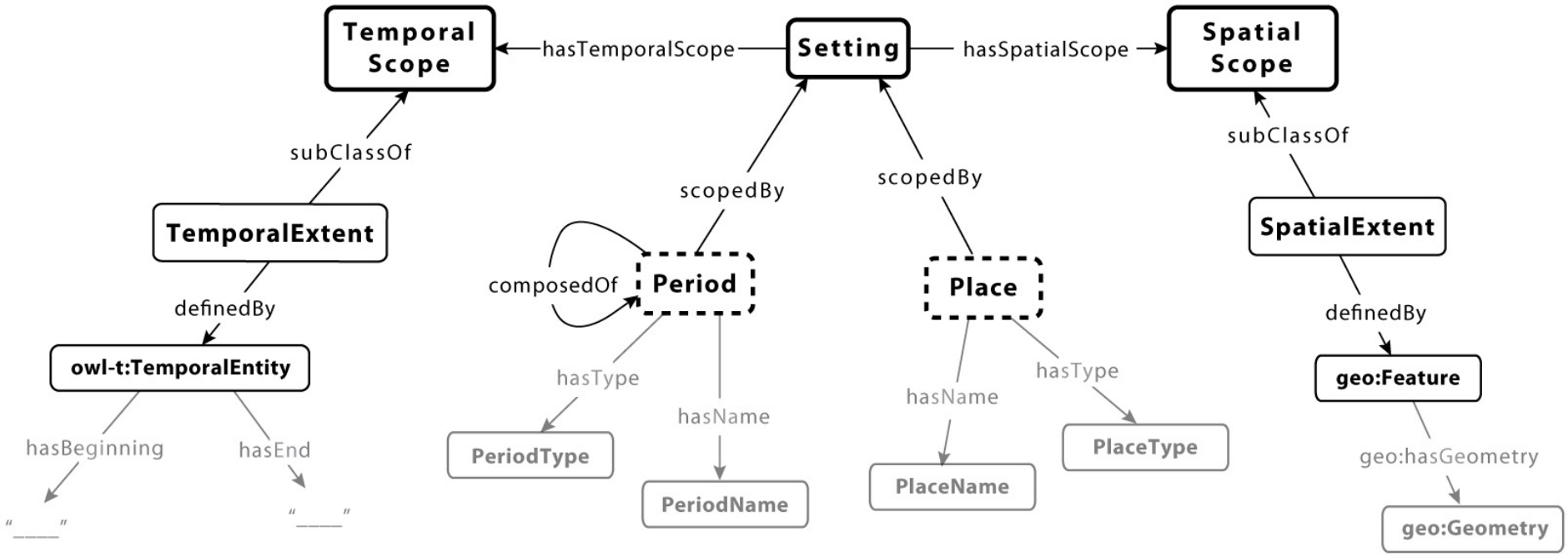
dynamic

V

"...general properties expected from information systems dealing with place"

Setting pattern

A few years before the WHG project began, I did some work with Krzysztof Janowicz and Carsten Keßler, developing an ontology design pattern for setting, reflecting the fact that both places and historical periods have spatial and temporal scopes and extents. Trying to model the way that the answer to where is often "here, then" and the answer to when can be "then, here"



K. Grossner, K. Janowicz & C. Keßler (2014)
Place, Period and Setting for Linked Data Gazetteers
in Berman, Southall, Mostern (Eds.) *Placing Names*

GeoJSON

https://geojson.org/

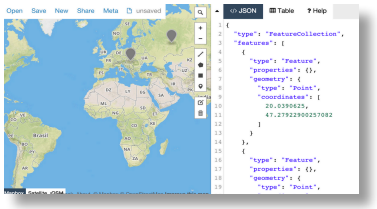
I needed a data format to join space and time – why start from scratch? GeoJSON is very widely used for web mapping applications.

> GeoJSON is a format for encoding a variety of geographic data structures.

OpenLayers™

Leaflet

mapbox



geojson.io

GitHub

QGIS

esri

```
{ "type": "FeatureCollection",
  "features": [
    {
      "type": "Feature",
      "properties": {      },
      "geometry": {      },
    },
    ...
  ]
}
```

GeoJSON models Features in FeatureCollections, with only three required properties: type, geometry, and properties. Properties is free-form, intentionally. Geometry is further specified. New elements can be added, with some restrictions, called "foreign members."

GeoJSON

```
{ "type": "FeatureCollection",
  "features": [
    {
      "type": "Feature",
      "properties": {      },
      "when": {      },
      "geometry": {      }
    },
  ]
}
```

So why not add "when" – at the level of the entire feature. In 2017 I began drafting a GeoJSON-T spec. It is still draft and provisional, though it has been getting more attention and traction recently.

GeoJSON-T

```
{ "type": "FeatureCollection",
  "features": [
    {
      "type": "Feature",
      "properties": {      },
      "when": {  },
      "geometry": {
          “type”: “GeometryCollection”,
          “geometries”: [{
              “type”: “MultiPolygon”,
              “properties”: {      },
              “coordinates”: [      ],
              “when”: {  }
          }]
    }, ... ]
}
```

And if the geometry is itself a collection of geometries, why not allow a “when” for each?

GeoJSON-T

```
"when": {
  "timespans": [
    {
      "start": { "in": "nnnn-nn" },
      "end": {
        "earliest": "-nnnn",
        "latest": "nnnn-nn-nn"
      },
    }
  ],
  "periods": [
   {
    "name": "Hellenistic Period",
    "uri": "http://n2t.net/ark:/99152/p0mn2ndq6bv"
   }
  ],
  "duration": "P100Y",
  "follows": "http://mygaz.org/p_00123",
  "label": "for a century in the Hellenistic period"
}
```

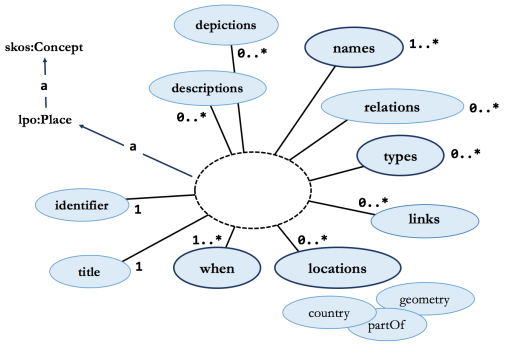
I then proposed some standard properties of "when" objects: timespans, named periods, duration, follows for sequences, and a label

https://github.com/kgeographer/geojson-t

GeoJSON-T

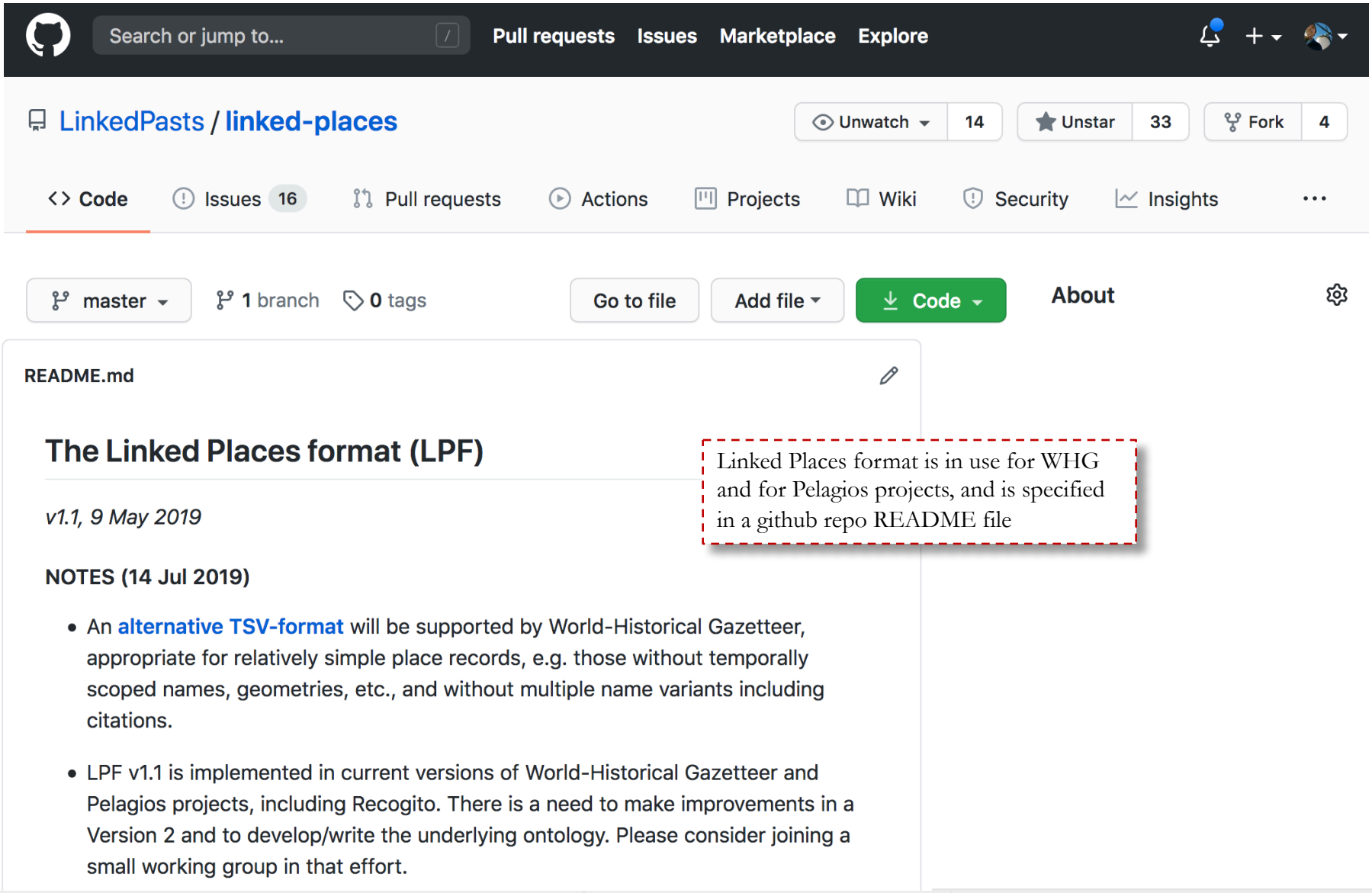
```
{
"type": "FeatureCollection"
"@context": http://linkedpasts.org/lp-context.jsonld,
"features": [
    {
      "type": "Feature",
      "properties": {"id": " ", "title": " " },
      "geometry": { ..., "when": {  } },
      "when": {      },
      "names": [{ ... , "when": {  }}],
      "types": [{ ..., "when": {  }}],
      "relations": [{ ..., "when": {  }}],
      "links": [{      }],
      "descriptions": [{      }],
      "depictions": [{      }],
    }, ...
]}
```

When it came time to develop a standard contributon data format for WHG, I extended GeoJSON-T, and threw in JSON-LD compatability to make it RDF. You can see *optional* "when" objects can be used to temporally scope an entire feature, singleton geometries or collections, names, types, and relations with other places.



Linked Places

https://github.com/LinkedPasts/linked-places



Linked Places format is in use for WHG and for Pelagios projects, and is specified in a github repo README file

"@context": http://linkedpasts.org/lp-context.jsonld,

```
{"@context": {
  "id:": "@id",
  "type": "@type",
  "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
  "rdfs": "http://www.w3.org/2000/01/rdf-schema#",

  "lpo": "http://linkedpasts.org/ontology/lpo_latest.ttl#",
  "lawd": "http://lawd.info/ontology/",
  "gvp": "http://vocab.getty.edu/ontology#",
  "aat": "http://vocab.getty.edu/aat/",
  "tgn": "http://vocab.getty.edu/tgn/",
  "crm": "http://erlangen-crm.org/current/",
```

“lpo”: “http://linkedpasts.org/lpo_latest.ttl”

```
  "features": {
    "@id": "lpo:hasFeature",
    "@type": "geojson:Feature",
    "@container": "@set"
  },
  "properties": "geojson:properties",

  "geometry": "geojson-t:geometry",
  "geometries": {
    "@id":"lpo:setting",
    "@type": "lpo:Setting",
    "@container": "@set"
  },
  "geo_wkt":"http://www.opengis.net/ont/geosparql#asWKT",
  "periodo": "http://n2t.net/ark:/99152/#",
  "ccode": {"@id": "gn:countryCode"},

  "when": {"@id": "lpo:when"},
  "timespans": {
    "@id": "lpo:timespan",
```

There is a draft context file, which refers to a Linked Pasts Ontology (lpo:)...

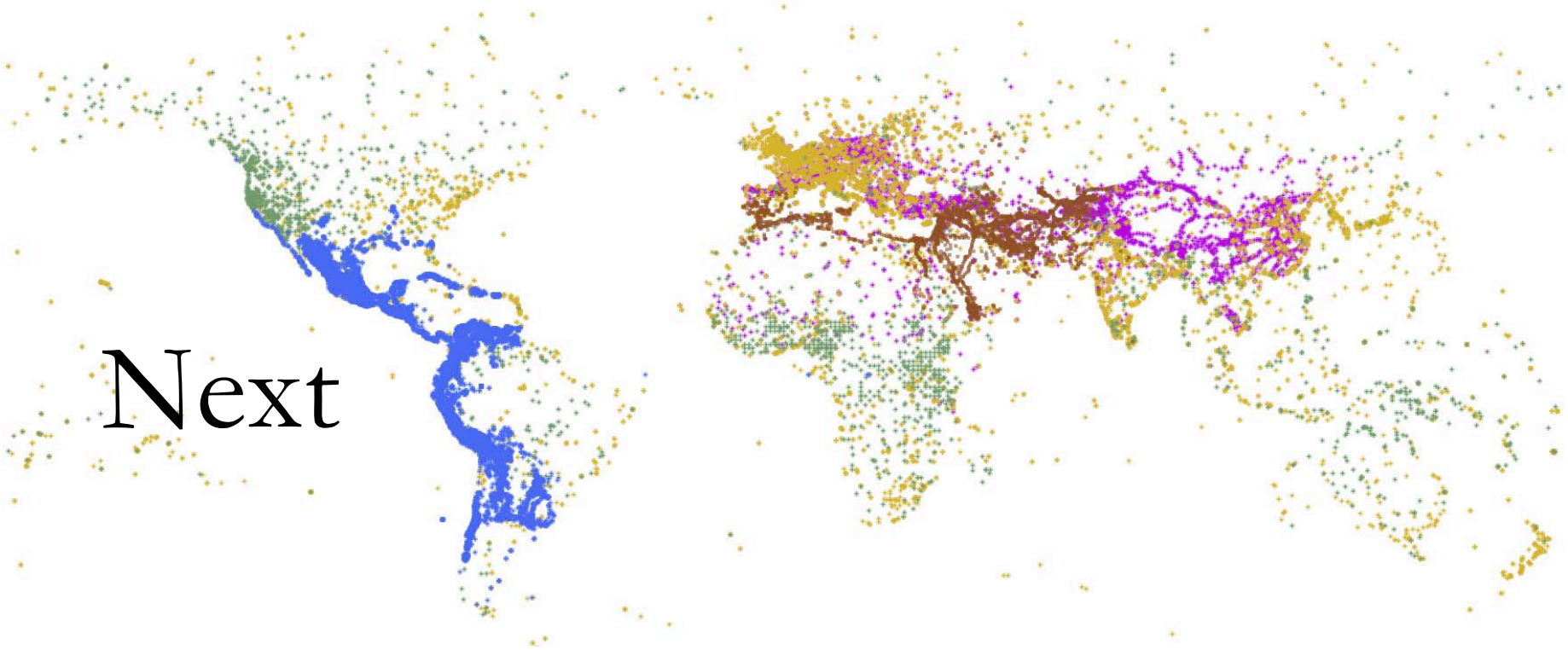
lpo: <http://linkedpasts.org/ontology#>

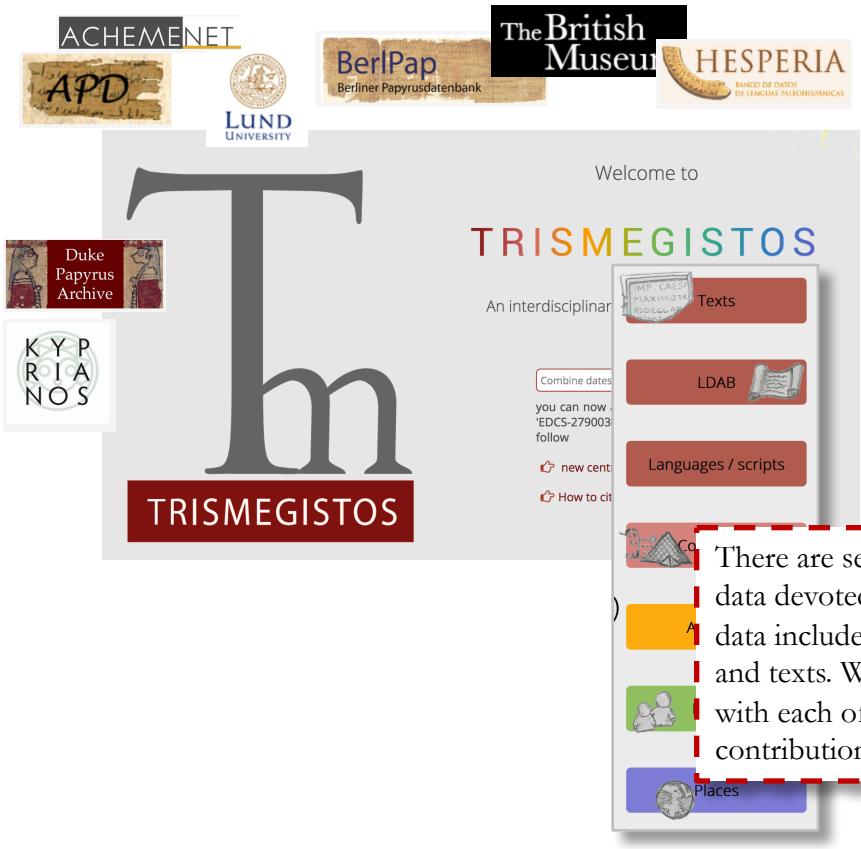
```
## LPO version 1.1. Richard Light, started 13 March 2020

@prefix lpo: <http://linkedpasts.org/ontology#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix time: <http://www.w3.org/2006/time#> .
@prefix xml: <http://www.w3.org/XML/1998/namespace> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

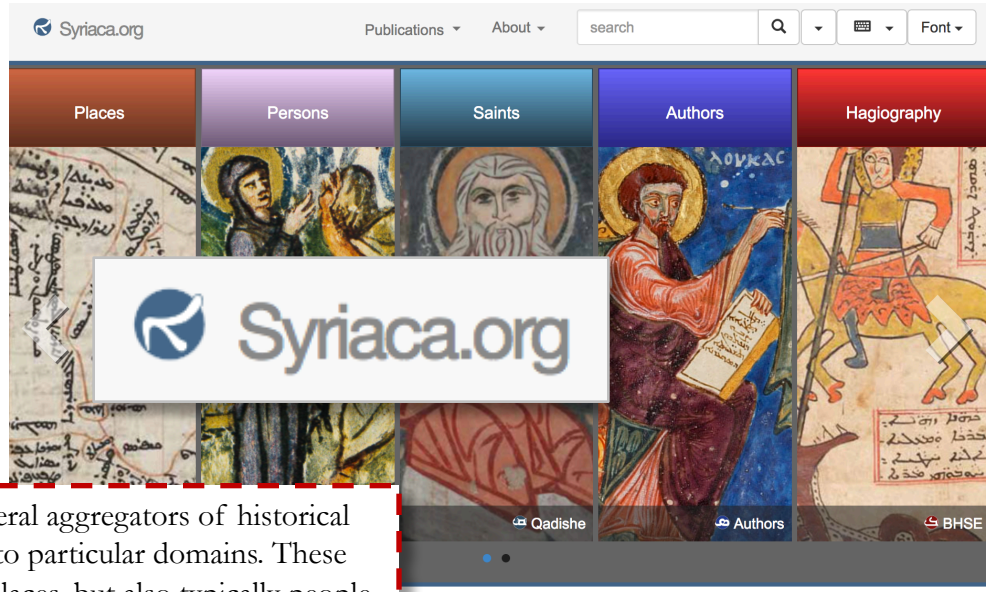
```
lpo:when  a owl:FunctionalProperty ;
  rdfs:range lpo:Timespan ;
  rdfs:comment  "Relates an instance of an entity or
   for a period of time it existed or was valid. For
   association with a particular name, its spatial lo
   place. That TemporalScope could be metric (a time:
   defining a named historical period (e.g. a PeriodO

lpo:timespan a owl:ObjectProperty ;
  rdfs:range lpo:Timespan .

lpo:has_start  a owl:ObjectProperty ;
  rdfs:subPropertyOf  time:intervalStartedBy ;
  rdfs:domain lpo:Timespan ;
  rdfs:range  time:ProperInterval ;
  rdfs:comment "" .
```

The Linked Pasts Ontology is essentially aspirational at this point. I realized as this process went along that the order of development was unusual, or unorthodox, or both. There is a format in use to describe places, which has an underlying formal model that has not yet been committed to valid RDF. I'm not sure what to make of that – except that the model works, and that the ontology will eventually be recorded properly. We (Rainer Simon and I) know what we mean by the terms we've adopted, but haven't yet formalized their semantics.

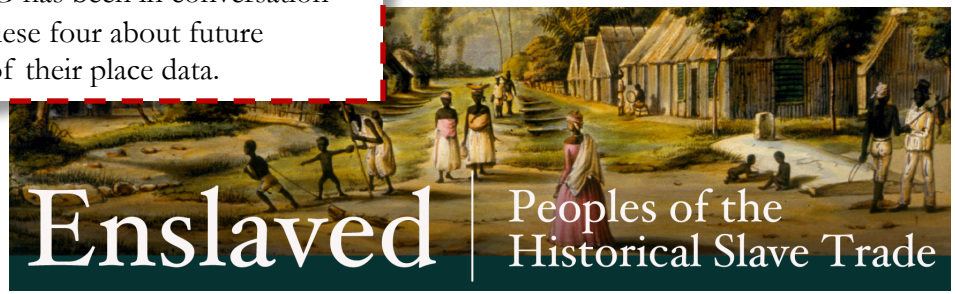
Next

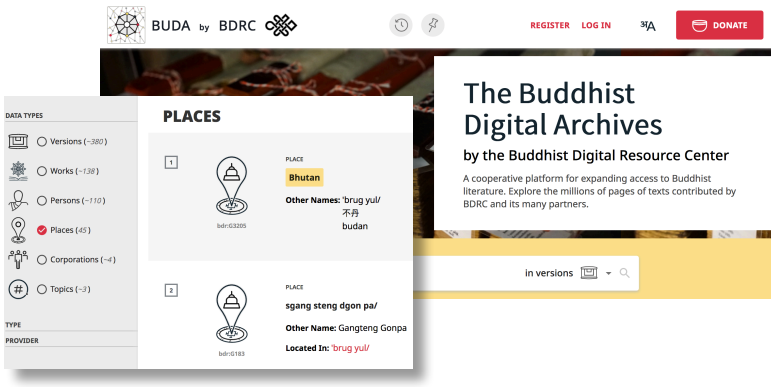




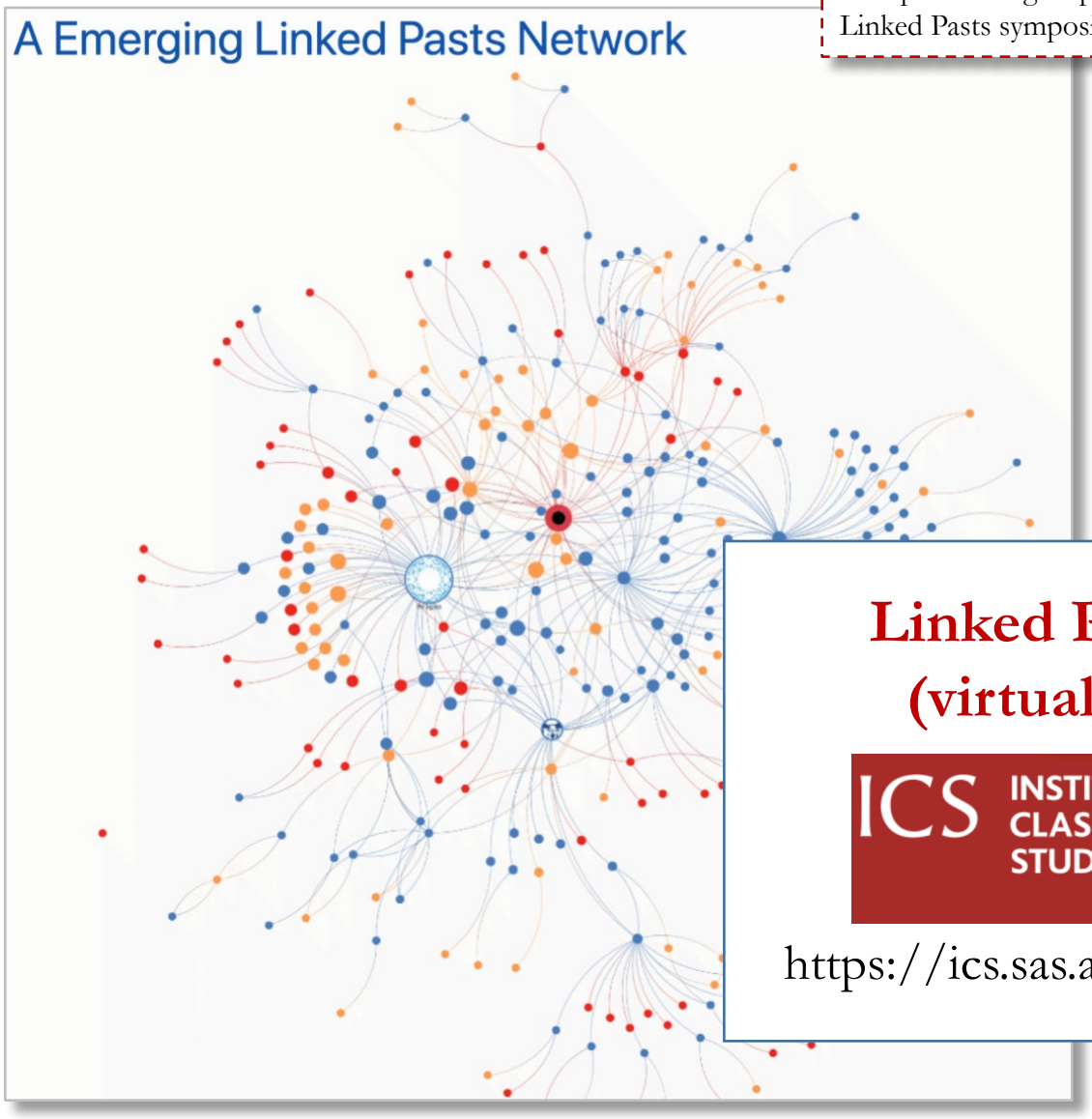


There are several aggregators of historical data devoted to particular domains. These data include places, but also typically people and texts. WHG has been in conversation with each of these four about future contributions of their place data.





- Biographies of the Enslaved (Henry Louis Gates and Steven Niven) View project data
- Free Black Database (Brian Mitchell) View project data | View this journal data article
- Legacies of British Slave-ownership (Keith McClelland) View project data
- Louisiana Slave Database (Gwendolyn Midlo-Hall) View project data | View this journal data article
- Maranhão Inventories Slave Database (Walter Hawthorne) View project data | View this journal data article
- Voyages: The Trans-Atlantic Slave Trade Database (David Eltis) View project data



One place that groups like these meet is the Linked Pasts symposium, now in its sixth year.

Linked Pasts I	London
Linked Pasts II	Madrid
Linked Pasts III	Stanford, CA
Linked Pasts IV	Mainz
Linked Pasts V	Bordeaux

Linked Pasts VI: London (virtual 2-16 Dec 2020)

ICS INSTITUTE OF CLASSICAL STUDIES | SCHOOL OF ADVANCED STUDY UNIVERSITY OF LONDON

https://ics.sas.ac.uk/events/event/22792

Start a huge, foolish project
Like Noah

– Rumi

A poet's directive I have always taken to heart.

Note that not all huge projects are foolish!

http://whgazetteer.org

https://github.com/LinkedPasts/linked-places

https://github.com/kgeographer/geojson-t

@WHGazetteer
@kgeographer 

karl@kgeographer.org