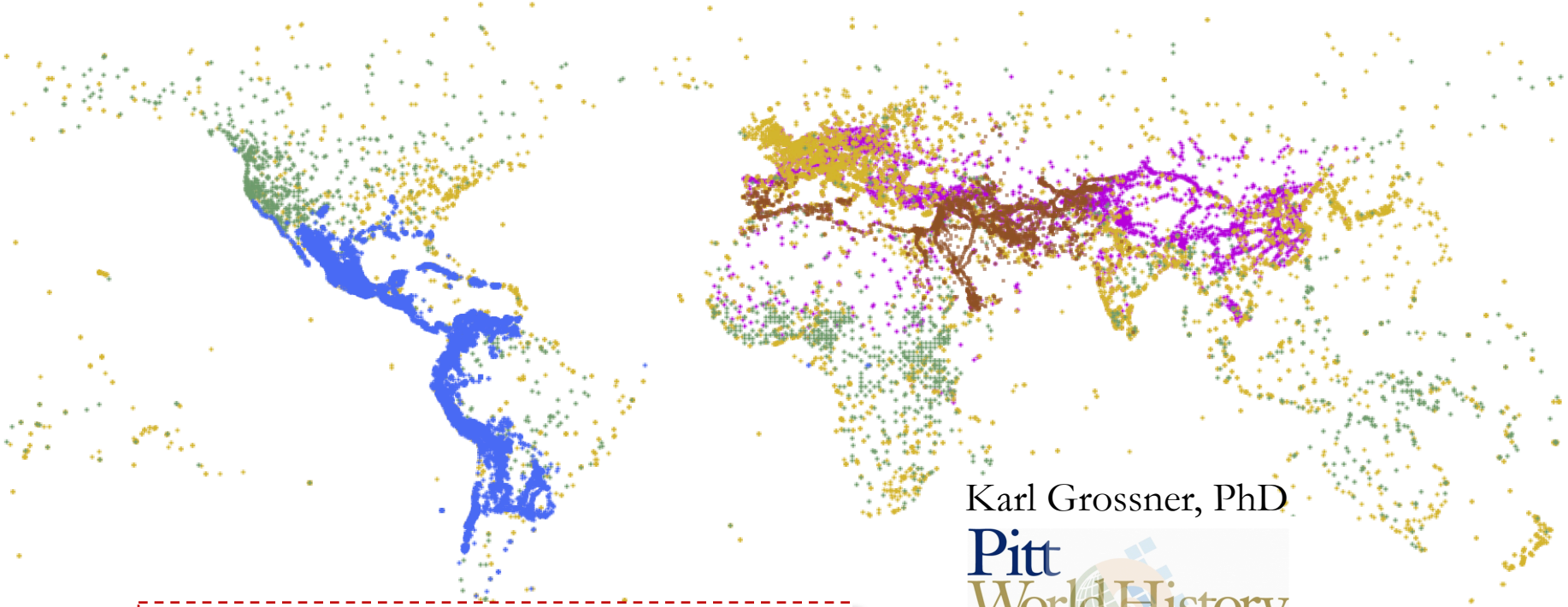
Representing Place in World Historical Gazetteer

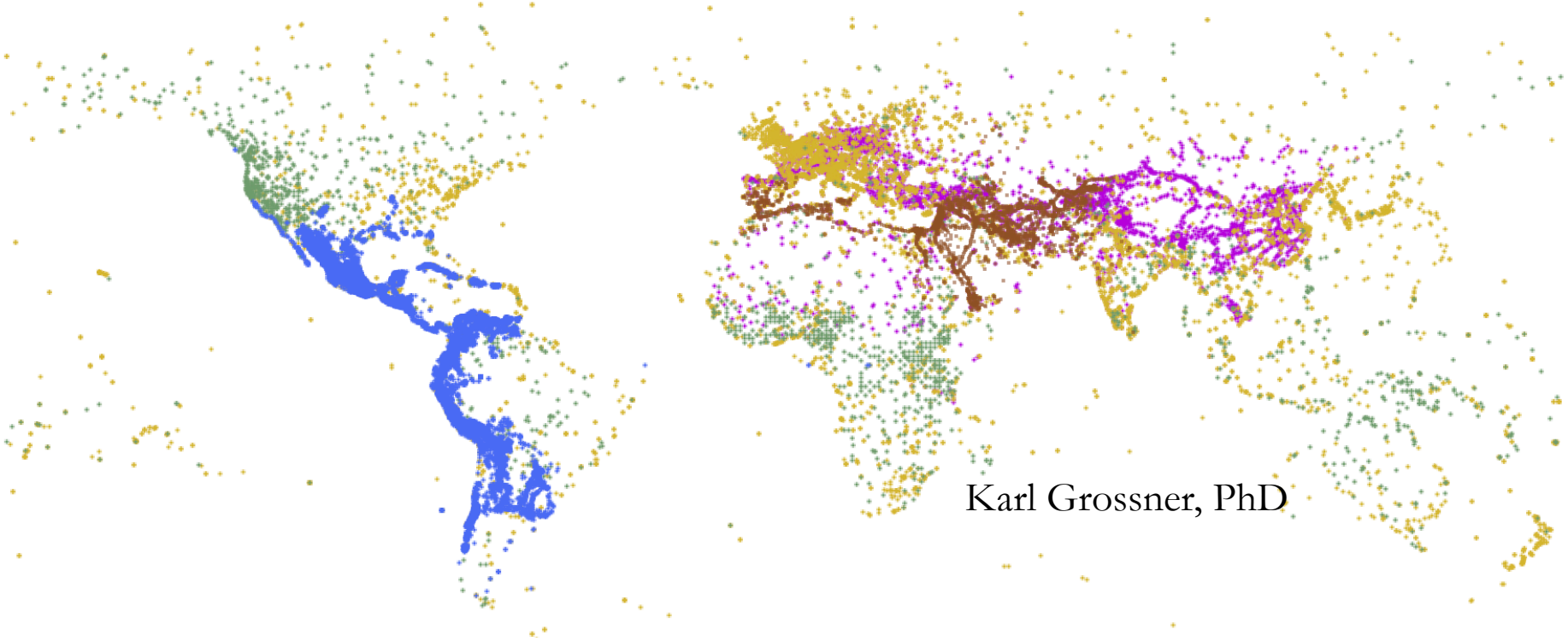


Karl Grossner, PhD
Pitt World History Center

spatial@ucsb
27 Oct 2020

I'm very pleased to be back at UCSB and the Spatial Center, my almae maters. This talk is about representing knowledge of places in the new World Historical Gazetteer platform, and about my now 10-year engagement with Digital Humanities.

Computing Place with World Historical Gazetteer



Karl Grossner, PhD

One day I hope to give a talk with this title. My personal research agenda can be summed up by the phrase "computing place." I see this platform as supporting, among other things, my own research on computation of historical cultural landscapes. Quality of analysis depends in part on quality of description.

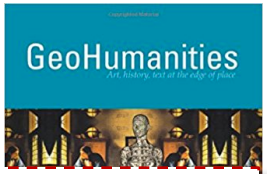
Outline

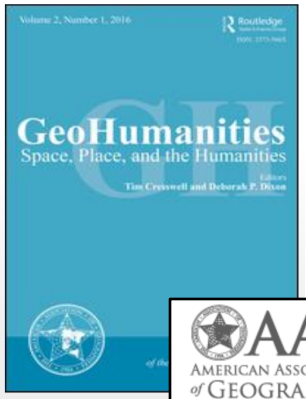
- Some Context
- Place
- World Historical Gazetteer
- Linked Places
- Next

Context

Geo-, Spatial-, Geospatial Humanities

A research trajectory



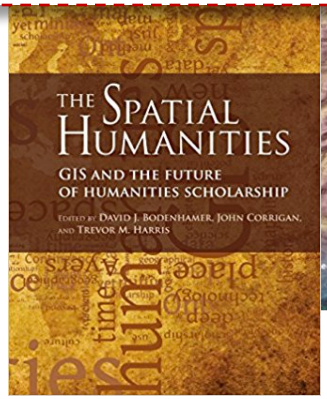


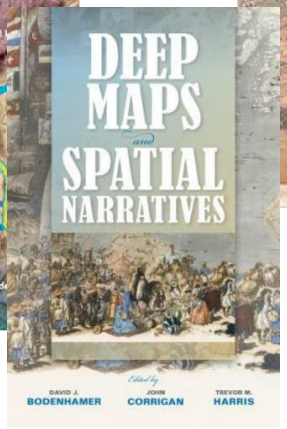
GeoHumanities SPECIAL INTEREST GROUP
Alliance for Digital Humanities Organizations

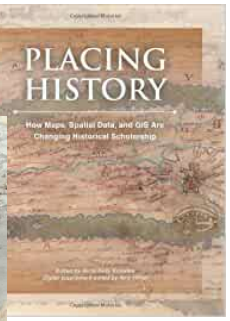
Geography and GIS/GIScience have been connecting with the humanities in several ways over the last couple of decades.
Disciplinary perspectives vary. Anyone who attents the AAG conference will know that cultural geography is well represented, and digital methods are rare. In the AAG journal GeoHumanities, qualitative methods predominate. Historians and geographers have engaged in developing Historical GISes. Historians are generally not keen on including the terms "geo" or "geography" with "history" – claim to want only spatiality.

GeoHumanities

Spatial Humanities/History







Stanford University
Spatial History Project

ROY ROSENZWEIG Center FOR History AND New Media

UNIVERSITY of VIRGINIA LIBRARY
SCHOLARS' LAB

Spatial Humanities
TEXTS, GIS & PLACES
Univ of Lancaster

Historical GIS
SSHA

Summer 2019, UCGIS hosted a meeting on Geospatial Humanities, and produced a special issue of the International Journal of Humanities and Arts Computing, Volume 14 (1-2), 2020

GeoHumanities

Spatial Humanities

Geospatial Humanities

University Consortium for GEOGRAPHIC INFORMATION SCIENCE

The Geospatial Humanities: Transdisciplinary Opportunities for the for GIScience Community

June, 2019 Washington DC

"While humanities scholars have increasingly shown interest in GIS and GIScience, the reciprocal interest is not often as expressed..."

"...a possible methodological path to the **geospatial humanities...**"

Whether or not a new term is needed (or would ever be adopted) advancing the connections between GISci and the Humanities is desirable

Alberto Giordano, Shih-Lung Shaw and Diana Sinton

The Geospatial Humanities: Transdisciplinary Opportunities

International journal of humanities and arts computing, 2020, Vol.14(1/2)

https://www.euppublishing.com/toc/ijhac/14/1-2

2013

GeoHumanities SPECIAL INTEREST GROUP
Alliance for Digital Humanities Organizations

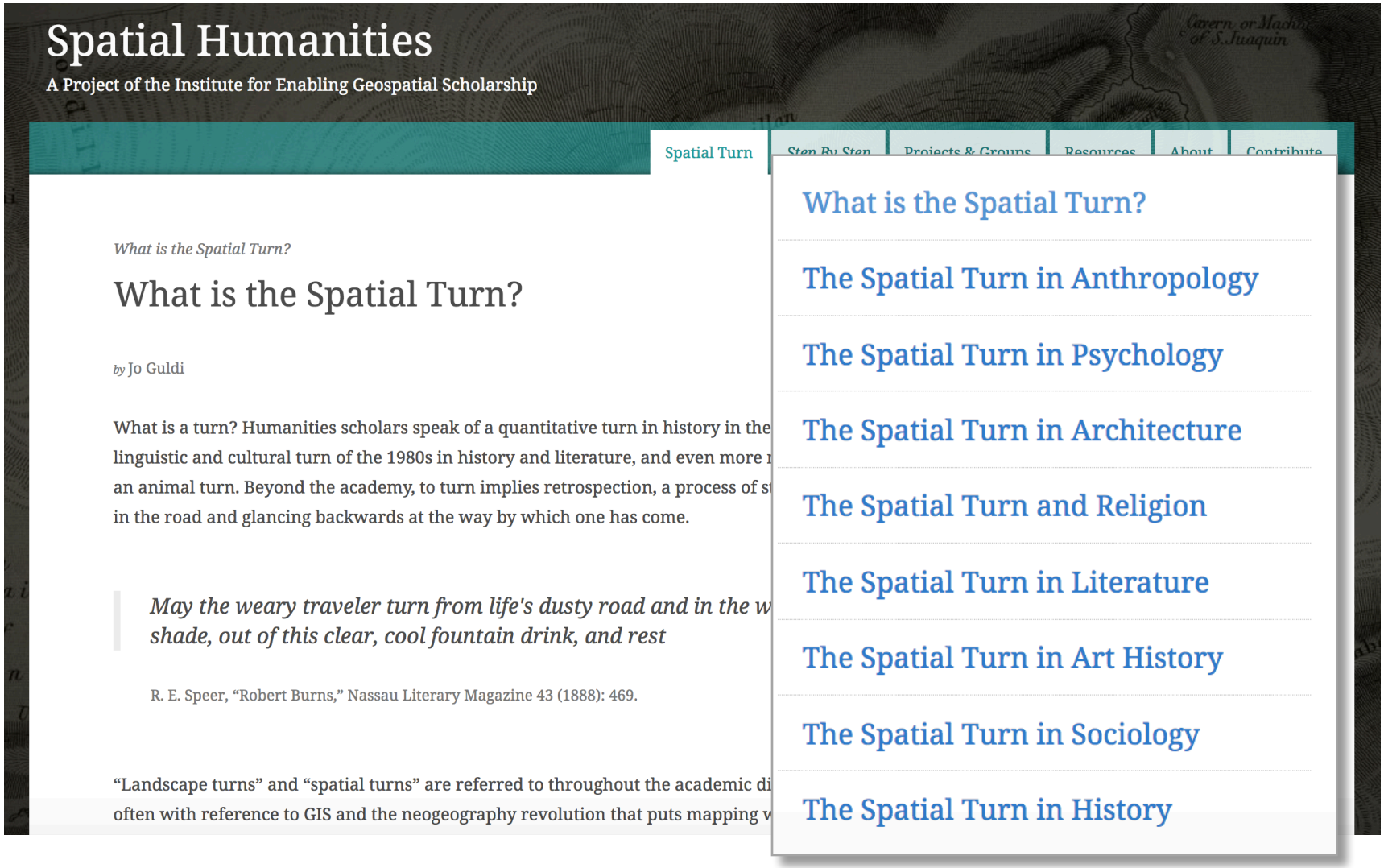
Literary Studies, History, Philology, Linguistics, Archaeology, Arts

"...a venue for pooling knowledge and best practices for relevant existing **digital tools and methods**, to foster the collaborative development of **shared resources and new tools and extensions to geospatial software**, and to keep humanist scholars at large informed about the possibilities and inherent pitfalls in their use."

2015

Bible Geography
Black Geographies
Cultural Geography
Ethnic Geography
Feminist Geographies
...

I entered this meeting ground between history and geography via ECAI at Berkeley, then HGIS and the Social Science History Association. As you probably know, historical research is carried out in most of the social sciences, but many historians do not consider themselves scientists at all. In time I made my way to Digital Humanities, and 2013 co-founded one of the first special interest groups in the Alliance of Digital Humanities Organizations – ADHO. My goal was to make clear to that community that GIScience is NOT static, and if there are deficiencies in geospatial methods and software, these could be addresses, IF the two domains met. In 2015, AAG launched a GeoHumanities journal dominated by qualitative methods and critical theory. Appropriation!!



Jo Guldi, Associate Professor of History, SMU
http://spatial.scholarslab.org/spatial-turn/what-is-the-spatial-turn/

Representing historical knowledge in geographic information systems

(Grossner 2010)

- **digital historical atlases**
- spatial-temporal RDBMS
- "story/deep" maps
- web maps generally
- desktop GIS (esri ArcMap, QGIS)

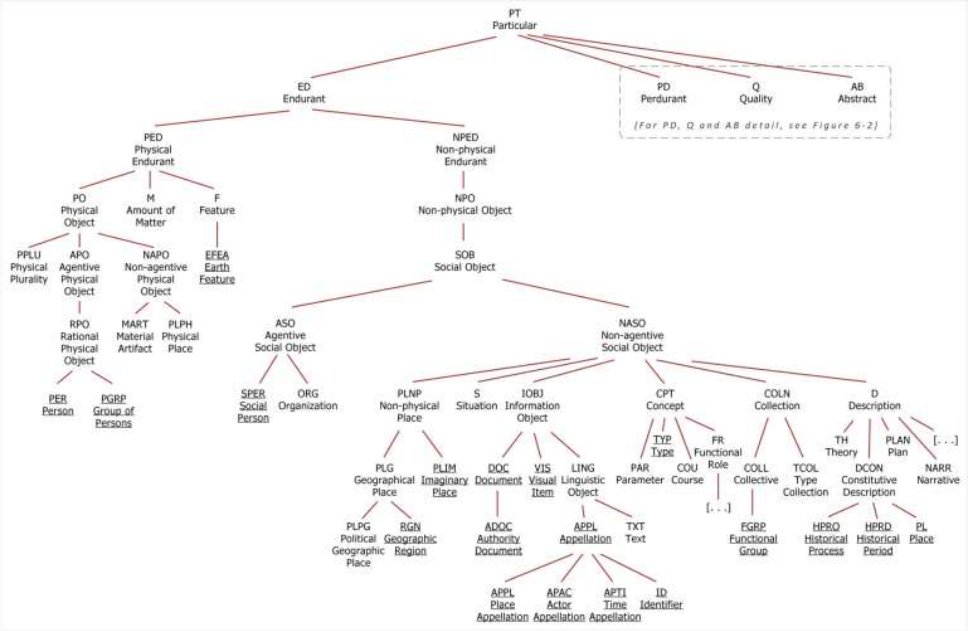
~~**gazetteers**~~
~~triple stores~~

My background: a 2010 dissertation at UCSB. In my work, "geographic information system" is a generic term, not at all limited to desktop GIS software. Gazetteers are certainly geographic information systems, but I did not consider them in that work. I worked with relational databases queried with SQL, and not triple stores queried with SPARQL.
The work did delve into geo-semantics and their formalization, and I relied on work from the Muenster group at the time – with whom UCSB had a strong connection.

Committee: M. Goodchild, H. Couclelis, K. Clarke, M. Raubal, R. Mostern
Muenster allies: W. Kuhn, K. Janowicz, C. Keßler

Spatial History Ontology (SHO)

"...the introduction of logical formalisms to historical knowledge representation in GIS should be attempted incrementally. Fortunately for those who want to make maps right away, a useful level of reasoning can be accomplished in a relational model..."



One product of my dissertation work was a Spatial History Ontology, which I instantiated in a relational database, PostgreSQL. I aimed for what I called "a useful level of reasoning," and "sufficient logic."

Extending DOLCE

(and adopting its First Order Logic notation)

I studied historical atlases extensively, and delved into historiography, knowledge representation, and cognitive science. I ended up extending DOLCE, and defining entities, relations, and axioms in first order logic, as DOLCE does.

"...all Achievements x are either directly constituted (DK) by some Activity y for some time t during the interval bounding the Achievement, or specifically generically constituted (SK) by some Activity throughout that interval (sho11)."

$$\forall x(ACH\ (x))\ \exists y\ ((DK(x, y, t)\ V\ SK(x, y))\ \Lambda\ A(y))$$ (sho11)

In English:

Events, and occurrents generally, are composed of one or more activity, or "temporal substance" – much as material objects are composed of physical substances in some combination.

"sufficient logic"

I discovered I could express the ontology in the RDBMS, and achieve most of the logical constructs and reasoning I was after.

Table 6-1. Logical capabilities compared

RDF/S or OWL capability	Achieved in RDBMS
Schema definitions, e.g. PoliticalEvent hasType rdfs:Class hasParticipant hasType rdf:Property	Tables for Class and Relation ('property'); all instance records have foreign key to Class and/or Relation
Class and property propagation (is-a; sub-relation) Property intersection (A *subproperty-of* B; A *subproperty-of* C → if x A y, then x B y, x C y)	Hierarchy established with parent_id value for each; recursive queries using SQL's `WITH`, `WITH RECURSIVE` and `UNION`
Class definitions	Table columns incl. custom data types (`ENUM`); check constraints incl. `NOT NULL; DEFAULT`
Property definitions - Domain and range - Cardinality	Check constraints
Transitivity of parthood, and is-a relations	Recursion, using SQL's `WITH`, `WITH RECURSIVE` and `UNION`
Complex classes, e.g. NewClass ≙ intersectionOf [ConditionA, ConditionB,...]	Materialized views
Differentiating individuals, e.g. owl:distinctMembers, owl:allDifferent	`UNIQUE` constraints, including primary keys
Instance checking	SQL: `SELECT...WHERE...`
Graph query patterns	SQL: `SELECT...WHERE...AND...AND...FROM...JOIN`
Reification; blank nodes	Association classes
Unions and intersections	SQL: `UNION` and `INTERSECT`

"representing information about the world in a form that a computer system can utilize"

KR meets DH

- maps w/time
- networks of weighted relations
- linked data discovery
- uncertainty

KR: $\forall x(ACH(x)) \exists y ((DK(x, y, t) \vee SK(x, y)) \wedge A(y))$

"events are composed of activities for some period"

DH: "huh?"

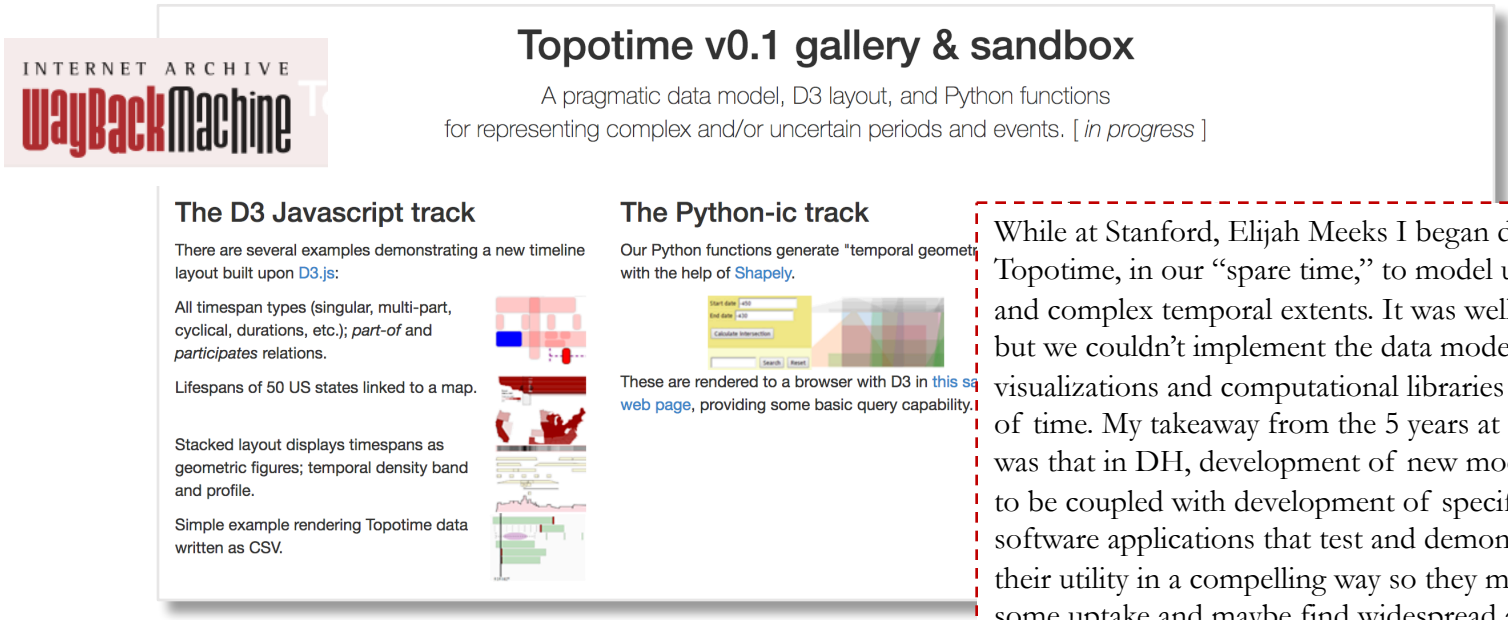
or

"so what?"

or

"my data doesn't look like that; what is cost/benefit?"

I left UCSB and went to Stanford Libraries to work as a Digital Humanities Research Developer. My ideas, for example of event-entered data systems did not get a great reception with my DH colleagues. DHers wanted visualizations of quantified relations between people, objects, concepts, and places. They wanted to represent uncertainty and the absence of data, and my advocacy of event-centered data fell flat.



While at Stanford, Elijah Meeks I began developing Topotime, in our "spare time," to model uncertain and complex temporal extents. It was well received, but we couldn't implement the data model, visualizations and computational libraries for lack of time. My takeaway from the 5 years at Stanford was that in DH, development of new models had to be coupled with development of specific software applications that test and demonstrate their utility in a compelling way so they might get some uptake and maybe find widespread adoption.

Figure 1 – A timespan with uncertain start and end

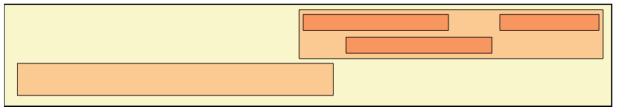


Figure 2 – A composite event rendered as its timespan with two levels of sub-event parts.

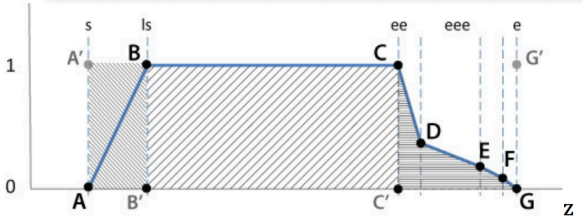
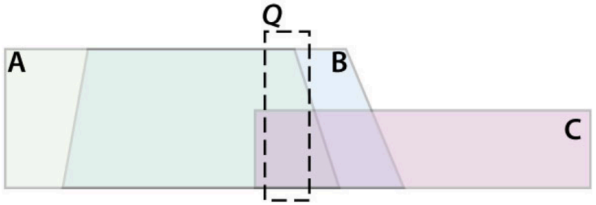


Figure 5 – Percentages of "potential certainty"



KR meets DH

KR: $(PLG(x) \vee GOBJ(x)) \wedge \exists y(APPL(y) \wedge identifies(y, x) \rightarrow \exists t(T(t) \wedge identifies(y, x, t)$ (sho38)

If a Place referent has a name, that name is valid at some time

DH: "huh?"

or

"so what?"

Back in my SHO, I expressed the temporal scoping of place names this way...something I needed, but not interesting to Dhers or directly useful to them.

Place

Defining Place

- experienced space (paraphrasing Yi-Fu Tuan)
- "a meeting up of histories" – Doreen Massey
- a function of events and activity that *have happened* there, and *can happen* there – K. Grossner
  - buildings, monuments, streets, squares, parks built
  - lives led, works created, performances
  - commerce, conflict, meetings, dominion
- places are dynamic, they change over time
- the answer to "where?"

My interest in describing places led me to humanistic geographers' conceptions of Place, which led me first to event-centered models and formats

CIDOC-CRM

Of course that led to investigating CIDOC-CRM, which is a popular "ontology" within the digital humanities, at least within the GLAM domain (for which it was developed). It's definition of Place does not correspond with broad requirements within DH. Scope notes for place_is_defined_by were recently amended to account for "phenomenal places," which are not differentiated formally in CRM

E53 Place

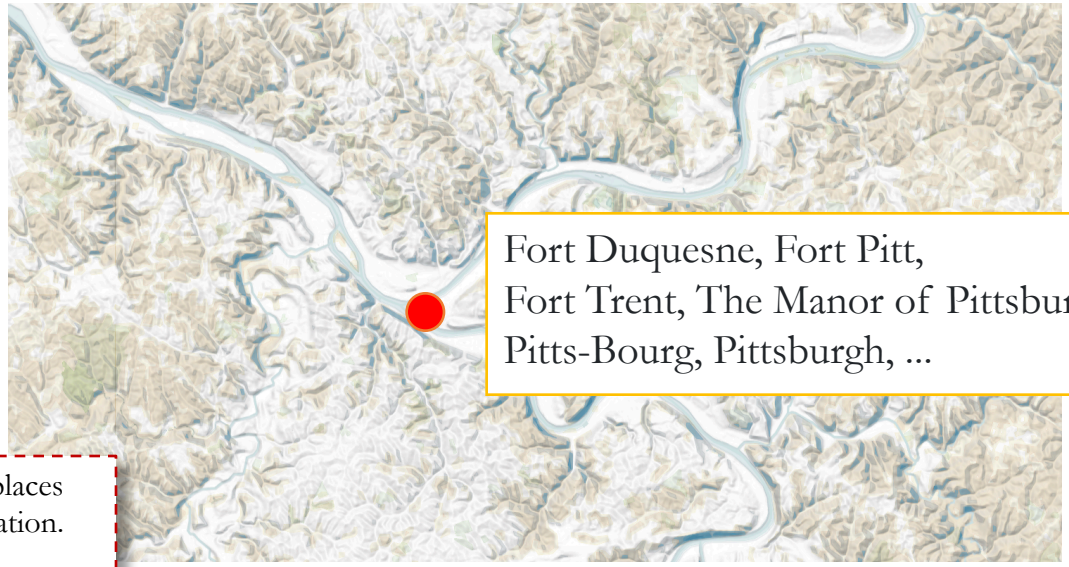
This class comprises **extents in space**, in particular on the surface of the earth, **in the pure sense of physics: independent from temporal phenomena and matter**.

P168 place_is_defined_by

"phenomenal places"
Note that it is possible for a place to be defined by phenomena causal to it or other forms of identification rather than by an instance of E94 Space Primitive.

E93 Space Primitive

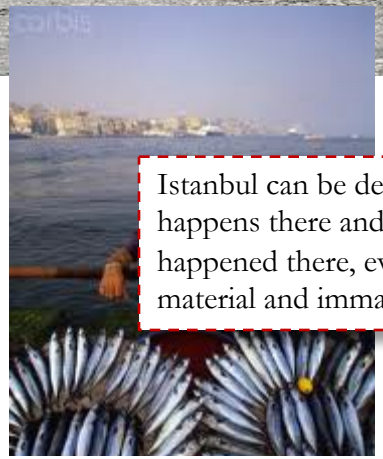
This property associates an instance of E53 Place with an instance of E94 Space Primitive that defines it.



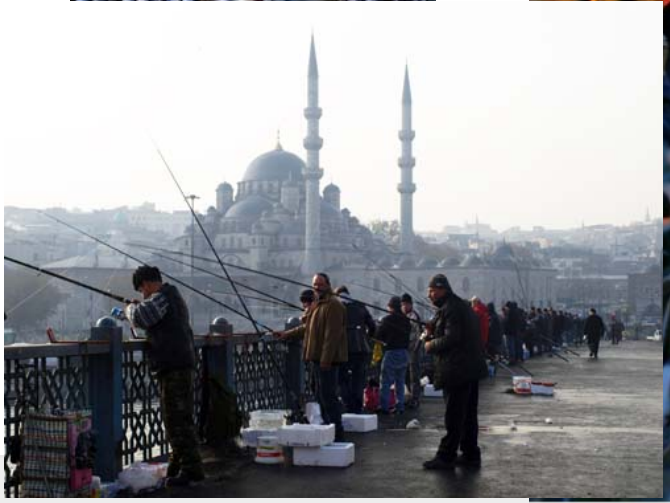
Fort Duquesne, Fort Pitt,
Fort Trent, The Manor of Pittsburgh
Pitts-Bourg, Pittsburgh, ...

Place is not Location, and places are not defined by their location. The red markers indicate a representative location for the cities of Pittsburgh and Istanbul, and that is all they represent.

Byzantion, Byzantium, Constantinople, Constantinopoli, Constantinopolis, Costantinopoli, Estambul, Istamboul, Istambul, Istanbul, Konstantinopel, Konstantinoupolis, Kustantiniyah, Mikligard, New Rome, Stamboul, Stambul, Tsargrad, Tsarigrad, İstanbul, ...



Istanbul can be defined by what happens there and what has happened there, evidenced in material and immaterial culture.



Pécs, Hungary

Pécs always was a multicultural city where many cultural layers are encrusted melting different values of the history of two thousand years. Hungarians, Croatians and Swabians still live in peace together in economic and cultural polarity.

The same goes for Pecs, Hungary, which is typical of all places in a way – what is there is evidence of what has happened there and what happens there now.

The structure important to several religions – dominant or ascendant at different times

Cultural practice of performance, costumes

(Some) of what happens there now...

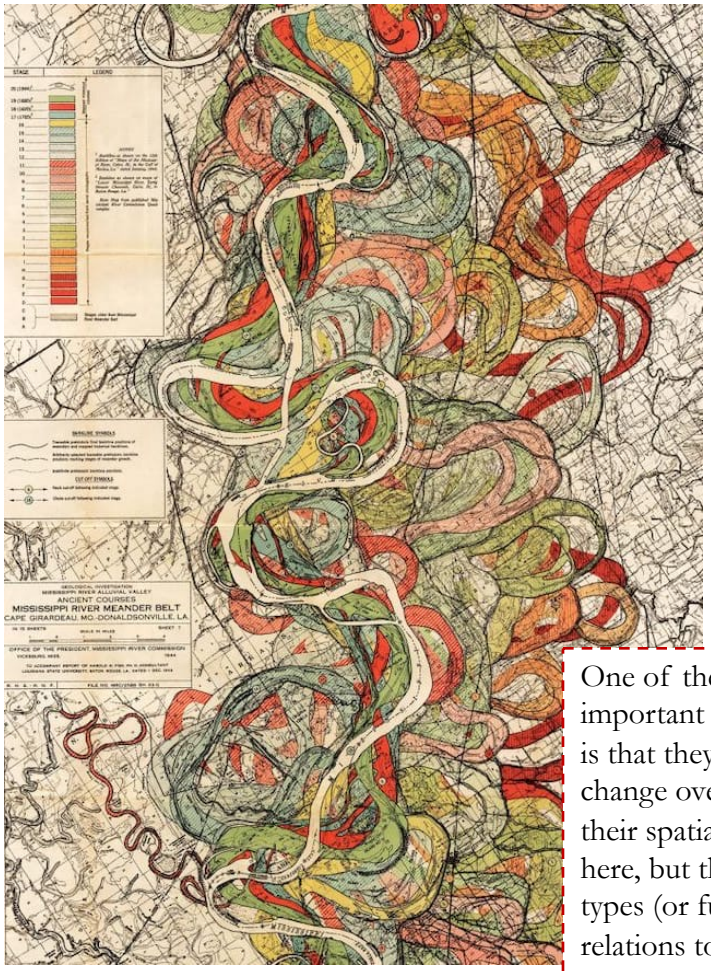
Purves, R. S., Winter, S., & Kuhn, W. (2019). **Places in information science**. *Journal of the Association for Information Science and Technology*, *70*(11), 1173-1182.

"A **place** is an object resulting from **a shared identification of a location**. As an object, it may become a part of a network and participate in events."

"An **object** is **a uniquely identifiable entity existing in space and time** and having well-defined properties as well as relations with other objects."

I very much like this recent paper on 'Places in information science.' I particularly like its pragmatic approach, relating the difficult term 'place' with information systems – each of which have particular requirements.

"...we can productively work with many existing definitions about places and move toward **a shared understanding of the general properties expected from information systems dealing with place.**"



Mississippi River

One of the very most important attributes of places is that they are dynamic – they change over time. Not only their spatial extents, as shown here, but their names, their types (or functions), and their relations to other places – their membership in networks and their parthood in regions and territories.

"Poland"

World Historical Gazetteer

"What do humanists want?
What do humanists need?
What might humanists get?"

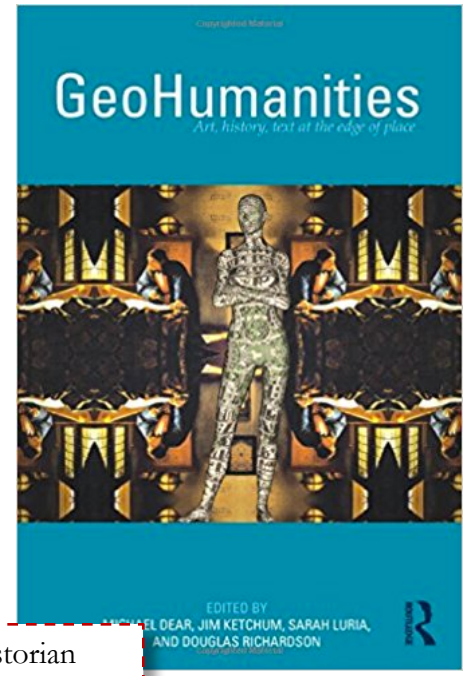
Peter K. Bol (2011)

In Dear, M. et al (Eds.) *GeoHumanities: Art, history, text at the edge of place*.

a. Maps and some spatial analysis

b. A world-historical gazetteer

c. ??? ("we'll see")



In 2011, Harvard historian Peter Bol asked these questions in a chapter with that name. His answers follow.

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

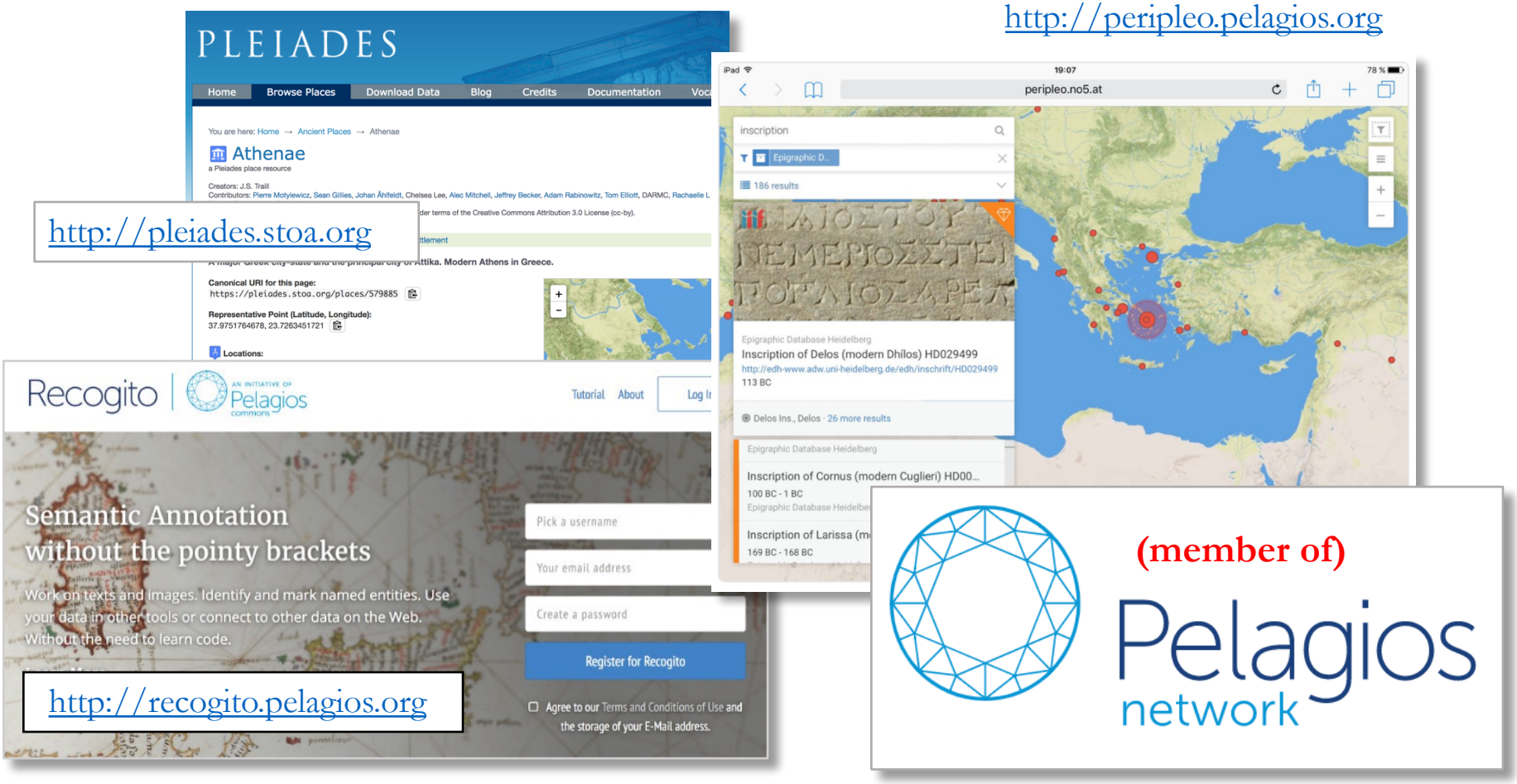
Karl Grossner
Technical Director & Lead Developer

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

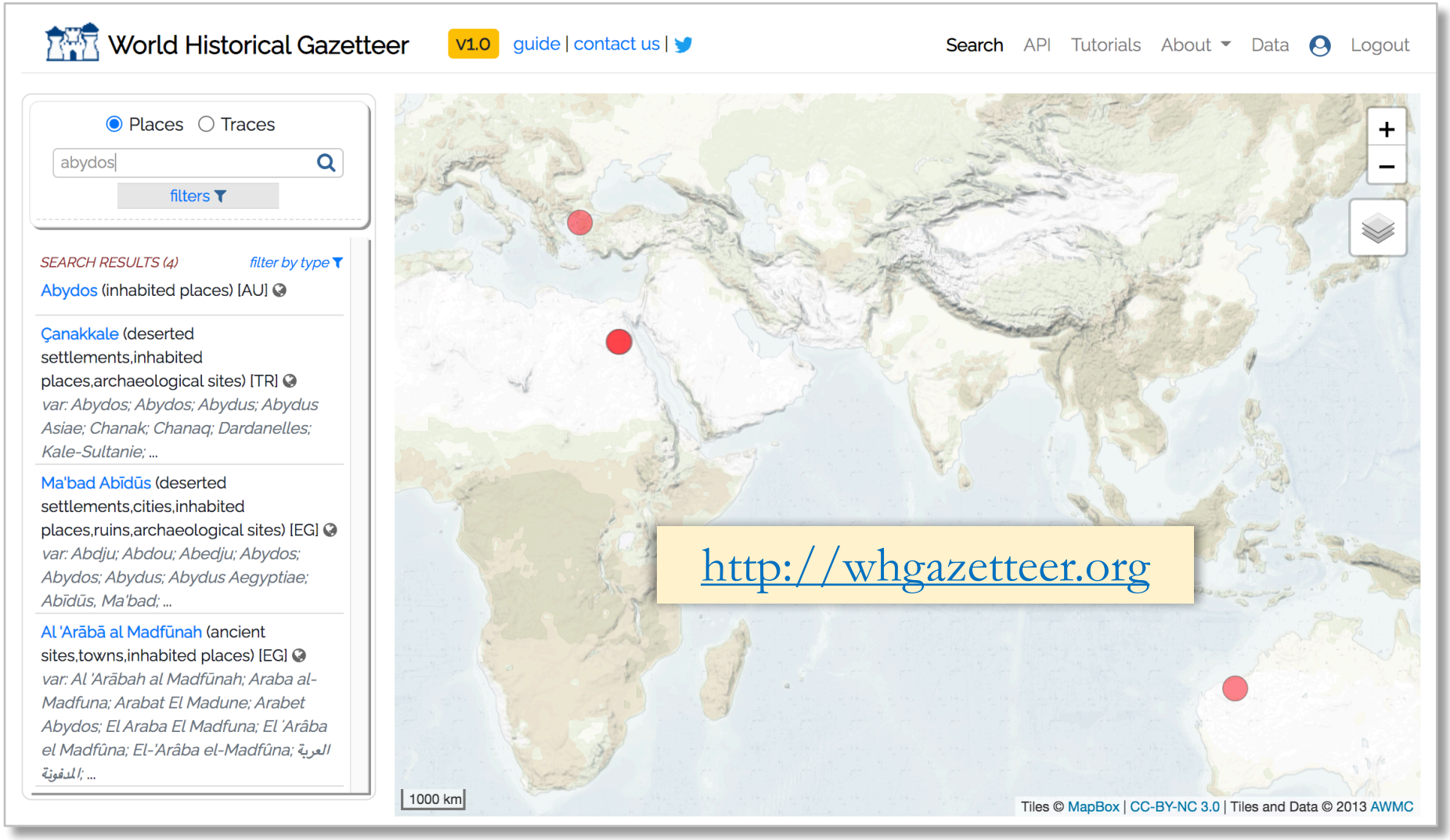
Trace annotations

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)

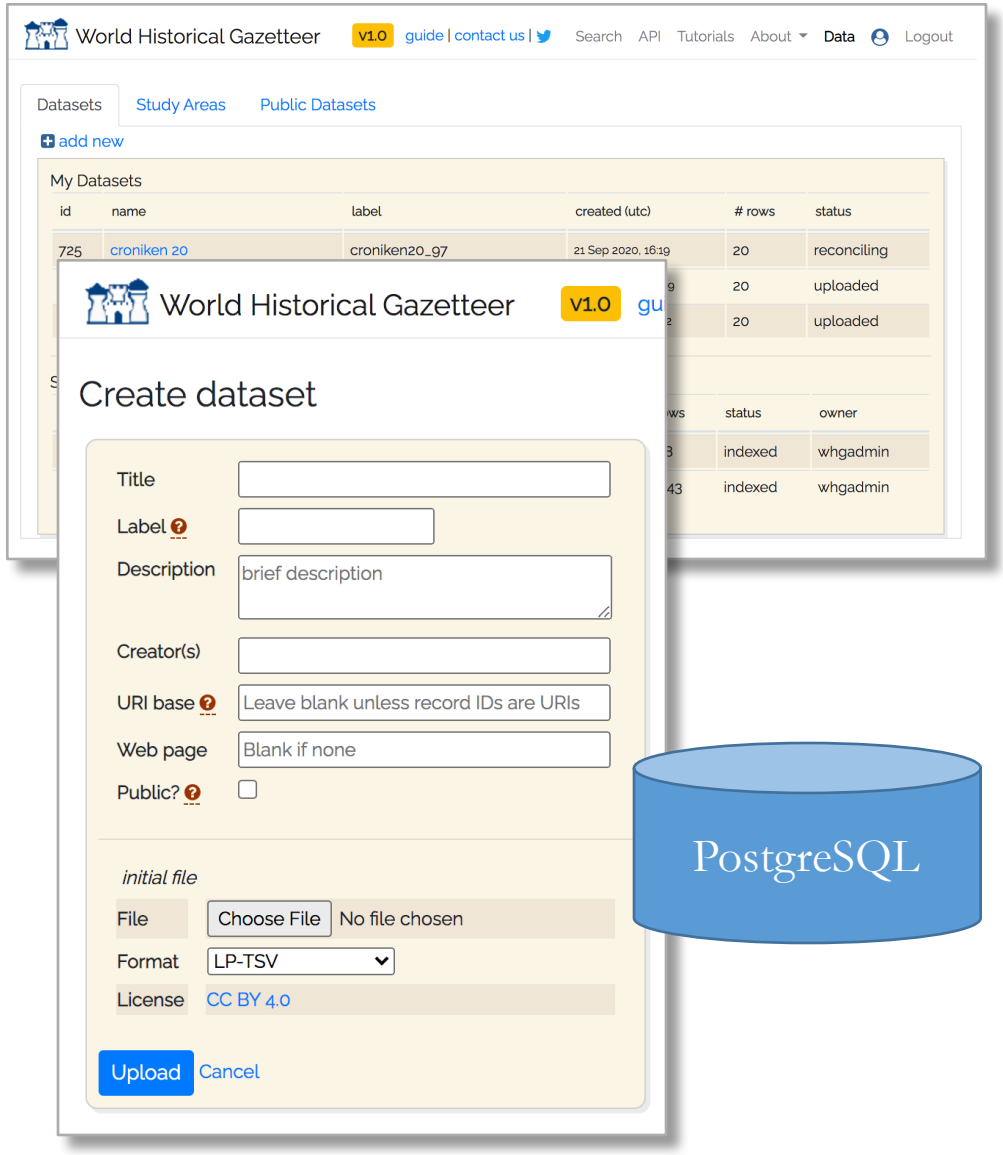
D-Place

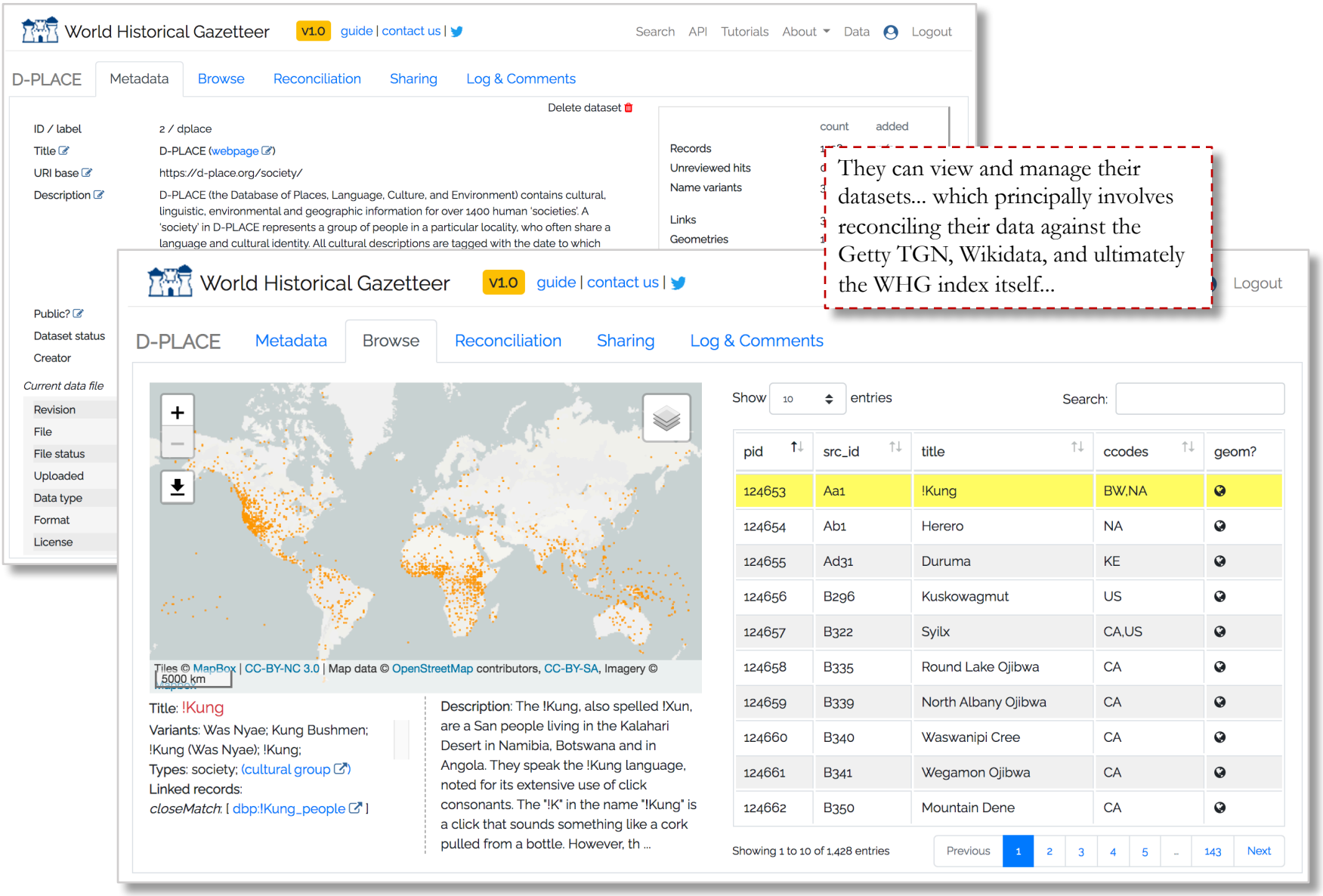
Enslaved

Old World Trade (OWTRAD)

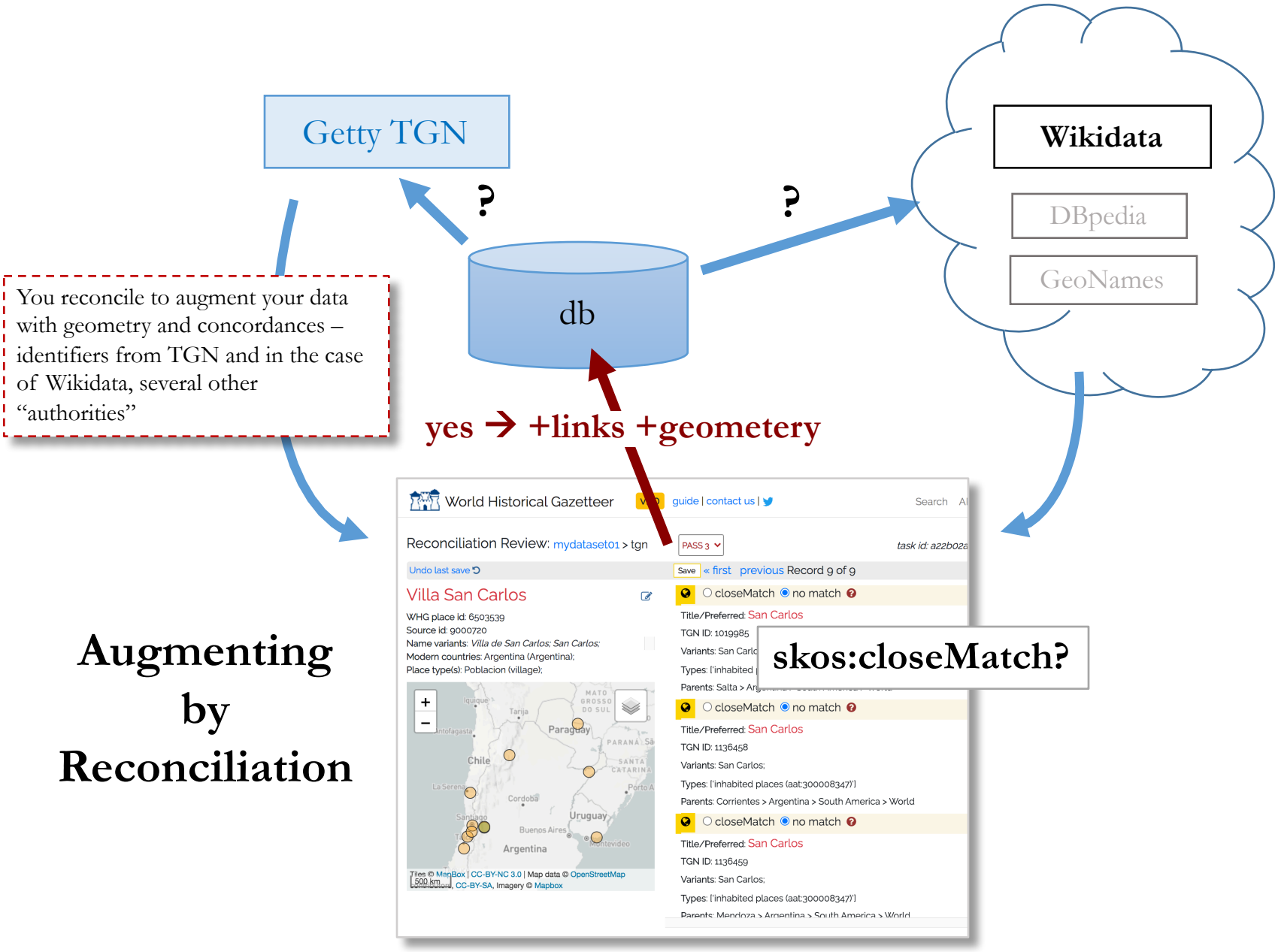
Pleiades

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



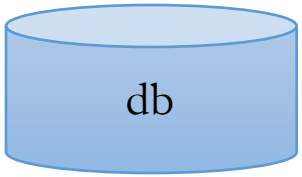


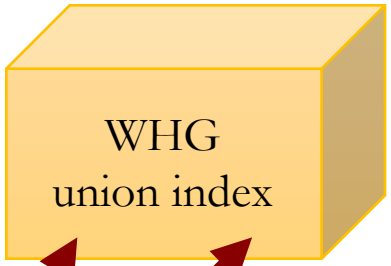
They can view and manage their datasets... which principally involves reconciling their data against the Getty TGN, Wikidata, and ultimately the WHG index itself...



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)



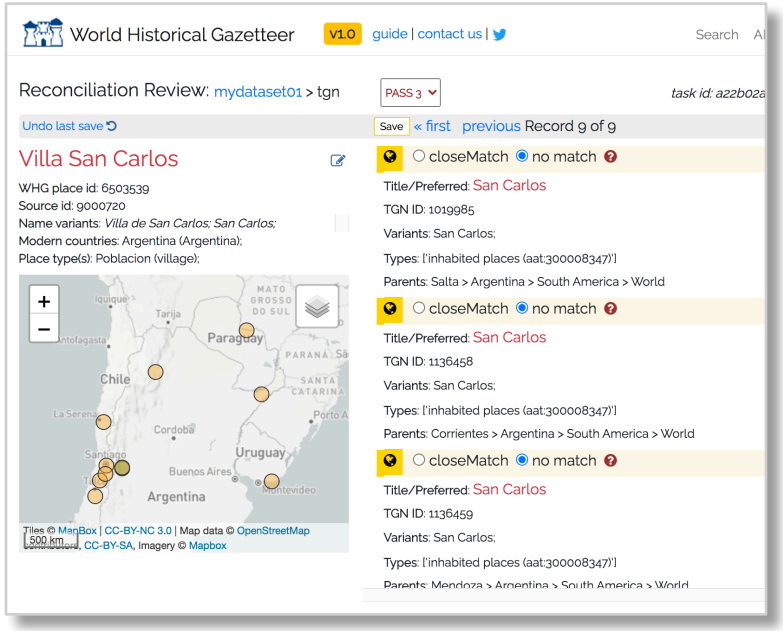


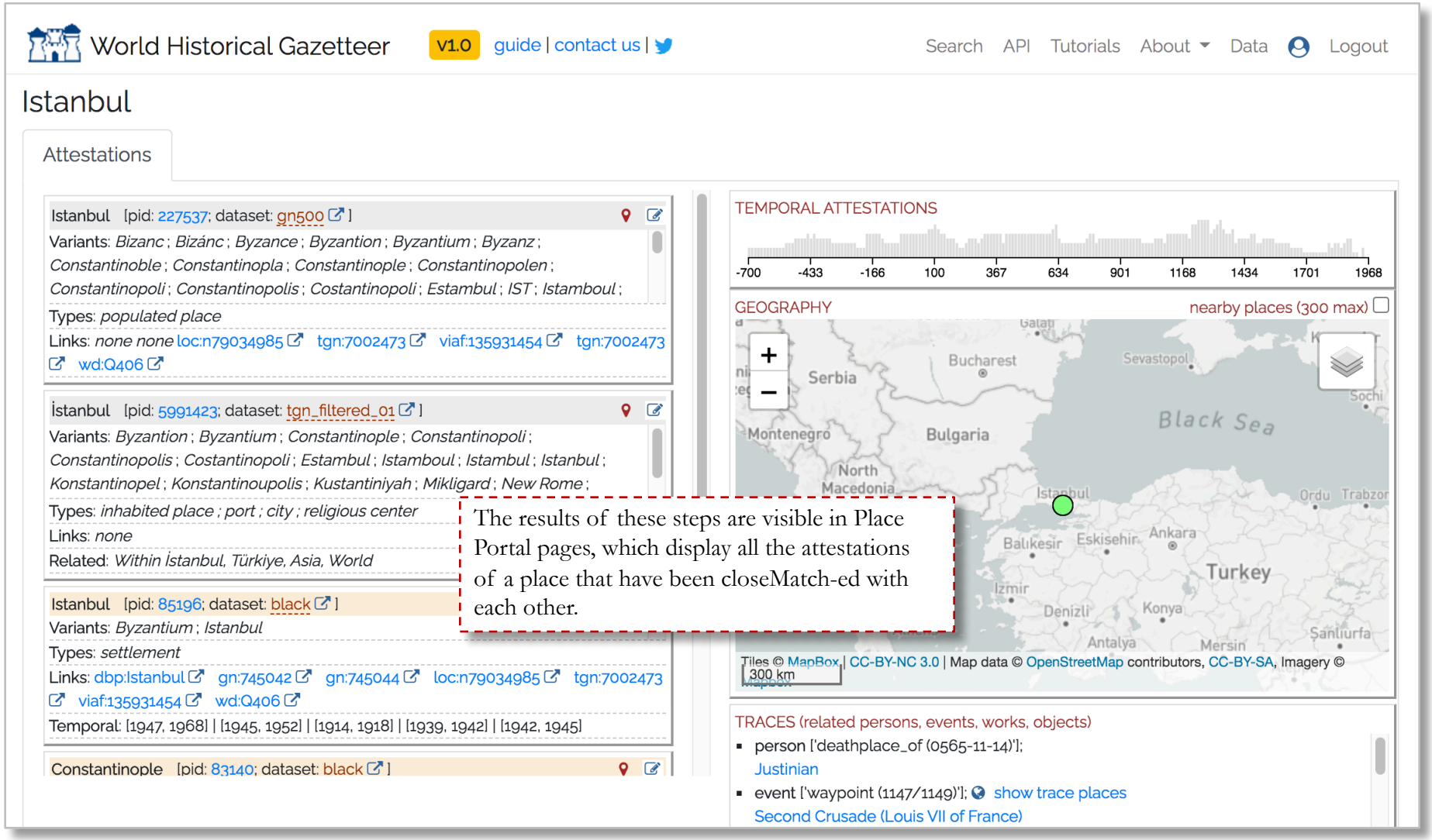
yes → child

no → new record

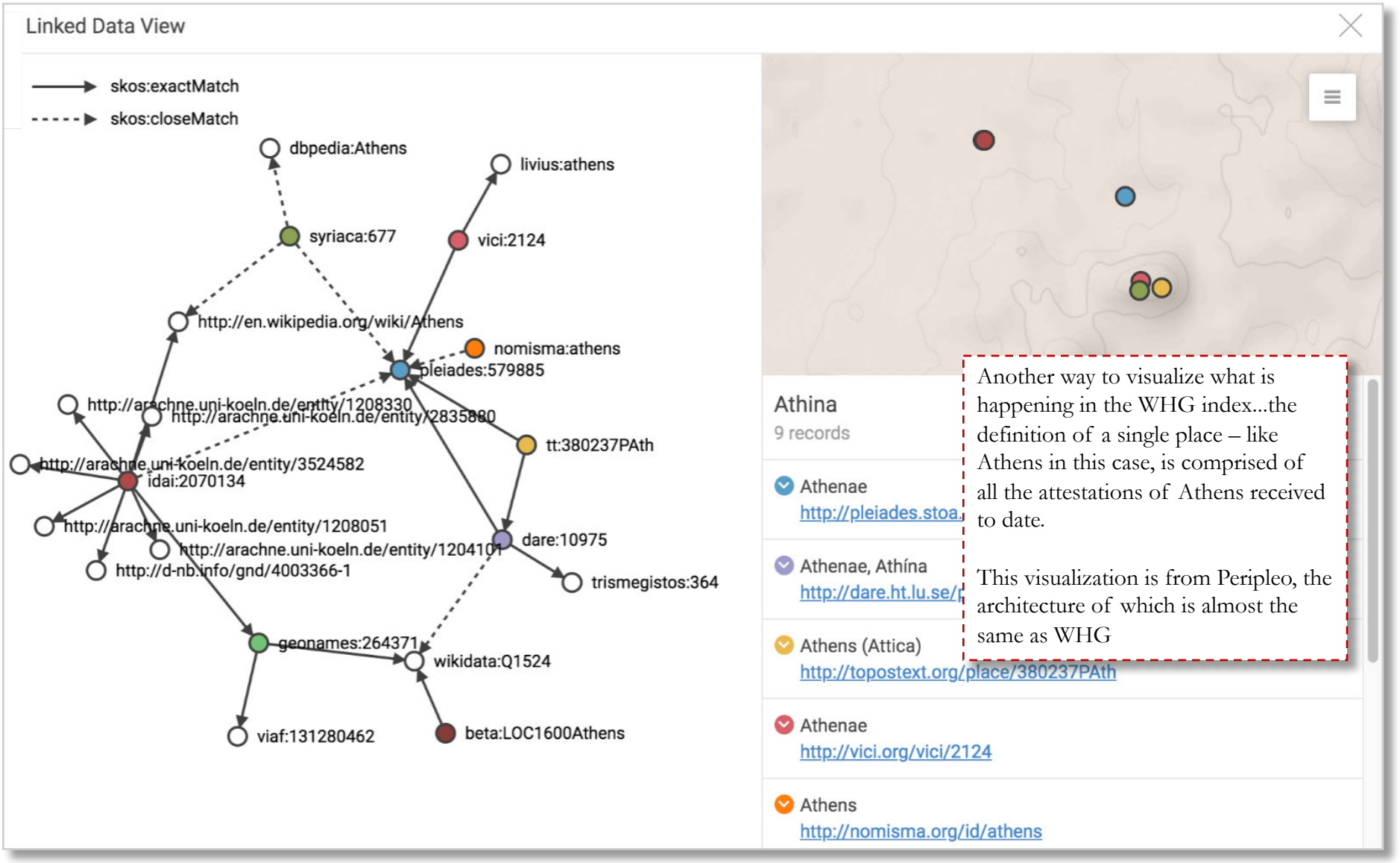
skos:closeMatch?

Accessioning to WHG index

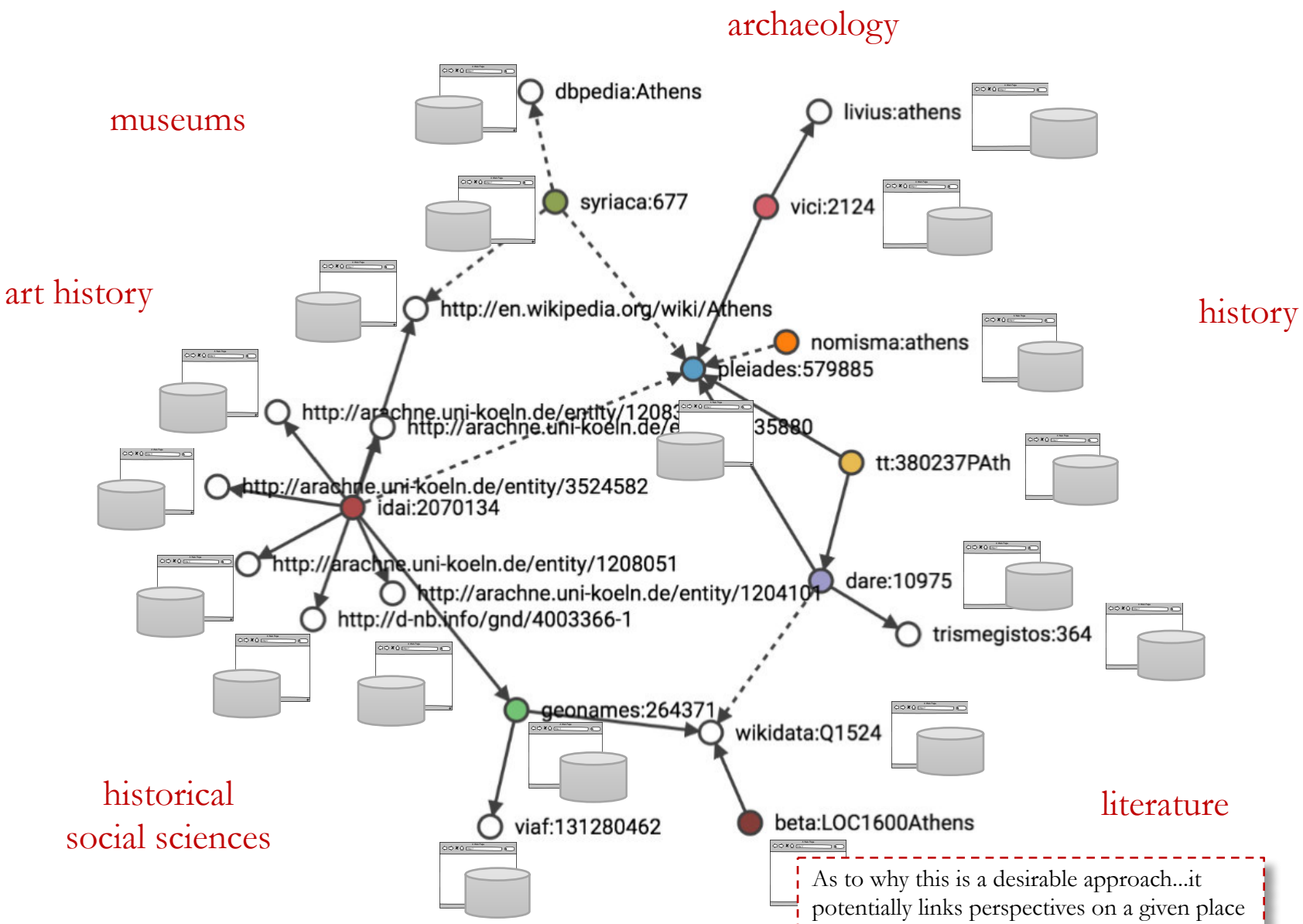




The results of these steps are visible in Place Portal pages, which display all the attestations of a place that have been closeMatch-ed with each other.



http://peripleo.pelagios.org



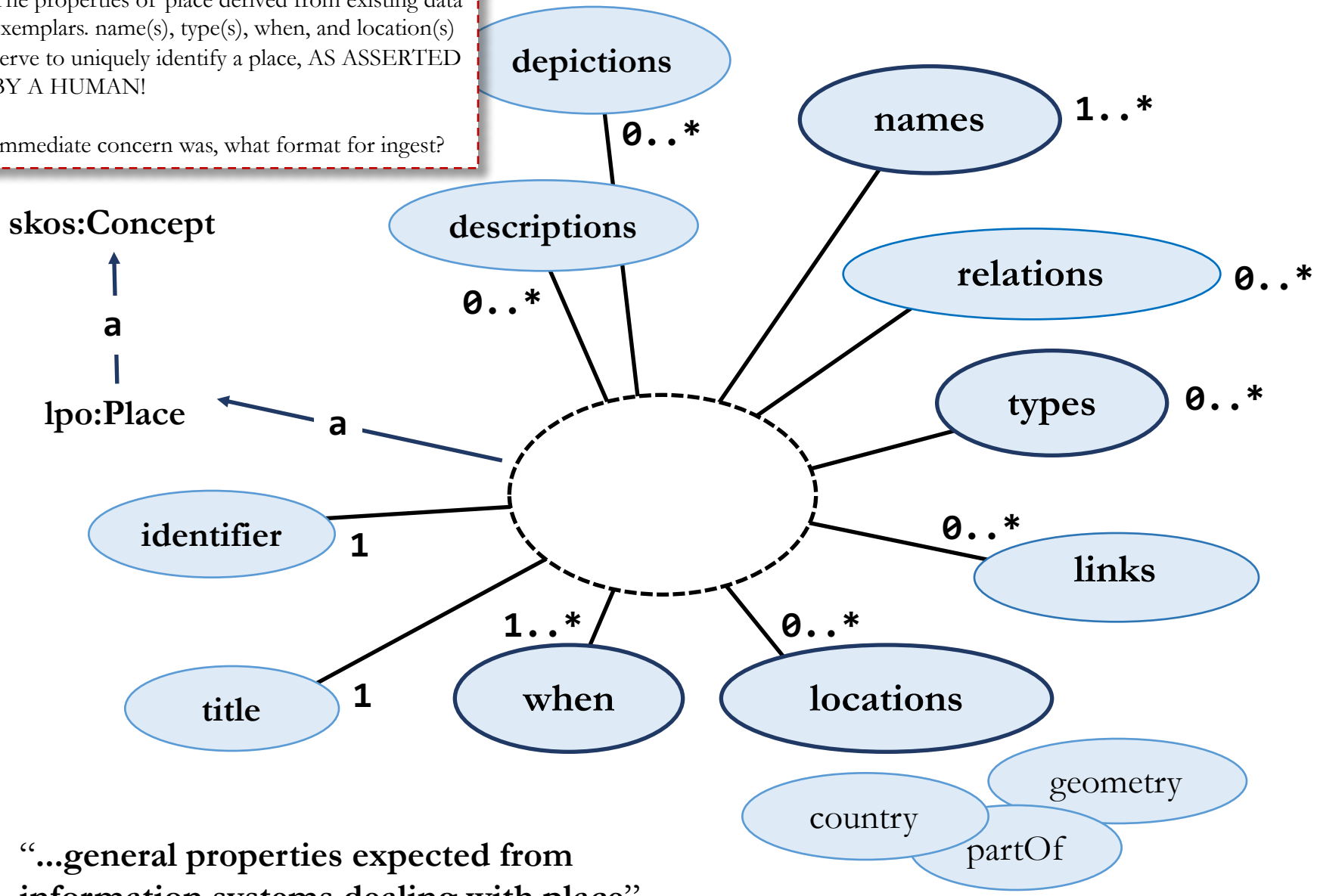
As to why this is a desirable approach...it potentially links perspectives on a given place from many fields

Linked Places

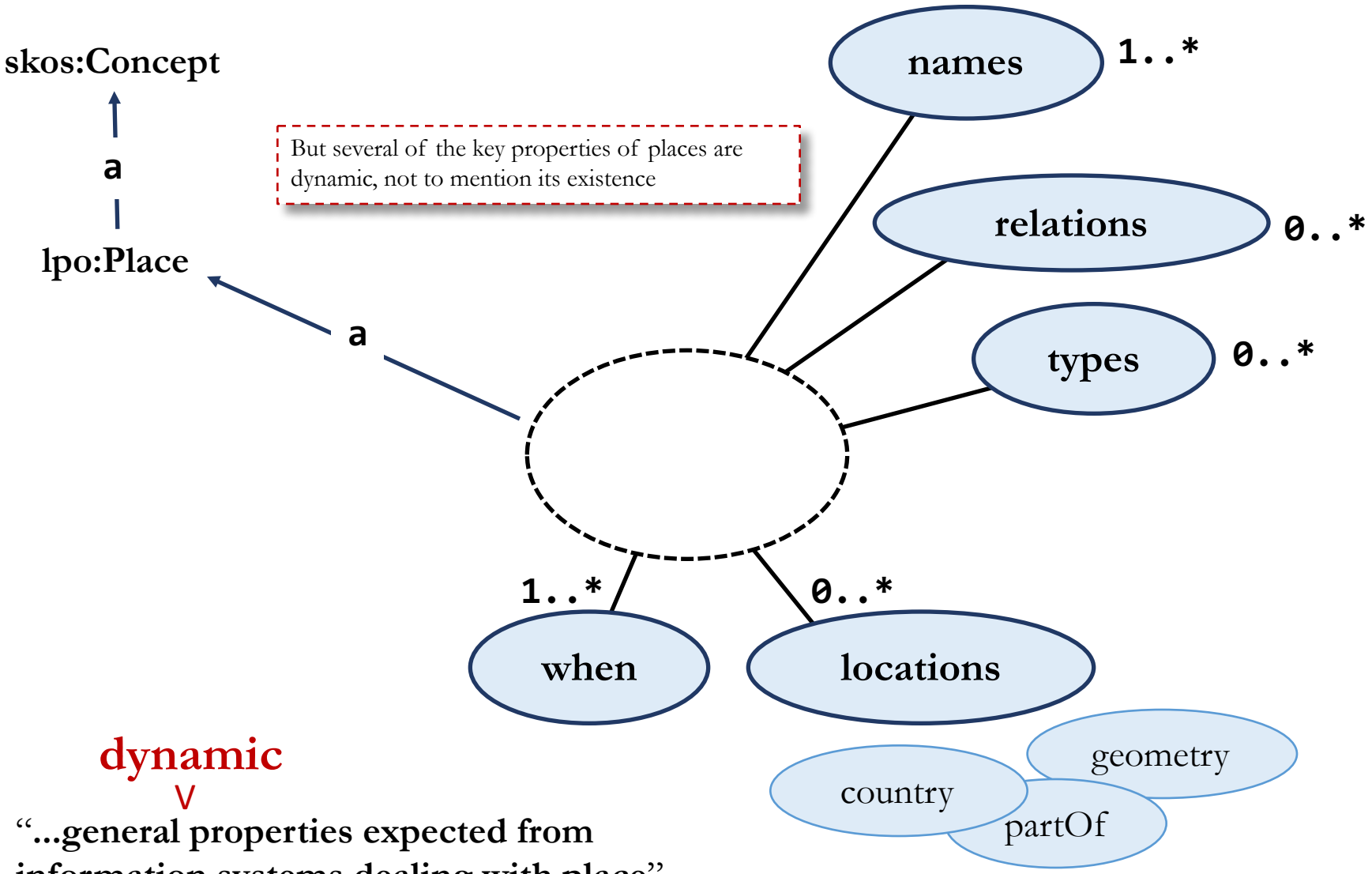
A model and a format

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"



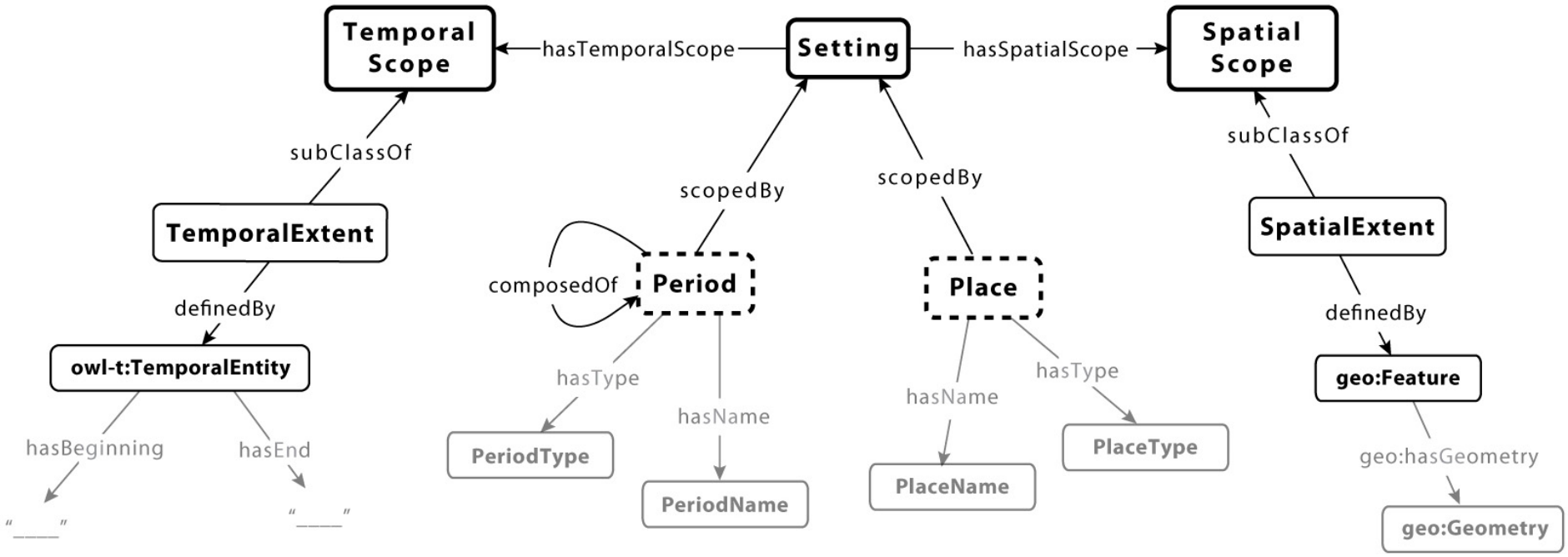
dynamic

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/

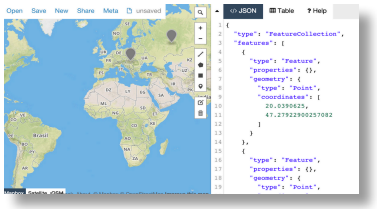
I needed a data format to join space and time – why start from scratch? GeoJSON is very widely used for web mapping applications.

> 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"
}
```

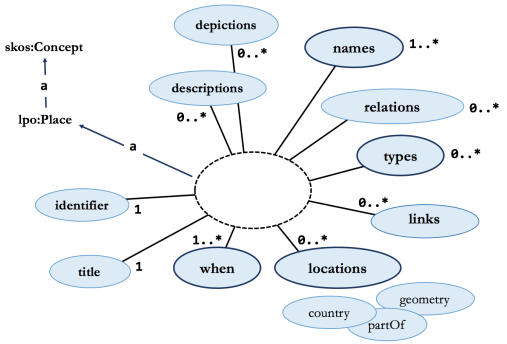
I then proposed some standard properties of "when" objects: timespans, named periods, duration, follows for sequences, and a label

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

GeoJSON-T

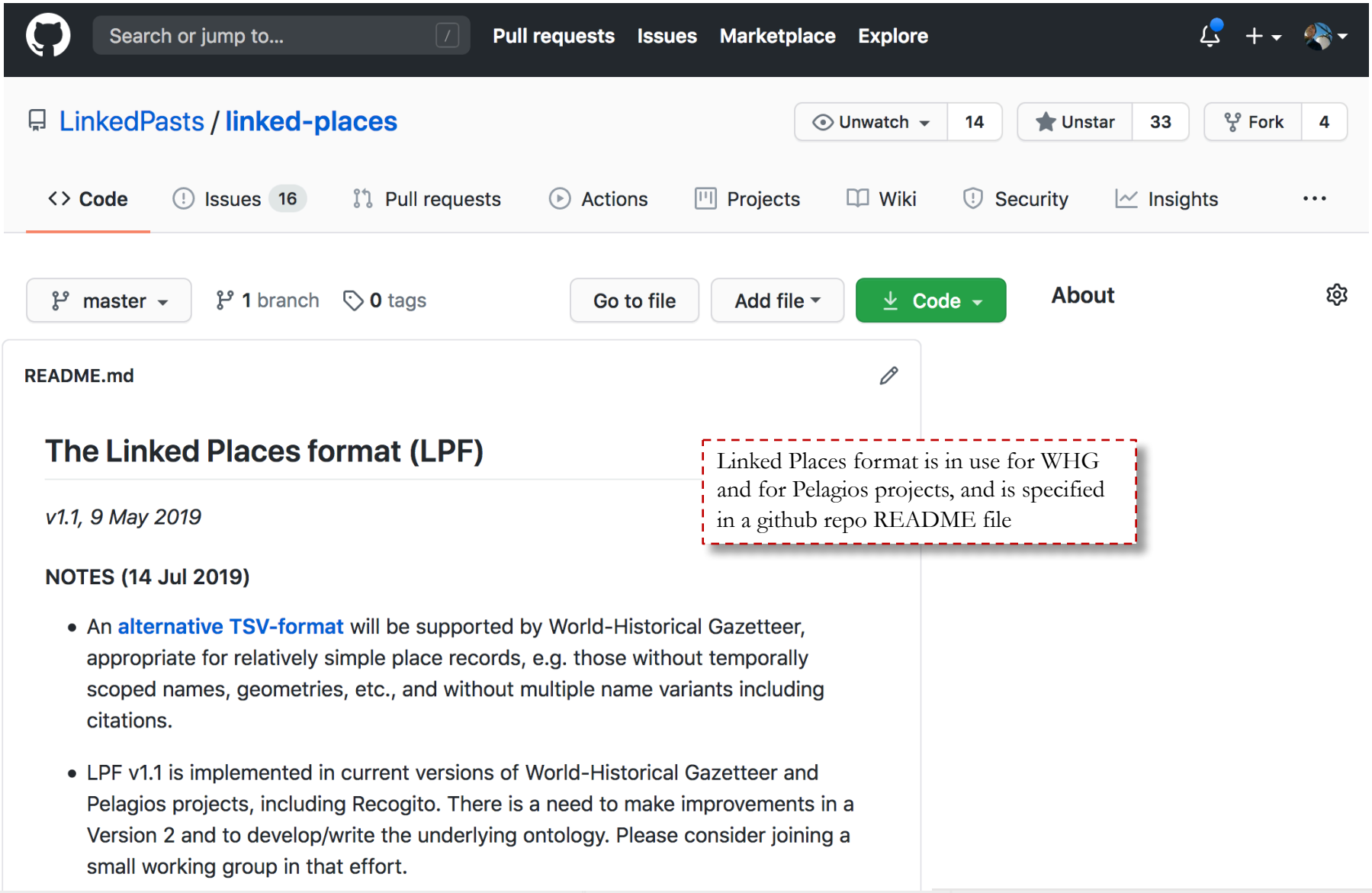
```
{
"type": "FeatureCollection"
"@context": http://linkedpasts.org/lp-context.jsonld,
"features": [
    {
      "type": "Feature",
      "properties": {"id": " ", "title": " " },
      "geometry": { ..., "when": {  } },
      "when": {      },
      "names": [{ ... , "when": {  }}],
      "types":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

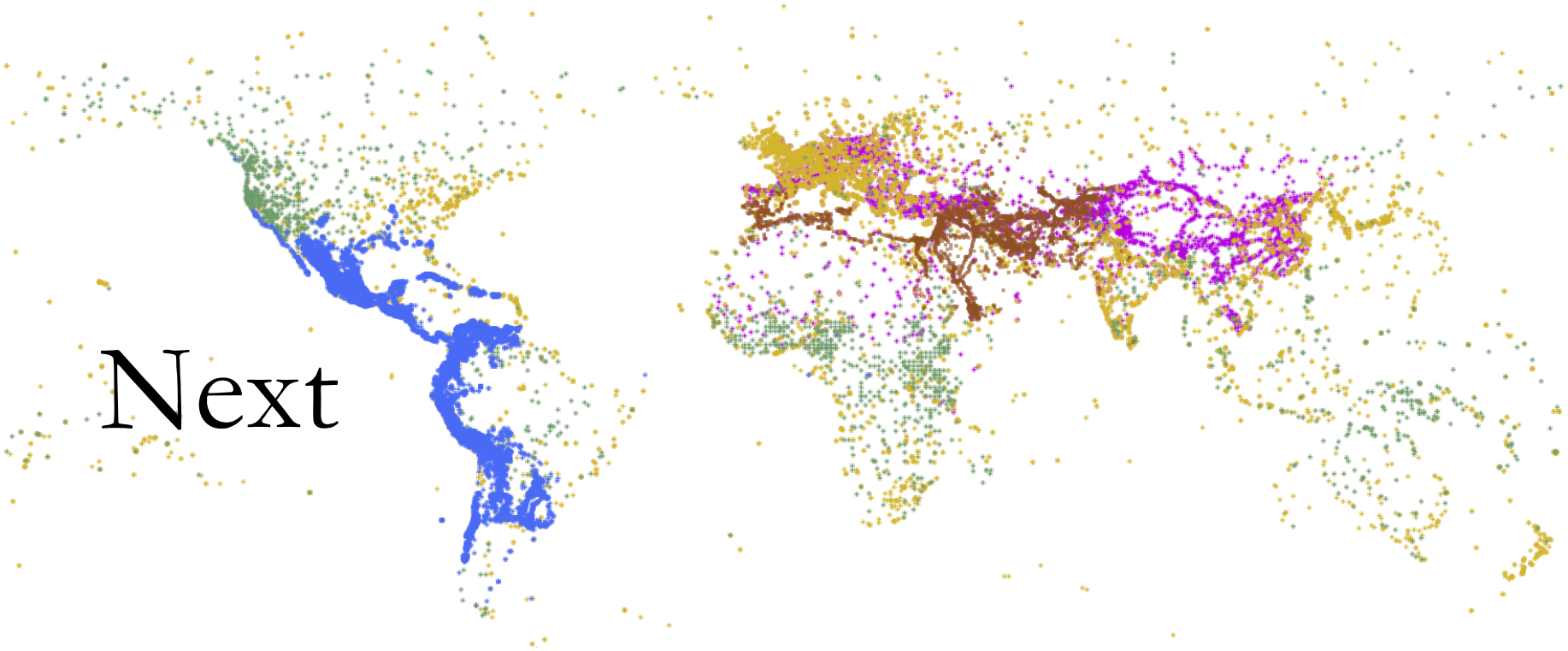
There is a draft context file, which refers to a Linked Pasts Ontology (lpo:)...

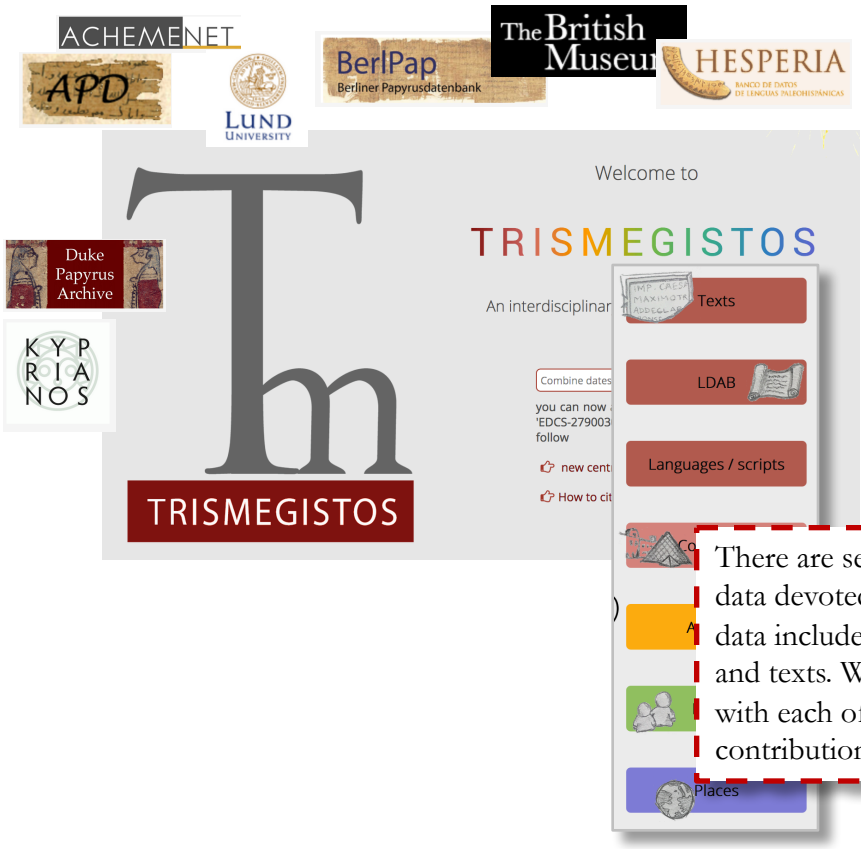
lpo: <http://linkedpasts.org/ontology#>

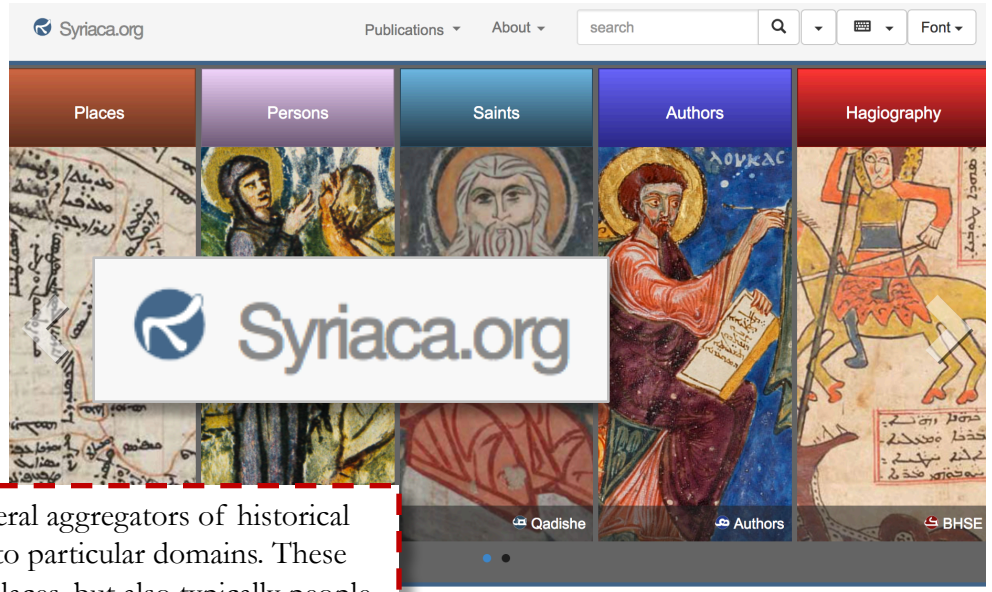
```
## LPO version 1.1. Richard Light, started 13 March 2020

@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 "" .
```

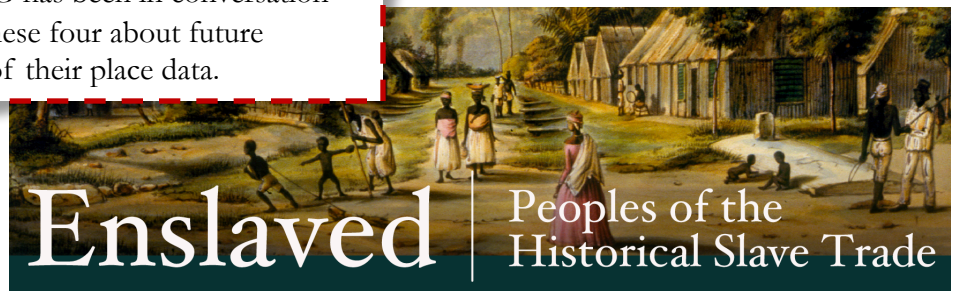
The Linked Pasts Ontology is essentially aspirational at this point. I realized as this process went along that the order of development was unusual, or unorthodox, or both. There is a format in use to describe places, which has an underlying formal model that has not yet been committed to valid RDF. 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.

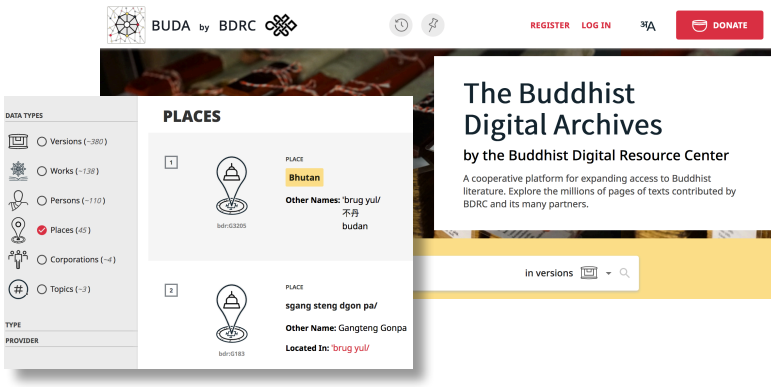

Next



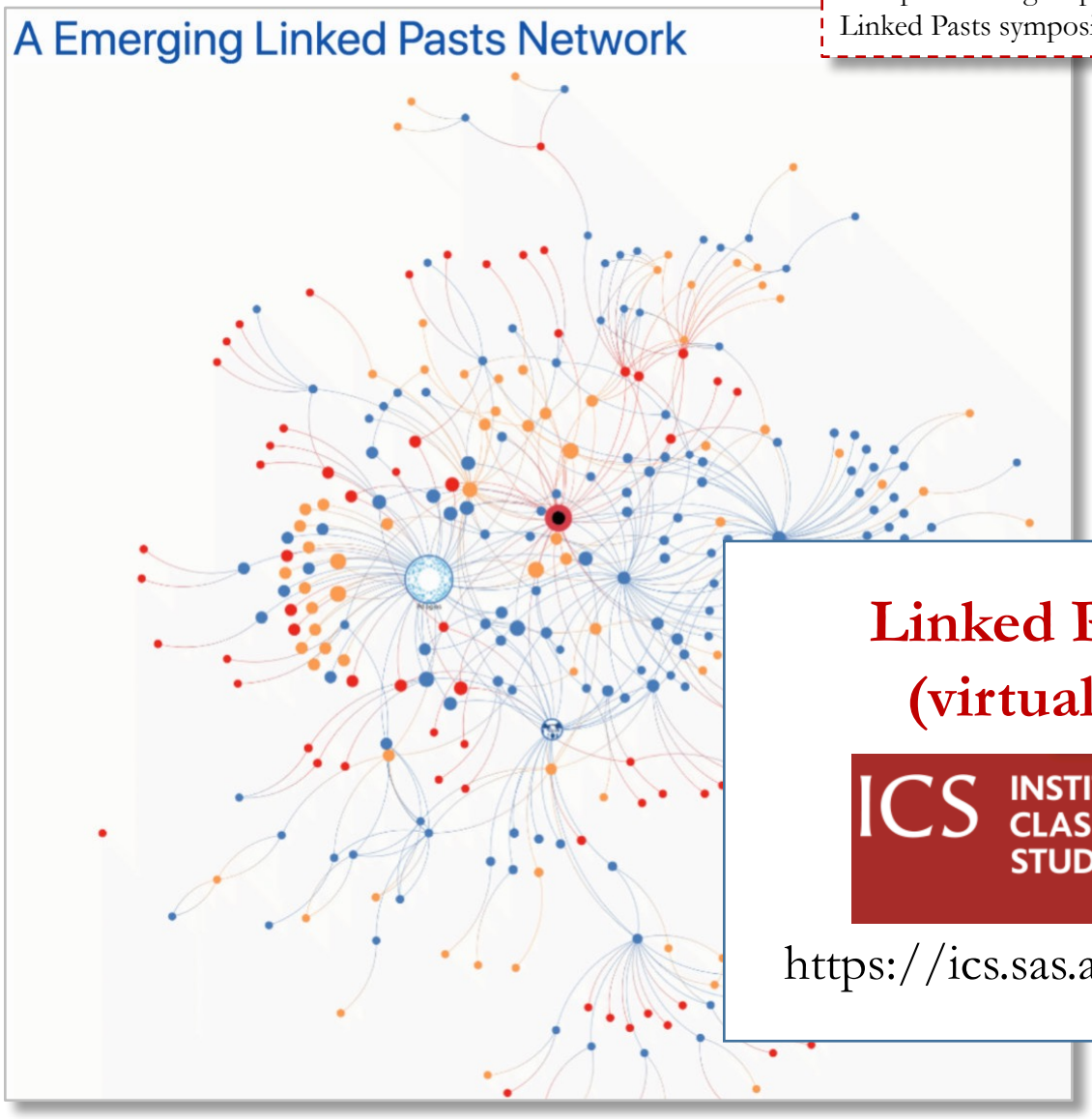


There are several aggregators of historical data devoted to particular domains. These data include places, but also typically people and texts. WHG has been in conversation with each of these four about future contributions of their place data.





- 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