

Guidelines

About these guidelines:

- We envisage the guidelines not as prescriptive, but rather as a checklist of questions and issues that someone starting on an Urban Gazetteer should somehow formulate answers to or be aware of. (We do however give several options and sometimes give preferences)
- We also envisage that people starting on an Urban Gazetteer may have hardly any experience with Gazetteers in general, which is why we aim to explain every step on the way
- We depart from the idea that Urban Gazetteers are not essentially different from settlement (or other larger scale) gazetteers, but we have experienced that they are in some respects more complex (e.g. chronology, hierarchy, toponyms), which is why we consider it worth the while to treat them as a separate category
- We especially aim at guidelines and consequently also at Urban gazetteers that are 'multidisciplinary', i.e. that take into account that scholars from different disciplines have another take on what places are, and that enable them to communicate with each other
- These guidelines are aimed at the Pelagios Community

What is an Urban Gazetteer?

Two features are essential in our eyes:

- An Urban Gazetteer in its most basic form is a list of items that are conceived of as *places* within (the boundaries of) a city
- An Urban Gazetteer generates URI's (Uniform Resource Identifiers) for each item on this list (which means that it in any case needs to be connected to a web domain to which the URI leads)

Note: other attributes (descriptions) can be added to the places on this list (and are regularly added: such as geolocations, variant names, references etc.), but are not strictly necessary and can be decided on depending on scope and goals of the gazetteer (see later)

What is the function of an Urban Gazetteer?

- An Urban Gazetteer serves to *identify* places within a city
- An Urban Gazetteer serves to *link* digital resources that use the URI's of the Gazetteer to describe/identify the places in their project (resource/database) [either within a single resource or between several resources]

So: in other words, why would you want to make an Urban Gazetteer?

- You have (or plan) a digital project in which places within a city play a certain/prominent role, and you want to make sure that all the information (metadata) about each place is attached to the correct and unique place 'record'
- You want to link two or more separate digital resources containing information about places in a city, so that the information about each place in these resources can be connected, merged, shared or exchanged
- [You want to visualize the places in your project(s) on a map, which is facilitated by an Urban Gazetteer in which the places are georeferenced]
- [You want to do various kinds of queries – visualizations]

Checklist for Urban Gazetteers

A. Goal

What is the purpose of your Gazetteer?

- See the options above:
 - o a starting – single purpose – gazetteer Note: a single purpose gazetteer in practice is often much more than a gazetteer proper, but integrates a gazetteer into an environment that does a lot of other things as well, often with a specific focus (e.g. MoEML, Mapping Visions of Rome).
 - o a super-gazetteer (to link place data already present in existing resources)

B. Intended audience

Who do you envisage will use the Gazetteer?

- Only you
- The people that use / view your specific project?
- The people from other projects that would like to connect to it?

C. Content

C1. Which items

– or: in other words: which places (do you envisage) will be part of your Gazetteer?

E.g.: all the places in Naples mentioned in the Atlante di Napoli

E.g.: all the churches in Rome built in the 12th century

E.g.: all the places in Athens that are now part of existing digital resources x, y and z

E.g.: all the structures identified at a certain archaeological site

Some advice:

1. Start with a manageable amount of places (manageable in terms of the amount of time and money you have available)

E.g. it is often useful to start with a proof of concept, a pilot that works, which can be the basis to apply for more funding

2. Consider the 'representativeness' of the dataset you envisage. To what extent is it random or given by certain biases / circumstances? How can you select a dataset in such a way that it is also possible to draw conclusions from what is on the list? [example]

E.g. If you cannot add all the places you envisage in one go, try to proceed step by step (layer by layer, level by level), rather than randomly adding places. It may make more sense (this is especially relevant users who might draw conclusions on the basis of the set) to first add all churches and then proceed to all printshops, than first adding all places that are on pages 1-10 of a certain topographical dictionary. [In the same vein, it may make more sense to first add complete buildings, before moving over to regions or parts of buildings.]

Of course there may be good reasons to do it differently, but then make sure you are transparent about the nature of your data selection

C2: Where are you going to derive your content from

-- from what resource will you draw the places that will be part of your Urban Gazetteer?

E.g: a print resource or digital resource:

- a topographical dictionary (all places in Rome in the LTUR)
- a text (e.g. all places mentioned in the Mirabilia)
- a map (all places mentioned on a historical map of Rome)
- an existing digital resource (all places in Naples that are part of Census)

Advice:

- Start by checking all the digital resources concerning the city you work on, and see to what extent these data can be used. (E.g. Word Gazetteer / Pleiades / Wikidata / Geonames)

C3. How are you going to turn this content into an Urban Gazetteer

(how will you transform them into a list?)

- OCR
- Manually typing them
- Using export (dumps) from existing resources (but then the hard work starts: how will you merge and align them? how handle duplicates?), if possible filtered for a certain category (e.g. from the Pleiades gazetteer select only the ancient temples inside the city of Rome)

Links to: technical options, e.g. whether to use a database with easy interface, how to export and import etc.

D: Data Model

The answer to the questions:

- How do you [conceptualize and] define the places (= items that each get their own URI) on your list?
- How do you circumscribe/describe these places (what kind of attributes/metadata do you add to each item)?
- How do envisage [conceptualize] the relations between place (the item) itself and its different attributes?
- How do you envisage the relationship between the various places (the items)?

Note: most of the issues that are specific for Urban Gazetteers will come up in this section, since defining what a place is within in city (e.g. when it ceases to be one item but should be broken up into two or more items) turns out to be very complex.

Advice:

- Before you start, check out existing Gazetteers and see how they conceptualized this – what is their datamodel?

Question 1: How do you define what counts as a place in your gazetteer, or in other words: what kind of 'items' each get their own URI on your list?

Help: If the city is the highest level that defines your gazetteer, what then is the next level thing that you distinguish: a name, a building, a coordinate, a location?

Subquestions:

- Do you consider a building as place (= item), or do you distinguish between buildings (as objects) and their geographical location (expressed by coordinates) as separate items (each with their own URI)
- Do you consider a toponym as a place (= item)?
- Do you only consider separate buildings?
- Do you also consider higher level entities, like regions?
- Do you also consider parts of buildings as separate places (a frieze of a temple?)

Preference:

We do have a preference for (a version of) the Pleiades model, because:

- it has a concept of place that can be both very specific (e.g. a temple) as well as general (e.g. a settlement)
- it has a concept of place that can have various attributes, but can also exist (and function as link) if some or most of the attributes are missing. Thereby it is especially useful to link resources of different kinds (archaeological/material data as well as literary toponyms)

Place:

Something can count as a place within a city:

- If it belongs to a city = is of a lower hierarchy than the city level
- If its existence (whether in reality or in the imagination) is referenced by at least one source

NB: it can range from the material remains of a building without a specific name but that is mentioned in an archaeological report, to a place in the city that is mentioned in a medieval legend, but of which no location is known, or that may have never really existed.

In an Urban Gazetteer each place item gets a URI and is assigned a name. However, this name should not be conceived of as 'the real name' or a separate attribute, but rather as a conventional name, by which humans (instead of machines) can talk about this place. It is however advisable to decide on beforehand where to base this conventional name on (and which language to choose).

Attributes

A place (= a single item) can have various attributes, of which the following three belong to (as we consider it) the core of an Urban Gazetteer:

Name: the name(s) by which this place is known

Geolocation: the geographical location that this place is known (or thought) to occupy

Reference(s): the source(s) from which this place is known

NB:

- Places can have more than one of each of these attributes (more than one name, more than one location e.g. in the case of different suggestions in secondary literature, more than one reference).
- Places can also lack one or more of these attributes: they can be unnamed (if it concerns material remains that are not identified - only known as e.g. House 121) or unlocated (e.g. only known from literature but location unknown).
- However, places can not be unreferenced: there need to be at least one source (whether material or documentary) to justify the existence of the place as item in the gazetteer

Name:

The name(s) by which this place is known (i.e. all names pointing to one place have to be grouped in this one item)

You have to consider how you want to deal with different names, different spelling, changing names, names in different languages, same names for different places, differences between primary and secondary sources, places with no (original) name etc., so as to enable people to find the place they are looking for.

Advice: see how MOEML has added all different names to the records

Geolocation:

The geographical location that this place is known (or thought) to occupy

You have to formulate an answer to the following questions:

- What do you mean by location? e.g. is it a point, a polygon etc. Advice: use standard notation and clarify which notation you use.
- What do you base the geolocation on? E.g. material remains, archaeological reports, literary/documentary sources etc.
- What is the positional accuracy of your geolocations? E.g. certain, uncertain, approximately

You will probably be confronted with issues like the following:

- If you find a reference to 'the house of Caravaggio' in two different documents, none of which mentions a location, you cannot be sure that the same place is referenced. Do you make one place 'House of Caravaggio' with the two references combined or two: 'House of Caravaggio A' and 'House of Caravaggio B'? What you choose is partly dependent on whether you want to count locations in some way.
- In case of a region it may be impossible to draw a line. In such cases it may be possible to define it as the area to which places x, y and z belong. See: hierarchy

References

You will have to formulate an answer to the following questions:

- What kind of references do you want to add for the authority of your items?
 - Primary sources, secondary sources, links
- How many references do you want to add?
 - Just a few vs. as much as possible?
- How do you distinguish between references that are 'necessary' to justify the existence of the place in the gazetteer and references that offer the user extra information (links - further reading - images)?

Further attributes / metadata:

Beyond these basic attributes that are typical for Urban gazetteers, what kind of further metadata do you want to add?

Or: how generic do you want the gazetteer to be vs. how much will it be tailored to a specific research plan?

Examples of further attributes:

- Type of place. To which category does the place belong (e.g. church, temple, bridge ..). If possible, use existing standards (ontologies), like the Getty Thesaurus
- Images
- Dates (e.g. when something was built / demolished) See further: chronology
- Relationships with other places (e.g. belongs to / is part of / is in the neighborhood of). See further: hierarchy
-

Chronology

A question that you will have to consider with regard to the data model: is the time factor important and if yes, where do you put it?

Is it important to know and record:

- When a place was 'built', restored, demolished?
- When a place was named A, and changed names into B?
- When a place named A on location X moved to location Y?

If so: is it possible or maybe even better to record these dates as text (in an information field, to be read by humans) or added as dates with which it is possible to calculate (make a timeline etc.)? (it could be better in instances where a name may be officially changed, but the place is still known with the old name)

If option 2:

- How exact do these dates need to be (exact day vs. period like 'Renaissance')? Advice: use standard periodization (e.g. PeriodO)
- Where do they have to be added in the datamodel: e.g. in the field concerned (name / location?)

Hierarchy - Relationships

Another question that you will have to consider is how and where you conceptualize the - whether or not hierarchical - relationships between the places in your gazetteer.

The places in the gazetteer are likely to relate to each other in certain ways (e.g. close to - belongs to - includes - crosses (with streets) etc.

- How do you conceptualize these relationships? Advice: some relationships are 'standard' in Geojson

Especially if a gazetteer includes places of smaller and greater granularity (e.g. ranging from parts of buildings to regions), places are likely to relate to each other in an hierarchical way.

- How do you conceptualize these hierarchical relationships? Options: an hierarchical tree structure (as in Census.de) or with a relationship that is specified as 'belonging to'.

Consider the following issues:

- What is the goal of these relationships? (e.g. to define a region it may be useful to record which places belong to this region)
- How do you control these relationships? (so that they are not too many and therefore meaningless) Advice: use standards

E: Technical implementation

Database

How will you organize and store your data?

What you choose, depends on:

- How simple or sophisticated/complex is your data model?
- Who is going to add the data; how user friendly do you want the interface to be?

Depending on the answers you can select a database programme (see appendix I: suggestions). This can range from a simple excell-sheet to a relational database with all kinds of built in visualization options

URI's

How are you going to generate URI's: where is your gazetteer hosted?

Export options

Which options do you add for exporting your data? Advice: enable as many as possible, but at least the most regular (csv; geojson)

Visualization

Using a map is a very useful and common method to visualize the places in an Urban Gazetteer, but it is not the only method, and it also comes with a number of challenges:

- Which map(s) do you use? Is it open source or not? Is it a realistic map or stylized? Is it modern or historical?
- Not all places can be located, some places have more than one location: how do you cope with that?
- If you have many places, the map will become too crowded: how do you cope with that?
- How is the time factor included in the map visualization?
- Do you add names on the map?

F: Sustainability

Issues of hosting, longevity, guaranteed stable URI's, documentation etc.

G: Documentation

How do you document your Urban gazetteer (= connected to sustainability as well)?

H: An Urban Gazetteer, and then?

If an Urban gazetteer is envisaged as a means to connect existing datasets, how are these connections made in practice. Will the urban gazetteer function as 'aggregator'? Will it also show incoming links etc.?

Appendix I: Suggestions of tools, programmes, open source software, best practices

Appendix II: Index of terms