Challenge-derived design practices for a semantic gazetteer for medieval and early modern places

Philipp Schneider [a,*], Jim Jones [b,c], Torsten Hiltmann [a] and Tomi Kauppinen [d]

[a] *Institut für Geschichtswissenschaften, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*
E-mails: philipp.schneider.1@hu-berlin.de, torsten.hiltmann@hu-berlin.de
[b] *Institut für Geoinformatik, WWU Münster, Heisenbergstraße 2, 48149 Münster, Germany*
E-mail: jim.jones@uni-muenster.de
[c] *Universitäts- und Landesbibliothek Münster, Krummer Timpen 3, 48143 Münster, Germany*
E-mail: jim.jones@uni-muenster.de
[d] *Department of Computer Science, Aalto University School of Science, FI-00076 Aalto, Finland*
E-mail: tomi.kauppinen@aalto.fi

Abstract. In recent years gazetteers based on semantic web technologies were discussed as an effective way to describe, formalize and standardize place data by using contextual information as a method to structure and distinguish places from each other. While research concerning semantic gazetteers with regard to historical places has pointed out the importance of enabling the creation of a global and epoch-spanning gazetteer, we want to emphasize the importance of taking a domain oriented approach as well – in our case, focusing on places set in medieval and early modern times. By discussing the topic from the historians' perspective, we will be able to identify a number of challenges that are specific to the semantic representation of places set in these time periods. We will then do a survey of existing gazetteer projects that are taking historical places into account. This will enable us to find out which technologies and practices already exist, that can meet the demands of a gazetteer that considers the time specific geographic, social and administrative structures of medieval and early modern times. Finally we will develop a catalogue of design practices for such a semantic gazetteer. Our recommendations will be derived from these existing solutions as well as from our epoch-specific challenges identified before.

Keywords: Linked Data, Digital Humanities, Historical Gazeteers, Historical Places, Digital History

1. Introduction

If the humanities and especially historical research are going to make use of the methods and techniques provided by computer sciences, in order to enhance and improve their methodological scope, they need to represent their domain and the knowledge belonging to it in a formalized, standardized, and machine-readable way. To this end, it is necessary to develop models in order to be able to capture the historical data, to enrich and to process them.

When studying the development of human societies from a historical perspective, three dimensions are of

---

*Corresponding author. E-mail: philipp.schneider.1@hu-berlin.de.

major relevance: people, time and space. In this paper we will discuss how to model historic places and spaces, focusing on the central European area and with regard to the medieval and early modern era, by means of information technologies. We propose the usage of well-established standards that enable data to be easily shared, linked, enriched with data from different domains, and reused for a wide variety of research questions.

Standardization inevitably leads to simplifications regarding different aspects of our understanding of historical places. We must therefore examine how such a need for simplification can be met while preserving the required level of complexity. According to the geographer Yi-Fu Tuan, a place can be understood as a "a center of meaning constructed by human experience" [1, p. 152]. According to this definition, a place in general can be anything as small and concrete as a gate, or as vast and abstract as an empire. Thus to understand and work with places, it is crucial to distinguish them from other ones alongside three dimensions: the geographic extent of the place, its *meaning* and the *construction* of the same meaning by "human experience". The latter aspect, the *construction of meaning*, can manifest in events taking place in relation to a geographic location. The processing of such events by human beings can subsequently lead to *attestation of a meaning* to the place involved in these occurrences. The *meaning* itself is either represented through events that have taken place in relation to a geographic location in the past or in certain material objects, to which a *meaning* has been attested. A simple example to illustrate this could be the former Parisian fortress *Bastille Saint-Antoine*. This building had during the early modern period a (administrative) meaning as a prison and a (symbolic) meaning as an image of oppression. An event, like the Storming of the Bastille in 1789 was able to construct a new meaning, now being a symbol for a fight *against* oppression, and thus forming a new relation between the geographic place and its significance for a group of human beings.

The actual distinction of places is commonly achieved by using specific geographical data that can be structured either with maps (thus focusing on the geometric extent) or by enriching place names with contextual information, which can be done with gazetteers. Both approaches can be represented with the help of computer-based methods and used for automated analysis. Historical cartographers are employing GIS for their needs while the structure of a gazetteer can be modeled by using relational or graph databases. Of course, the idea of a gazetteer pre-dates the digital age. For an overview on the use and genesis of gazetteers throughout history, see [2].

For historians of the pre-modern eras, it is not only important where the places described in historical sources are geographically located. Rather, it is of much more interest interest in which way they were related to certain groups of people and agents wielding governmental, juridical or religious power or influence over it, as well as how different places were related to each other. Such a contextual approach focusing on what humans attribute to a place rather than on it's geometrical extent offers a more accurate description of what distinguishes a place from another. The name as well as the location of a place is just a designation for a space that is in some way meaningful to someone, while the relation of a place to the cultural and temporal setting in which it has certain properties, make it unique as an entity [3, p. 56][4, p. 138]. Multiple and complex relations can be modelled by using digital gazetteers. A gazetteer uses an implicit structure to arrange and distinguish places by their various properties [5, p. 1042]. In its basic structure, described by Linda Hill, a place in a gazetteer consists of at least one *designation*, one localizing *footprint* (as geo coordinates) and a *type* [6, p. 107].[1] But how does a gazetteer need to be structured if it is to be applied to medieval and early modern places and spaces? What properties have to be used to model our notion of historical reality as closely as possible?

Some challenges related to the development of historical gazetteers have recently been discussed by [4, 5, 9]. However, these works have mostly focused on the 19th, 20th and 21st centuries. Places, set in these epochs, are more commonly distinguished by clear borders due to an increasingly high administrative penetration of space since the 1800s. A standardization of administrative institutions inside the states means fewer ambiguities to consider [10, pp. 99–100, 106].

Other works have taken a broader approach with upper-level ontologies and encompassed a model for human activity in relation to places [11]. Recent projects try to connect existing gazetteers to create an interconnected model of a global (historical) gazetteer. An example for such a project is Pelagios, interconnecting places and documents with a focus on antiquity [12]. However, such a global gazetteer has to be

---

[1]An emerging standard using this specification is *OGC Points of Interest*; see [7] and [8, p. 72].

created as a "federated system", derived from multiple gazetteers, highly specialized on certain demands [13, p. 87].

This paper will take a step back from this global perspective and adopt a domain-oriented focus to the distinct problems related to the representation of places set in medieval and early modern times, with regard to constraints specific to Europe, especially to the Holy Roman Empire. Places from these epochs share a few properties that are not inherent to later periods. The main reason for this is a different kind of society predominant in Europe before the 19th century. Other than modern democratic cultures, medieval and early modern societies where split into multiple subsocieties which existed next to one another or in conflict to each other. This also meant a much more ambivalent construction of meaning attested to the places that were relevant for these societies. A lack of clearly defined administrative units or a general fuzziness of borders between places are only two prominent examples for such ambiguities. In part, these problems are made even more complicated by incomplete or conflicting historical source material. Modern gazetteers on the other hand can more easily disregard such aspects to structure place data.

In Sections 2 and 3, we discuss the general methodological challenges that arise when creating a gazetteer for historical places – like modelling time, (proto-)administrative hierarchies or conceptualization, under special and more detailed consideration of the medieval and early modern situation. In Section 4 we examine how these problems have been addressed by existing gazetteers and ontologies. These are not dedicated explicitly to medieval and early modern places, but may in some cases be adapted for these specific domains. With this survey, we are able to get an overview of the technologies currently in use and the methodological problems that have already been addressed. Finally, we summarize these results in a catalogue of design-practices for the creation of historical gazetteers that take into account the particular properties of medieval and early modern places, and look ahead at necessary future developments.

In terms of technology we promote the usage of Semantic Web technologies for the creation of gazetteers. The standardization of data and the rules on how to organize them, which are intrinsic properties of Semantic Web technologies, enable a high degree of interoperability among different data sources. This allows researchers to use the same model to represent historical places and to interlink their different collections of data within the web of data. The formalized set of rules is developed and published as an *ontology*, a "document or file that formally defines the relations among terms" [14], as a specific conceptualization of a given field of interest. One should note, that such a conceptualization is not universal but represents a specific view on reality [15, p. 84]. Especially an ontology for historical places is therefore always shaped by certain research traditions.

2. Basic concepts to structure historical place and spatial data

To begin with, it is imperative that place data are systematically structured, so that places can be better distinguished within gazetteers. In general, there are two basic approaches for structuring place data: *formal* (relying on geospatial data) and *informal* (relying on the name of the place as a unique identifier) [6, p. 19]. To these existing approaches, we propose to also add *conceptual* and *event-based concepts* as approaches to distinguish place data in a more distinct way - the latter concept is further explained in section 3.5. In existing projects, a combination of these approaches is often used.

Formal gazetteers whose structure relies on geospatial data often build on geographic coordinates, which correspond to the place's position on Earth.[2] Depending on design and complexity of the gazetteer, this position can be modelled in the form of points, lines or polygons. In reality, all places bearing a meaning for human beings occupy a certain area in space. Therefore, a model based on single pairs of coordinates is already a strong abstraction, containing a considerable level of vagueness despite the geographical precision. This point-based approach nevertheless has many advantages compared to the ones based on areas. Medieval and early modern urban areas, for example, consisted not only of its inner region encompassed by a fortification. Inside the wall multiple spheres of influence could exist. Suburban areas outside the walls can also be considered part of a town and were subjected to continuous changes in size over the course of time. Therefore, a gazetteer that represents towns as polygons has to capture many more changes than a gazetteer that represents them as points. However, the

---

[2]See for example ISO 19112, an international standard for the implementation of a gazetteer based on geospatial features; see [8, p. 70].

required level of granularity is rarely covered by historical sources.

A *conceptual* approach distinguishes the places by classifying them using defined categories or place concepts. Mostly a place can be identified by more than one place concept. In this case it is to note, that a semantic gazetteer should always be able to represent dependencies between its objects.

If marking the geographic position of a place is necessary but the use of coordinates is impossible or not desired (e.g. due to lack of accurate data), places can also be distinguished by using Qualitative Spatial Reasoning. This can be implemented by representing the location of places via topological relations (see for instance [16]). The discovery and visualization of these relations can be facilitated by means of applying Semantic Web technologies and known GIS (Geographic Information Systems) standards [17]. In this case, a place is described in its abstract spatial relations to other places. For example, if one wants to determine which towns and villages were within the dominion of the 16th century Prince-Bishopric of Münster, the extension of the territory[3] in kilometers or miles is not relevant. The same applies to questions regarding neighboring territories, or enclaves. It is important that such information is in some way stored as properties of the Prince-Bishopric of Münster. Because most places and territories prior to the 19th century lack clearly defined borders, this approach has many advantages for modelling the fuzziness of historical places. When implementing Qualitative Reasoning an underlying calculus is needed. In the case of spatial reasoning with topological relations, the *Region Connection Calculus (RCC8)* is oftentimes used [18].

3. Challenges

Modelling places and territories in the pre-modern periods poses many conceptual and methodological challenges. In this section we elaborate on these challenges, including specific ontological issues regarding different ways to describe a place as a concept. We also discuss which problems arise when dealing with multiple and changing toponyms and with modelling temporal, territorial and hierarchical structures. Finally, we discuss the issue of capturing data provenance with regard to modelling historical places.

3.1. Toponymy

Modelling the designation of places over time poses a challenge in itself. A town in the contemporary Federal Republic of Germany is called *Münster* by their inhabitants, but bears the official name *Münster (Westf.)*. There are sources from the ninth century which refer to this place as *Mimigernaford* [19, p. 283]. Another source dated to the year 1244 calls it *Monasterium* [19, p. 284]. This simple example shows two problems one encounters when modelling historical places. Firstly, place names change over time. Secondly, even at one point in time multiple designations can refer to the same place, which is already the case when taking into account multiple languages.

To address these problems, a gazetteer must distinguish between the place as such and its designations.[4] The structure of a gazetteer must thus offer the possibility to attribute different names to the same place. Furthermore, it should be possible to assign those names, concrete dates or time spans of usage to those names referring to their the date or time of their use. This approach would not only offer a clear disambiguation of the place data but also allow to recount the history of the toponymy itself, and could, thus, yield data for onomastic studies. It is also possible that sources contain multiple forms of a toponymy as misspellings or real alternatives may be of interest. However, the distinction of closely related alternative place names has to be made by domain experts.

3.2. Place concepts

In order to describe places based on historical and contemporary human-made attestations, categories have to be developed. Especially when dealing with historical places, the development of these categories already implies an interpretation. This interpretation is either based on historical research or, when using concepts directly derived from historical sources, based on a specific worldview, held by the creators of these sources. When focusing on the latter, a gazetteer cannot be merely understood as a collection of places, but as a "cultural gazetteer" [4, p. 141]. In addition, there are two major aspects to consider when creating concepts for categorizing places.

---

[3]In this example territory is grasped as an administrative unit of some sort. The example works as well in a narrower form in which territory is understood for example as a parcel of land.

[4]For a more detailed discussion of technical solutions for these problems, see [20] with regard to the career biographies of researchers.

Firstly, one can distinguish between *fiat objects* and *bona fide objects* (*physical objects*). Fiat objects are virtual spaces conceived by men [4, p. 135]. Mostly they are grounded in some legal or administrative concept, for example a diocese, a republic, or a kingdom. Physical objects are places that can be observed in the actual world, like buildings and towns (here understood as physical entities, not as political concepts) but also natural objects like rivers or trees. Especially in medieval and early modern times, it is easier to attribute exact geodata to these places than to fiat objects.

Although fiat objects and physical objects can be treated differently, it must nevertheless be possible to create dependencies and relations between them. A physical place like a church building can be located inside an administrative unit, which has its own properties. This means, that at least some of the properties concerning the administrative unit must be valid for the church as well. Therefore, the actual place and the meaning of the place attested by humans are distinguished by two concepts. Another example for this could be Vehmic oaks, which served as a court in medieval times. Modelling the tree object and the attested meaning as a place of jurisdiction as two concepts can also make it easier to represent changes to both aspects separately. The function of any specific Vehmic oak ended at some point in time, the tree itself on the other hand can exist much longer and even get another cultural meaning, attributed later in time.

Secondly, it is possible to use either specific/contextualized or general terms for place concepts. Place concepts like *duchy*, *republic*, *prince-bishopric*, or *parish* are specific terms, and their meaning is related to a certain historical context, while general concepts can be understood as concepts of broader categories, encompassing a multitude of place concepts. Examples for general concepts could be *secular dominion*, *theocratic dominion* and *ecclesiastical administrative unit*. Regarding the use of terms for concepts, a semantic gazetteer for historical places should not solely rely on one of these approaches. The assumption that a duchy and a republic can be grasped by only one general concept, like secular dominion, is a very broad historical simplification. Therefore a gazetteer always has to maintain relations between the place concepts it uses, depicting heredital dependencies between its domain oriented terms.

The given examples for specific fiat concepts illustrate an additional challenge related to medieval and early modern places in particular. Not all historic ways of ruling can be grasped the same way as contemporary administrative units. Going further back in time, a clearly defined authority ruling a certain territory is less likely to be found. Instead, administration existed in the form of various privileges, e.g. different forms of jurisdiction or the right to raise taxes. These, in a given territory, could be enforced by different agents – e.g. a prince, the church, a town, or a local nobleman. Sometimes the same or similar privileges could be claimed by multiple ruling actors for the same group of people. Therefore an accurate depiction of an administered space with a multitude of privileges based only on fiat concepts would require the creation of a large and heterogeneous number of such concepts. A more accurate way would be to model ruling as relations between agents, privileges and places. This way, the vast ambiguity of medieval administration, as well as small conflicts of interest, could be captured more effectively. An example for the connection of data about places and people to model rulership is provided by the *Hull Domesday Project*. The structure for the underlying database is mapped from the historical source text it represents, the Domesday book [21, 22].

Place concepts are similar to the *feature type* attribute defined by Linda Hill for a basic gazetteer [6, p. 107]. However we would like to go further than structuring places based on a thesaurus of feature types developed from a certain point of view on (historical) places. When structuring historical places not only the historical context should be included, but also historical point of views about the places. That way a gazetteer about historical places can be a database trying to depict the world as it was presumably perceived by the people of the past (by using the terms and concepts from historical sources) while it is also a tool to structure our way of thinking about history in terms of research traditions. A simple feature type thesaurus is not sufficient for such a task [23, p. 45].

To better grasp these challenges about multiple assertions how a place is regarded, we can again take a look at the history of the city of Münster. During the *Münster rebellion* from 1534 to 1535 the city was under the rule of an anabaptist group. They regarded the city as the center of their new kingdom and therefore detached from the prince-episcopat Münster it was part of. In the eyes of the bishop, the city was still part of his territory. At the same time historians could want to classify the city for the years 1534–1535 as neither, but as *theocratic dominion*. A gazetteer for historical places should be able to include these multiple assertions.

3.3. Hierarchies

An extensive gazetteer should be able to structure not only the places as singular entities, but also model relations between them. We can distinguish semantic and geometric relations between places. In this section we will focus on the semantic relations between fiat places, while the issue of geometric relations has been discussed in section 2. When considering semantic relations a historical gazetteer should especially be able to represent the membership of a place in an administrative structure. Modelling these relations has several advantages. When for instance the name of a town is not known to a user, the place can nevertheless be found by its context information e.g. by its association to a certain political entity. Furthermore, the position of a place in a hierarchy of rulership already delivers a basic understanding for the people associated with a place. Their rights, duties and status as ruled subjects can in part be derived from the legal status of the place they inhabit. This legal status can express itself in its relation to other places.

When modelling places as defined administrative units, one has to decide between distinguishing the levels in a general (*administrative level 0*, *administrative level 1*, *administrative level 2*, ..., *administrative level n*) or in a specific way (*parish*, *diocese*, *bishopric*). However, while creating such a model, one should keep in mind that those apparently evident hierarchies are the product of a long and complex development. It would thus be necessary to keep the model of these hierarchies as flexible as possible.

Firstly, for the medieval and early modern period, we can distinguish between an *ecclesiastical* and a *seigneurial* hierarchy, which may be understood as different layers of society. On the ecclesiastical layer, a place could belong to a diocese, while on the seigneurial layer it is part of a certain county or duchy. Furthermore, there could be different territorial social groupings that can be modeled with additional hierarchies. An example for this would be the districts of certain craftsmanship.[5]

Until this point, we presumed that hierarchies were not in conflict with each other and therefore not contradictory. But places can also be part of two or more *competing* hierarchies. This can be the case when multiple authorities lay claim to a single place.[6] Hierarchies can be a way to model the disputed state of such a place.

To illustrate the problems described above we can take a look at the jurisdictional situation in the city of Münster from the 16th to the 18th century. As a city set in the prince-episcopat of the same name, it fell under the jurisdiction of the diocesan ecclesiastical court (*Offizialatsgericht*), which was part of the ecclesiastical jurisdictional hierarchy [25, p. 265]. At the same time the prince-episcopate employed a seigneurial court (*Hofgericht*) which was responsible for charges not concerning religious matters. On the other hand, the diocesan ecclesiastical court did claim non-religious jurisdiction as well [25, p. 265]. To make matters even more complicated, the (seigneurial) judge, responsible for the city of Münster, was admittedly appointed by the prince-bishop, but the city council could in some way control the administration of justice in the city by employing lay judges [25, p. 262]. One could argue that we have to distinguish between three layers of jurisdictional hierarchy of which the city of Münster was part of: ecclesiastical, seigneurial and the control by the city council. All these hierarchies would in general be in conflict with each other.

3.4. Territories

If territories are modelled in a gazetteer, it has to be asked how continuity and changeability of their borders and coverage are structured in the ontology. Furthermore, processes like merging or splitting of territories need to be modelled, so some relations for stating if there is any continuity between such operations should be created.

The representation of the territory as an area poses another problem. When dealing with territories that are encompassed by clearly defined boundaries, it is possible to store area information as polygons to maps. However, with medieval and early modern territories this is rarely feasible, although there were material (like using boundary stones, *Landwehren* or actual landscape descriptions) as well as symbolic ways (ritualistic statements of belonging, for example by staging processions) to mark the border of a territory [10, pp. 106–107].

Even if this kind of border demarcation is preserved, they are not necessarily an accurate mark of historical

---

[5]Especially itinerant craftsmen, for example tinkers, were organized in tinker-districts. These districts stood in relation to a territorial lord who was responsible for their protection; see [24, pp. 831–932].

[6]A contemporary example for such a place would be the Crimean peninsular.

territories. Borders were often in dispute, so that this status has to be captured as well. These considerations only apply, however, when there was indeed a spatial concept of borders, thus when a dominion was linked to a territory. The polygon-based representation of administrative structures, which is used in modern maps, does not reflect the medieval or early modern situation. In the Middle Ages, 'ruling' did not mean *ruling over space*, but *ruling over people*. Homogeneous domains often did not yet exist in today's form, but only developed over time into modern territorial states, demarcated by their clearly defined borders [10, p. 103]. Over the course of this development, the understanding and formation of territorial spaces were constantly changing [10, p. 100]. In cases when borders existed before modern times, they were an approximation and rather separated different spheres of influences than different territories.

Clear boundaries were more simply established for small spaces in which ruling agents were defined, and where they could unambiguously be marked, e.g. by walls, as it is the case for towns, or part of towns like cathedral immunities [26, p. 10].[7] But the existence of a marked boundary does not imply the exact representation of the historical administrative conditions. Regarding the city of Münster we know that at least since the 9th century the centre of the place was surrounded by the wall. But the city did not end at its walls but was surrounded by a number of singular households outside the walls. These buildings belonged to the place Münster as well. Here, the modelling of hierarchies comes into place, as discussed in section 3.4.

3.5. Temporal disambiguation

Because of the continuous changes of toponyms, place concepts, administrative affiliation or the mere existence of places, a historical gazetteer demands for a model of time. There are numerous possibilities to tackle this issue. Some projects, focusing on contemporary place data, simply use a *historical* or *former* tag for places and concepts that are no more. This is, of course, insufficient when using this kind of data for historical research. In the case of historical gazetteers,

it is important to relate places and concepts to time in a precise and specific manner. With such an approach one can clarify exactly what place in what condition, and in what time frame is referred to. There are much more complex approaches for dealing with spatio-temporal data which go beyond the scope of this project. Still, this section provides a brief overview on the subject.

Using GIS practices as an example, the easiest way to model time is to understand the whole data set as a approximative representation of the world at a certain point in time. This *temporal snapshot* can be created for multiple points in time by copying the whole data set and assigning a new temporal index. Such a *sequence of snapshots* would allow for temporal disambiguation [27, p. 4]. Although easy to implement, this approach does not model relation of continuity between the places – these would have to be created by the user [28, p. 6]. Moreover, it would demand a completeness of data for the different temporal snapshots, which for historical data is hard to achieve due to the fragmentary nature of the historical evidence.

In fact, there are two general ways of conceiving a model for time: the first is based on the concept of timespan in which a statement is *valid*, the second is based on the point in time at which something has changed, in other words it is based on the *event*.

When every statement in the gazetteer is attributed with a valid time, it is possible to distinguish different aspects of a place on a high granularity. Not only the place as a whole, but also individual aspects (e.g. its population or predominant religion) can be distinguished by different time intervals. On the other hand, the *valid time* approach has the disadvantage of not modelling relations of continuity between two time spans, just as in the snapshot approach. There are some solutions to this, most notably the *temporal spaceworm approach*. Here, different intervals relating to the same object are encompassed by another object to state the relations between them [29]. Within a time span on the other hand, too much continuity can be a problem. A time span usually is not tangible in historical data, therefore creating a time span is bound to a certain research tradition in classifying historical events and epochs. The beginning or end of a time span is only seldom reflected precisely in historical sources. For example, there are seldom historical records that state the exact date when a populated place was founded. More commonly, we only have a *terminus post quem* for its existence.

---

[7]One should note, that even such clearly defined borders usually only marked the core of a sphere of influence. The whole sphere of influence often expanded further to a peripheral state. An example for this was he *Bannmeile*. Covering the surrounding area of a town it assured certain economic privileges for the people living inside the city walls; see [24, pp. 675–676].

This problem, as well as the problem of continuity, can also be addressed by using the second approach to model time – an event-based approach. Numerous ways for modelling events exist on a theoretical level. In general the valid-time approach assigns time spans as attributes to places while with the event-based approach event-objects with or without a time index are modeled. These can be associated with one or more places. One possibility to further distinguish the events in use is to build an additional ontology to represent the events in the gazetteer.[8] Naturally, modelling historical events is an additional topic in itself and has its own methodological and technical challenges. One is the distinction between punctual (taking up a single point in time) and durable (taking up a time span) events. For a more detailed discussion, see [27, p. 53]. As every event is taking up a certain amount of time, a conceptual discussion has to take place regarding this distinction. To overcome these issues, events can also be understood as operations executed on objects and thereby changing the state of an object. Worboys suggests the following operations: *creation*, *continuation*, *disappearance*, *reappearance*, *transformation*, *death*, *transmission* (an operation executed by an object on an other object) and *cloning* [27, p. 6]. Such Operations offer a more accessible basis for creating an event-based time model. When conceiving a catalogue of operations it is again important to account for problems of uncertainty and ambiguity. One should note, that the event-based approach differs in some way from the concept of a gazetteer as a collection of enriched place data. Its focus shifts from the contextualization of places to the contextualization of human agency, which is then located in space [23, p. 1092].

A number of these problems results from the use of numerical dates, and arise from all the described approaches: Kauppinen et. al. distinguish three types of fuzziness regarding the representation of time: *uncertainty* (the exact date when a time span began or ended is not known), *subjectivity* (there are multiple diverging research traditions about the boundaries of a time span) and *vagueness* (expressions about a time span are made in reference to non-numerical concepts, e.g. "the event happened around the start of spring") [34, p. 549]. When dealing with populated places in the premodern era, one almost always has to deal with *uncertainty*. The problem can, for example, be approached

by using a known *terminus ante quem* (or *terminus post quem*). Using a day as a *terminus ante quem* is of course an approximation in itself.[9]

One also has to keep in mind that dates in historical sources may make use of different reference systems for dates. Like the Julian and Gregorian calendars, these can exist simultaneously. Although a standardization of these systems may be considered advantageous for comparability, a gazetteer drawing heavily from historical sources should be able to model different reference systems as well to keep the information provided by the sources.[10]

3.6. Provenance of data

Designing an ontology as well as modelling a database for historical places is based on the study of historical sources. It is therefore important to state from which sources place data, concepts and properties are derived, since historical data are rarely unambiguous and undisputed. For historians, it is vital to distinguish between different sources that may present different perspectives on the same event or which have different levels of trustworthiness. Thus, for instance, there is a difference whether data derives from a charter certifying a certain act of law between different parties, or whether your data derives from a chronicle written many years after the event by one of the interested parties, for instance after a dispute about this act of law.

On the other hand, one should note that not only actual places but also mythological places played a role in historical sources. Fictitious places were occasionally even depicted in maps. The most popular example may be the *Ebstorf Map*, but fictional places were also used in maps that were actually used for navigation, like portolan charts [38, p. 106]. It is therefore necessary to make a distinction between fictitious and actual places in an ontology for a historical gazetteer

---

[8]For examples of event-ontologies see [30] and [31]. A description of historical periods is provided by the event-gazetteer *PeriodO*; see [32]. For a more general overview on period gazetteers see [33].

[9]*Topotime*, an extension of the *GeoJSON* format, already addresses these problems. It allows for modelling events and historical periods based on the valid time approach but provides not only a start, and an end date, but also an optional earliest start and end date respectively; see [35, p. 2]. Formally this approach is based on the application of fuzzy set theory; see [34, p. 550].

[10]A few of the concepts discussed in this section are already implemented in the *OWL Time* ontology [36]. Most of the gazetteer projects, discussed in this paper, do not make use of this ontology. However there exists a proposal to integrate the OWL-Time Ontology as well as GeoSPARQL into the CIDOC CRM, taking spatiotemporal queries and qualitative reasoning into account; see [37].

taking into account early modern and medieval times [4, p. 139]. Since fictional places are not subjected to the same rules and as real ones, a semantic gazetteer should also put less constraints on mythological places than on real ones.

Furthermore, in the interest of citability, it might also be desirable to capture which person entered which data. This could become much more important in the future if a contribution to a database by a researcher should be counted as a form of publication [39, 40].

4. Ontological approaches

Existing digital gazetteers mostly try to cover the state of the contemporary world. Some also incorporate historical places, mostly by simply adding a *historical* tag to places and spaces that are no more. This is not sufficient for modelling historical developments, as has been pointed out above.

In the following section, we thus focus on ontologies and projects more aware of the necessity of a historical perspective. Moreover, due to their different perspectives, an overview of these projects also establishes the current state-of-the-art. The following prefixes are going to be used throughout this section:

```
dcterms: http://purl.org/dc/terms/
dnb: http://d-nb.info/gnd/
dnbt: http://d-nb.info/standards/elementset/dnb#
foaf: http://xmlns.com/foaf/0.1/
geo: http://www.opengis.net/ont/geosparql#
gndo: http://d-nb.info/standards/elementset/gnd#
gov: http://gov.genealogy.net/ontology.owl
owl: http://www.w3.org/2002/07/owl#
pleiades: https://pleiades.stoa.org/places/vocab#
pl-author: https://pleiades.stoa.org/author/
pl-places: https://pleiades.stoa.org/places/
pl-time:
    http://pleiades.stoa.org/vocabularies/time-periods/
prov-o: http://www.w3.org/TR/prov-o/#
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs: http://www.w3.org/2000/01/rdf-schema#
skos: http://www.w3.org/2004/02/skos/core
xsd: http://www.w3.org/2001/XMLSchema#
```

The *CIDOC CRM* [41] is a heavy-weight ontology, conceived for managing items collected by museums and tracking their provenance, but open enough to model many other representations needed in the humanities. Initiated by the International Council of Museums in 1996, it can be considered the most extensive ontology in the cultural heritage sector [42, p. 174], becoming an ISO standard in 2006. The version used in this paper is 6.2.2. Implementations of the *CIDOC CRM* can be found at the project homepage [43] or at the Erlangen CRM/OWL project [44].

The *Gemeinsame Normdatei (GND)* [45] (common standardized data) is an authority file for a normed vocabulary used in cataloging literature. It has been developed by the German National Library by merging different standardized data models [46, p. 58]. It is the central authority file for Linked Data in the domain of cultural heritage in German-speaking countries to unambiguously identify elements as contemporary and historical persons, institutions and subject headings. It, thus, also includes contemporary as well as historical places and place names. The data model is structured based on an ontology [47]. Although *GND* and *CIDOC CRM* do not focus on places, we have to take them into considerations because of their status as de facto standards in the cultural heritage sector.

Wikidata [48] is a project hosted by the *Wikimedia Foundation*. It is constructed as a central repository for structured data within *Wikipedia*, from which all other Wikipedia projects can draw data. *Wikidata* does not have a fixed ontology, but is solely build up of so-called *Entities*. An entity can either be an *Item* (a specific object or an abstract class) or a *Property*. All entities consist of a name and a number of properties, each referring to another item.

Besides these three central projects providing an ontology to describe cultural heritage, a normed vocabulary in the shape of an authority file and a repository to collect structured data, we also examine two historical gazetteer initiatives which focus on different time periods.

Pleiades [49] is a project of the *Institute for the Study of the Ancient World* at the New York University. It covers geopolitical as well as geophysical places from the European ancient world. Its focus lies in the positioning of the places in a simple map view, but provides also an ontology to cover name variants or contributors.

The *Genealogisches Ortsverzeichnis (GOV)* [50] (Genealogical Gazetteer), finally, is a Gazetteer for historical places build up through crowd sourcing and hosted by the *Verein für Computergenealogie* (Society for Computer-Based Genealogy) under CC-BY-SA 4.0 license [51, p. 53]. It is aimed at historians and genealogists, covering Europe from the 19th to 21st century. Along with *Pleiades*, the *GOV* will serve as an example for highly specific domain ontologies for historical places.

Of course this list is far from being complete. With examples serving as *de facto* standards, an upper ontology and two domain ontologies, we can take a look at three different types of projects. Some com-

mon projects, like the *ADL Gazetteer Protocol* [6] or the *Getty Thesaurus of Geographic Names* [52] are left out, because they focus on the disambiguation of the places by name while the enrichment of the place data with contextual information plays only a minor role. The ontology provided by *Geonames* [53], which is often used to aquire coordinates, focuses solely on contemporary places. The most extensive gazetteer focusing on historical places, the *World-Historical Gazetteer* [54], is currently under developement. For an overview on ontologies focusing on modelling the administrative structure of places see [55].

In the section, we examine how these projects represent historical places, whether the challenges described above have been addressed by them, and in which way they did so. This is meant to provide conclusions for the modelling of a Historical Gazetteer for medieval and early modern places and spaces.

4.1. CIDOC CRM

The *CIDOC CRM*[11] does not offer a concrete catalogue for historical place concepts in itself. It provides an easy structure to implement such a catalogue with a contextualized approach. To represent a place as defined by it's geometric extent the entity `E53 Place` can be used. From its superclass `E1 CRM Entity` it inherits the property `P2 has type (is type of): E55 Type` to classify the different places to be incorporated in the ontology. An inheritance structure in a concept catalogue could be modelled based on the property `P127 has broader term (has narrower term) : E55 Type of E55 Type`. This would however not be interoperable since the terms used here are not standardized.

The question of the diverging toponyms can be addressed by the property `P87 is identified by (identifies): E44 Place Appellation`. The class `E44 Place Appellation` has four subclasses to not only refer to the name of the place (`E48 Place Name`) but also to its coordinates (`E47 Spatial Coordinates`), its address (`E45 Address`) and or to smaller sections of a given place as for instance a gate or the deck of a ship (`E46 Section Definition`) [41, p. 23]. Describing multiple names can be achieved by assigning multiple Appellation objects to one place. If a gazetteer based on this structure should distinguish place names as less important, the superclass of `Place`

---

[11]For an overview of the classes and properties discussed in this section, please consult Figure 1 on page 12.

`Appellation (E41 Appellation)` has the property `P139 has alternative form: E41 Appellation` to define a secondary appellation.

The problem of topological relations between different territories is tackled in the *CIDOC CRM* with three properties of the class `E53 Place`. These can model enclaves (`P89 falls within (contains): E53 Place`) as well as the fact that two or more territories overlap (`P121 overlaps with: E53 Place`) or border each other (`122 borders with: E53 Place`) – regardless of the actual spatial extension of the places. If this extension should be stored, the place-class also offers the property `P168 place is defined by (defines place): E94 Space Primitive` to assign geodata to an object.

To solve the issues of ambiguity and uncertainty with historical source data, Hiebel et. al. have provided a new approach in their extension of the *CIDOC CRM*, *CRMgeo*. They derive two subclasses from `E53 Place` to represent it: The first, `SP2 Phenomenal Place`, represents the place how and where it is actually located and shaped, regardless of known, unknown or uncertain geodata or any other human declaration, like its depiction on a historical map or a description of the place found in a source (thereby being regarded as a *physical* or *bona fide object*). The second one, `SP6 Declarative Place`, resembles the concept of *fiat object* described in section 3.1. With regard to ambiguity of historical data it can be used as our description of a `SP2 Phenomenal Place`, based on historical source, archeological findings or guessed approximations. Only this place can be assigned geodata (or a *Spacetime Primitive* in terms of the *CIDOC CRM*) [56, pp. 4–6].

This distinction clarifies if one models the actual historical place, or if one talks about a contemporary and possibly flawed representation of it. A fictitious example to illustrate this could be a medieval village that does not exist anymore. Through descriptions in sources about it and archaeological findings we can make educated guesses about its true location and shape. But regardless of how exact our findings may be, we can never be sure that the data we acquire represents all the information of the village as it truly was in the past. To separate our incomplete findings from the historic object, the *CIDOC CRM* offers two concepts to describe a historic place.

Concerning the representation of temporal change and the matter of temporal disambiguation, the *CIDOC CRM* mainly focuses on the event-based approach. There are two different variations. Firstly, the use

of the class `E4 Period`. A Period can be an `E5 Event` which can for example be further specialized by `E63 Beginning of Existence` or `E64 End of Existence`. Through `E4 Period` all these classes get the property `P7 took place at (witnessed): E53 Place`. Therefore, different events concerning a place, for example it's foundation, can be modelled.

By using `E7 Activity` – another subclass of `E5 Event` – pre-modern privileges held on a place by actors can also be modeled. In the *CIDOC CRM* an `E7 Activity` can be carried out (`P14 carried out by (performed)`) by an `E39 Actor`. Modelling privileges as `E7 Activity` would take the ambiguity of medieval and early modern administrative structures and rulership, introduced in section 3.2, into account.

With Version 6.2.2 some features from CRMgeo proposed by Hiebel et. al. [56] have become part of the *CIDOC CRM*. The concept centers around the class `E92 Spacetime volume` which represents a place at a given time period. Through its subclasses `E18 Physical Thing` and `E4 Period` a spacetime volume can either emphasize its physical or its temporal properties. The spatial properties of a spacetime volume are represented through the property `P161 has spatial projection` to `E53 Place` while the temporal information can be modelled with `E52 Time-Span`. With the subclass `E93 Presence of E92 Spacetime volume`, parts or snapshots of space-time-representation of a place can be modeled. With these changes it is possible to model a much more advanced interconnection between places and time.

To describe the provenance of any information modelled with the *CIDOC CRM*, it can be extended with *CRMInf*. This ontology provides a framework for making inferences and argumentation as well as modelling the credibility of statements. However, a full discussion of the *CRMInf* ontology would go beyond this paper. For an extensive discussion of the model – including the underlying concepts – see [57].

One should also note, that the classes, discussed in this section, are only meant to describe real objects. However the *CIDOC CRM* also provides the class `E89 Propositional Object` which can be applied for modelling mythological places as well. It's property `P67 refers to (is referred to by)` allows such an object to be linked to any other entity of the *CIDOC CRM*, which can be used to make any statements about a fictional place.[58].

4.2. Gemeinsame Normdatei (GND)

For the *Gemeinsame Normdatei (GND)*[12] the central class for places is `gndo:PlaceOrGeographicName`. It allows for a geographical, conceptual as well as a hierarchical disambiguation of places. In general, smaller and specific places like towns seem to be understood geographically, while bigger and more abstract places are categorized hierarchically. The geographical categorization is allowed by usage of the property `geo:asWKT` from the *GeoSPARQL* vocabulary[13] to assign georeferencing data to it, as is shown in Listing 1.[14]

Listing 1: An extract about the town of Münster (Westf)

```
_:ms geo:asWKT "POINT(7.62 51.96)"^^geo:
    wktLiteral ; a sf:Point .
```

A conceptual disambiguation of places can be done in different ways. Firstly, there are a few subclasses for `gndo:PlaceOrGeographicName`, with no formal distinction between fiat objects and physical objects.[15] By associating a place with one or more of these subclasses, a conceptual categorization is possible. Listing 2 shows the turtle representation of the Prince-Bishopric of Münster.[16]

Listing 2: An extract about the Prince-Bishopric of Münster

```
dnb:4279480-8 a gndo:ReligiousTerritory ,
    gndo:AdministrativeUnit , gndo:
    TerritorialCorporateBodyOrAdministrative-
    Unit ;
    gndo:broaderTermInstantial dnb:4160263-8
    ;
```

---

[12]For an overview of the classes and properties discussed in this section, please consult Figure 2 on page 13.

[13]*GeoSPARQL* is a standard of the *Open Geospatial Consortium* which is developing standards and technologies for processing geo data; see [59].

[14]http://d-nb.info/gnd/4040608-8

[15]The subclasses are called: `gndo:BuildingOrMemorial`, `gndo:ExtraterrestrialTerritory`, `gndo:FictivePlace`, `gndo:TerritorialCorporate-BodyOrAdministrativeUnit`, `gndo:MemberState`, `gndo:NameOfSmallGeographicUnitLyingWithin-AnotherGeographicUnit`, `gndo:Country`, `gndo:NaturalGeographicUnit`, `gndo:ReligiousTerritory`, `gndo:AdministrativeUnit` and `gndo:WayBorderOrLine`.
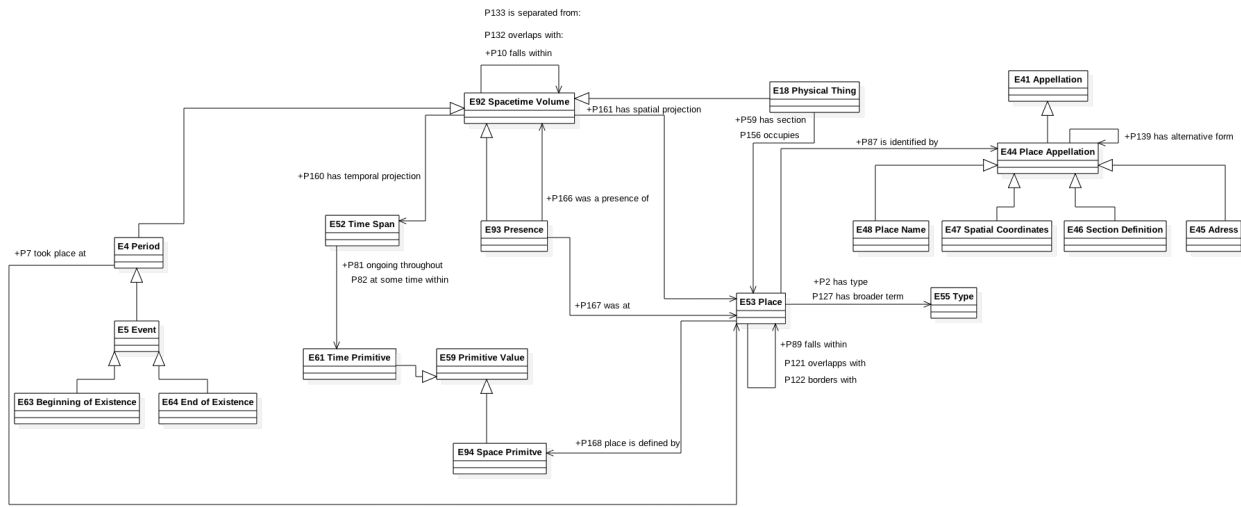
[16]http://d-nb.info/gnd/4279480-8

Fig. 1. UML representation of the *CIDOC CRM* classes discussed in Section 4.1

```
3    gndo:geographicAreaCode <http://d-nb.info
         /standards/vocab/gnd/geographic-area-
         code#XA-DXDE> , <http://d-nb.info/
         standards/vocab/gnd/geographic-area-
         code#XA-DE-NW> , <http://d-nb.info/
         standards/vocab/gnd/geographic-area-
         code#XA-DE-NI> ;
4    gndo:
         variantNameForThePlaceOrGeographicName
         "Fürstbistum Münster (Westf)" , "Mü
         nster (Westf) (Hochstift)" , "Hoch-
         und Niederstift Münster" , "Haut et
         Bas Eveche de Munster" , "Münster (
         Westf) (Staat)" ;
5        gndo:preferredNameForThePlaceOrGeo-
         graphicName "Hochstift Münster (
         Westf)" ;
6    gndo:relatedTerm dnb:4040610-6 .
```

As shown in line 1, the Prince-Bishopric of Münster can be understood as a `gndo:AdministrativeUnit` as well as a `gndo:ReligiousTerritory`. By using multiple inheritance, both the ecclesiastical and the secular aspect of dominion are captured for the territory.

The number of the subclasses provided by the *GND* ontology is rather limited. Therefore, objects can be further enriched with context by using the property `gndo:broaderTerm` or `gndo:broaderTermInstantial`, which means describing concepts not as part of a formalized ontology but as arbitrary descriptions. This can be seen in line 2 of listing 2. In this case, the referred URI leads to the concept *Hochstift* (Prince-Bishopric) [60]. All subjects can be further specified by adding an information text via gndo:definition and related terms via `gndo:relatedTerm`. In this way, a basic human-readable conceptual disambiguation is possible. An automatic processing of concepts attributed this way on the other hand is not given.

The *GND* ontology allows for the modelling of multiple names for one place, may they be contemporary or historical. However, the model does not distinguish, for instance, between different languages. One of the alternatives is marked with `gndo:preferredNameForThePlaceOrGeographicName` as the name by which the place is mainly addressed. Other variations are stored by using the property `gndo:variantNameForThePlaceOrGeographicName` as can be seen in listing 2 in line 4. There is no model for a temporal disambiguation of the names.

It is not possible to model territories as polygons with the *GND* ontology. A territory is defined as such by its conceptual and hierarchical structure. Coordinates can only be used to model the most northern, southern, western and eastern point of a territory. If a border itself is of interest for a gazetteer it is grasped as a place and can be modelled with the class `gndo:WayBorderOrLine`. This could for example be the case when one wants to model events taking place in context to the border.

For temporal disambiguation, the *GND* ontology uses the *timespan* approach. The dates of creation and termination can be modelled with the properties `gndo:dateOfEstablishment`, `gndo:dateOfTermination` and `gndo:dateOfEstablishmentAndTermination`. All three can be linked with either a literal describing the tem-
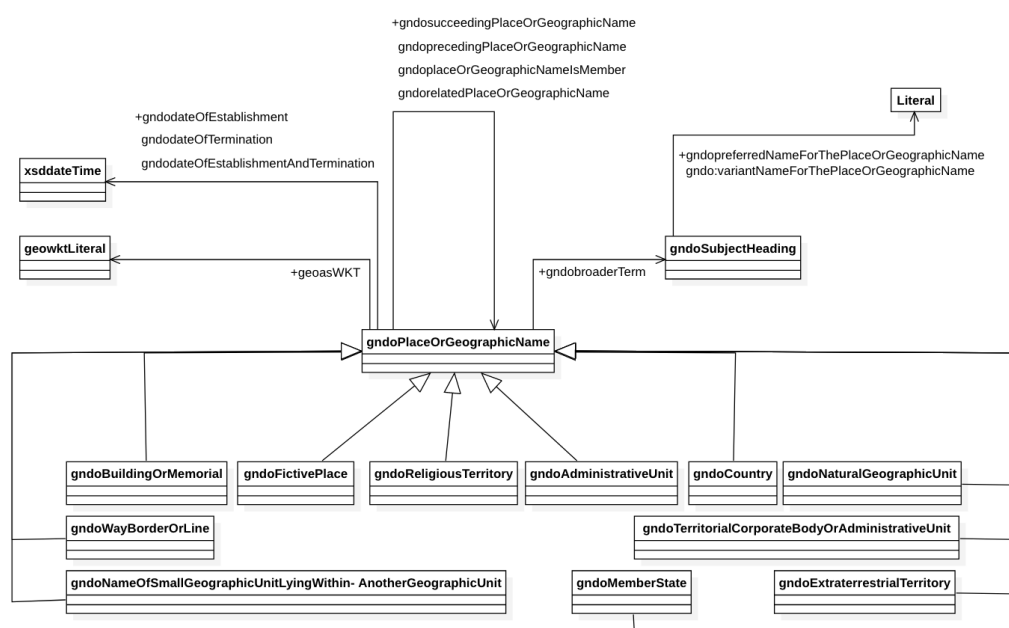
Fig. 2. UML representation of the *GND* classes discussed in Section 4.2

poral information or with another node of the type `xsd:dateTime`. To capture information of continuity, e.g. when places are merged or split, the properties `gndo:succeedingPlaceOrGeographicName` and `gndo:precedingPlaceOrGeographicName` are provided.

The hierarchy model of the *GND* ontology does not rely on a predefined hierarchical structure for its place data, but on historically contextualized relations between objects. Places can be linked by the properties `gndo:placeOrGeographicNameIsMember` and `gndo:relatedPlaceOrGeographicName`. The only other way to arrange places into an administrative hierarchy is the usage of `gndo:geographicAreaCode`. Therefore, it is not possible to distinguish multiple hierarchy-trees and arrange the places in them. For the Prince-Bishopric of Münster, this can be seen in listing 2 in line 3. The listed area codes stand for *Deutsches Reich*, *Nordrhein-Westfalen* and *Niedersachsen*. The first area code refers to a historical membership, the others are allowing for a simple spatial placing of the object, referring to the Federal states the Prince-Bishopric's former territory is currently part of. A temporal disambiguation regarding the membership of elements in hierarchies is not possible with this ontology.

Since the *GND* ontology is mainly designed to standardize terms for cataloging literature, it does not provide any means to capture provenance of the information concerning historical place data.

4.3. Wikidata

In *Wikidata*, it is possible to describe general as well as specific objects and concepts with entities. As with the *GND* ontology it is possible to conceptualize a place by relating it to a number of concepts and using multiple inheritance. Using the Prince-Bishopric of Münster (with the identifier `Q697254`) as an example again, we can illustrate the conceptual inheritance.[17] The place is understood as a concrete instance of (`instance of (P31)`) the concept *Prince-Bishopric (Q24298790)*. This (general) class is a subclass of the concept *Ecclesiastical principality (Q3403564)* which again is a specialization of *spiritual territory (Q1499065)* as well as *principality (Q208500)*.[18] Both concepts are differentiated further. This shows how the different aspects of a dominion in the Holy Roman Empire can be modelled with even higher granularity than with the *GND* ontology. This

---

[17]For a simple UML representation of the example see Figure 3 on page 14.

[18]A principality is defined by *Wikidata* as a "monarchical feudatory or a sovereign state, ruled or reigned over by a monarch with the title of a prince".

has the huge advantage, that these relations afterwards can be included in machine reasoning.

Like the *GND* ontology, *Wikidata* can handle multiple name variants with one marked as the preferred name. Aside from that there is a property called `Former place names (Q7458920)`. Because all values of properties can be attested qualifiers, the temporal disambiguation of historical names with the *valid time* approach is possible. However, in practice it seems to be more common for the *Wikidata* community not to distinguish between name variants and historical names. Both forms are instead stored as aliases to the preferred name in each language version. The alias field of *Wikidata* does neither provide the means for a temporal disambiguation nor is there a tag to mark a name as *historical*.

The primary intent of *Wikidata* is to provide structured information for the different *Wikipedia* projects. Therefore, it does not provide any means to store polygonal data. Territories can be depicted as SVG-images, but they can of course serve mostly for a visual representation and not an automatic analysis. Again like in the *GND* ontology, coordinates of the outermost points of a territory can be stored to approximate the extension of a place. On the other hand, a conceptualization with the classes `shares border with (P47)`, `enclave (Q171441)` or `exclave (Q933394)` can be used to model basic topological relations between different territories.

Temporal data is also modeled in part with topological relations. The respective properties are called `inception (P57)` and `dissolved or abolished (P576)`. To model continuity, `Predecessor (Q16881385)`, `replaced by (P1366)` or `followed by (P156)` can be used. A further development of the concepts used in *Wikidata* are the *Qualifiers*. These are values that can be assigned additionally to each other value of a property. A qualifier can either be an integer or an interval [61]. With this feature it would be possible to distinguish temporal information of each aspect of a place.

Because of the wide scope of *Wikidata*, its properties for hierarchical arrangement of places are not contextualized for historical territories but rather for the use with general concepts. It is possible to distinguish between the membership of a place in a country, in an administrative and in a jurisdictional district.[19] Be-

sides that, special roles of a place can be modeled, for example being the capital of a territory by using the properties `capital (P36)` or `capital of (P1376)` for the reverse relation. Multiple claims on a place can be modelled with `territory claimed by (P1336)`. Because *Wikidata* allows multiple inheritance as well for the fact that an object can of course be linked to multiple other objects, it is possible to arrange a place in multiple hierarchy trees.

Wikidata offers a way to capture the provenance with its properties `stated in (P248)` or `reference URL (P854)`. The first one relates to another *Wikidata* class or object while the second relates directly to another web resource.
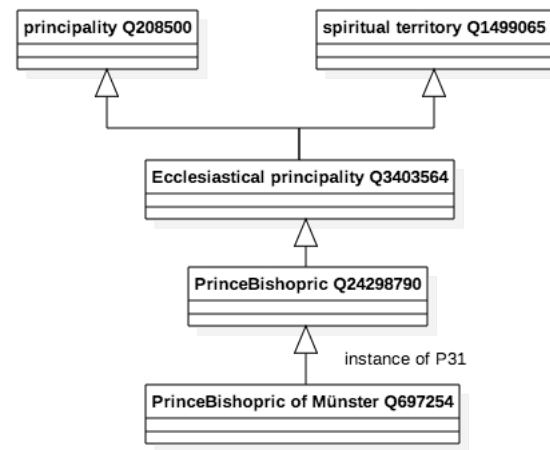

Fig. 3. UML representation of the Prince-Bishopric of Münster in *Wikidata*

4.4. Pleiades

The *Pleiades* ontology has been conceived to model places and spaces in a given historical time, in this case the Antiquity. Its purpose is to distinguish name variants taken from historical sources, and capture the time span in which they were used.

It thus uses specific place concepts that are embedded in a historical context. Although some of the concepts are more general than others, they are not related to each other in a structure of inheritance. Most concepts are defined in a documentation [62]. Because the project's definition is based on Yi-Fu Tuan's

---

[19]The names of the corresponding properties are `country (P17)`, `located in the administrative territorial entity (P131)` and `applies to territorial jurisdiction (P1001)`.

experience-based approach to places [1], the list consists not only of man-made objects like *bridges*, *forums* or *villas*, but also of natural objects like *forests* and *rivers*. Even fictional places are covered in *Pleiades* although there do not seem to exist specific concepts for them. Such places are marked as erroneous place as can be seen at the example of the mythical town Kikynethos.[20] For uncertain cases, there exist the Feature Types `unknown` and `undefined`. Some of the concepts are taken from the ontology of the *Getty Thesaurus of Geographic Names* to ensure some interoperability. One should also note that *Pleiades* does not model fiat places. All places – including the erroneous ones – are presumed to have a geometric location [33].

The *Pleiades* data model[21] provides a much more sophisticated handling of historical place names. At least one name in its original language has to be assigned to every place entity. To prevent encoding problems, the names are given in their romanized form; see [63]. For the city Istanbul, *Pleiades* lists 16 different names, such as Kōnstantinoupolis, Byzantion, Qustantiniya, etc., each with their own temporal attestation, by means of using the properties `pleiades:start_date` and `pleiades:end_date`. One or more epochs, in which the name was used, can be assigned by using `pleiades:during`. Listing 3 shows an extract from the RDF representation of Istanbul[22] focusing on the name Byzantion.

Listing 3: An extract about Istanbul

```
1  <https://pleiades.stoa.org/places/77402394/
       byzantion>
2    a pleiades:Name ;
3    dcterms:description "A place name from the
         TAVO Index (Vol. 1, p. 309)" ;
4    pleiades:during pl-time:early-iron-age-
         anatolia, pl-time:hellenistic-middle-
         east, pl-time:roman-middle-east;
5    pleiades:start_date -1200 ;
6    dcterms:modified "2013-09-09T15
         :40:34-04:00" ;
7    dcterms:title "Byzantion" ;
8    prov-o:wasDerivedFrom [ rdfs:label "TAVO
         Index" ] ;
9    pleiades:nameRomanized "Byzantion" ;
10   dcterms:creator pl-author:deblauwe ;
11   pleiades:end_date 640 ;
```

---

[20]https://pleiades.stoa.org/places/544356
[21]For a simplified UML representation of the data model see Figure 4.
[22]https://pleiades.stoa.org/places/77402394/rdf

```
12  owl:sameAs <https://pleiades.stoa.org/
        places/77402394/byzantion> ;
13  dcterms:contributor pl-author:fdeblauwe, pl
        -author:ekansa .
```

As one can see the properties `pleiades:start_date` (line 5) and `pleiades:end_date` (line 11) provide a valid time frame for the use of a certain name. One or more epochs in which the name was used can be assigned by using `pleiades:during` (line 4 in Listing 3).

Furthermore, the *Pleiades* ontology provides properties to model the accuracy as well as the completeness of toponyms that are taken from historical sources [64]. For both properties there exist three levels from `false/fragmentary` and `not reconstructable` to `accurate/complete`. Names can thus be used from a source that, for example, is known to use a wrong spelling of a name.

The *Pleiades* ontology does not provide a model to describe territories.

Concerning the time model, the properties `pleiades:start_date` and `pleiades:end_date` have already been introduced and can be seen in Listing 3 in lines 5 and 11. Statements in this ontology can not only be valid in time spans marked by dates, but also be assigned to an epoch or *time period* with `pleiades:during`. A time period also has an upper and a lower boundary in the form of timestamps. In terms of formal design, this feature bears no difference to the *valid time model* although it has the advantage that timespans are not arbitrary but contextualized with historiographic information.[23] This also prevents redundancies because the concept of an epoch has to be defined only once.

Since the *Pleiades* project focuses rather on the names of the places than on their underlying political structure, there is no elaborated hierarchical conceptualization. Basic hierarchies, like the membership of a city to a region, are simply referenced by using `rdfs:seeAlso`.

Finally, since the *Pleiades* project is based on the mentions of place names in literature and other source material, it also features elements to describe the provenance of Data. With the `prov:wasDerivedFrom` property taken from the *Provenance Interchange Ontology (PROV-O)*, a place object or a name can be referenced by literature or source material, as can be seen in listing 3, line 8.

---

[23]For a complete list of the predefined time periods see [65].
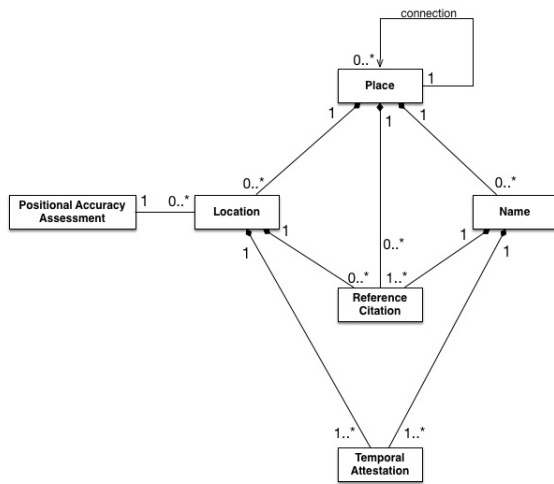
Fig. 4. UML representation of the *Pleiades* data model[24]

4.5. Genealogisches Ortsverzeichnis (GOV)

Whereas the *Pleiades* projects is dedicated to places and spaces in Antiquity, the *GOV* refers to a more recent epoch, from the 19th to the 21st century, with a particular focus on Europe. In 2014, the *GOV* contained approximately one million entries.[25]

The ontology of the *Genealogisches Ortsverzeichnis (GOV)* distinguishes between concepts for fiat objects and physical objects. Physical objects can be different buildings, towns, abandoned villages, churches or market areas while fiat concepts represent man-made virtual units, like secular, ecclesiastical and juridical administrative units. The general concepts are already contextualized. They model specific historical as well as contemporary ecclesiastical and secular administrative hierarchies.[26] There are no general concepts from which they inherit. In this case, different aspects of a special form of dominion are represented in the concept itself.[27] This highly contextualized approach in the *GOV* data model can be explained by the regional and temporal focus of the *GOV*.

The *GOV* ontology also allows multiple designations for a place. The RDF representation of the vil-

lage *Roztoka*[28] provides a good example to illustrate the model of toponyms in the *GOV* ontology:

Listing 4: An extract about Roztoka

```
<http://gov.genealogy.net/ROHOCKJO80CX>
  a <http://gov.genealogy.net/ontology.owl#
      GovObject> ;
  gov:hasName [
    a gov:PropertyName ;
    gov:timeEnd "1945"^^xsd:string ;
    gov:language "deu"^^xsd:string ;
    gov:value "Rohnstock"^^xsd:string
  ], [
    a gov:PropertyName ;
    gov:timeBegin "1945"^^xsd:string ;
    gov:language "pol"^^xsd:string ;
    gov:value "Roztoka"^^xsd:string
  ] ;
  rdfs:isDefinedBy "http://gov.genealogy.net/
      ROHOCKJO80CX/about.rdf" .
```

Listing 4 shows how two name changes of the village have been modeled in lines 3 and 9 by using `gov:hasName`. Besides the name of the whole object assigned by `gov:value`, every name gets a valid time assigned by `gov:timeBegin` and `gov:timeEnd`.[29] With the `gov:language` property it is possible to distinguish further name variants by using a language tag. The example above does not state since when the village Rohnstock exists. In general, the ontology provides no means to state if a date marking the conception or termination of a place is simply not yet given by a user or if it is unknown.

The *GOV* ontology completely lacks a data model for territories. All places that can be described as a territory are distinguished as such by their arrangement in a hierarchical structure alone.

With the `gov:timeBegin` and `gov:timeEnd` properties, the genealogical gazetteer only supports a simple *valid time* approach.[30] Time spans can be assigned not only to a name but any property of a *GOV* place. Therefore, this ontology allows to describe different aspects of an object in different frames of temporal reference.

The hierarchy levels used are specific and historically contextualized. On a conceptual level, there is a distinction between an administrative, an ecclesias-

---

[24]https://pleiades.stoa.org/help/pleiades-data-model
[25]http://gov.genealogy.net/search/index
[26]For a complete list of the concepts see [66].
[27]For Prince-Bishopric territories for example exists the concept *Hochstift*.

---

[28]http://gov.genealogy.net/semanticWeb/about/ROHOCKJO80CX
[29]See lines 5 and 10 in listing 4.
[30]See lines 5 and 10 in listing 4 for an example.

tical, and a jurisdictional hierarchy tree. Conflicting memberships at one of these trees can be resolved by the use of timestamps as is shown in listing 5, depicting the memberships of the *Freistaat Preußen*[31] over time.

Listing 5: An extract about Freistaat Preußen

```
1  <http://gov.genealogy.net/object_190317>
2    a <http://gov.genealogy.net/ontology.owl#
        GovObject> ;
3      gov:isPartOf [
4          a gov:Relation ;
5          gov:timeEnd "1866"^^xsd:
               string ;
6          gov:timeBegin "1815"^^xsd:
               string ;
7          gov:ref <http://gov.genealogy
               .net/object_218127>
8        ], [
9          a gov:Relation ;
10         gov:timeEnd "1870"^^xsd:
               string ;
11         gov:timeBegin "1867"^^xsd:
               string ;
12         gov:ref <http://gov.genealogy
               .net/object_264172>
13       ], [
14         a gov:Relation ;
15         gov:timeEnd "1945"^^xsd:
               string ;
16         gov:timeBegin "1871"^^xsd:
               string ;
17         gov:ref <http://gov.genealogy
               .net/object_190315>
18       ] .
```

For a human user, it is possible to see at one glance the hierarchies a place was part of at a certain time. To achieve this, the whole hierarchy tree for an object is traversed and then visualized when the page of an object is requested. The whole tree for the town of Münster is shown in Figure 5. Note that the nodes are named with historical terms to achieve a better distinction. The *GOV* ontology uses a similar concept as *Pleiades* for capturing provenance of property values. Here, the property `gov:sourceRef` can link an information to another resource as shown in the example in listing 6[33], line 7 describing the population of Prussia in the year 1871. The example links to a study. However, the *GOV* ontology does not distinguish be-

---

[31]http://gov.genealogy.net/item/show/object190317.
[32]http://gov.genealogy.net/item/show/MUNTERJO31TX
[33]http://gov.genealogy.net/item/show/object190317

tween historical sources, web resources and academic research.

Listing 6: An extract about Freistaat Preußen

```
1  <http://gov.genealogy.net/object_190317>
2    a <http://gov.genealogy.net/ontology.owl#
        GovObject> ;
3      gov:hasPopulation [
4          a gov:PropertyForObject ;
5          gov:source [
6            a gov:SourceReference ;
7            gov:sourceRef <http://gov.
                 genealogy.net/
                 source_190312>
8          ] ;
9          gov:timeEnd "1871"^^xsd:
               string ;
10         gov:timeBegin "1871"^^xsd:
               string ;
11         gov:value "24639706"^^xsd:
               string
12       ] .
```

4.6. State-of-the-art gazetteers and medieval and early modern places

Even if none of the projects discussed above is explicitly designed to take the specific problems of places in medieval and early modern times into account, a number of the challenges stated in Section 3 could already be considered resolved by them. Other issues yet remain untouched.

Challenges concerning categorization of place have been addressed insufficiently. The distinction between fiat places and physical places as well as general and specific concepts is only done by the *CIDOC CRM* as well as the *GOV* ontology. Since the *GOV* focuses on the 19th to 21st century, the ontology would have to be extended for usage for early modern and medieval places. Because the hierarchy levels are also places and, therefore, conceptualized, comparison between new levels and already existing *GOV* levels would still be possible. Specific concepts using terms that apply to pre-modern times can only be found at *Wikidata* where they are contextualized through integration in an extensive class hierarchy.

When modelling territories, a geographical approach is used by the *GND* ontology and *Wikidata* through approximating the extent of the area with single pairs of coordinates. Topological relations are provided by *Wikidata* and the *CIDOC CRM*.
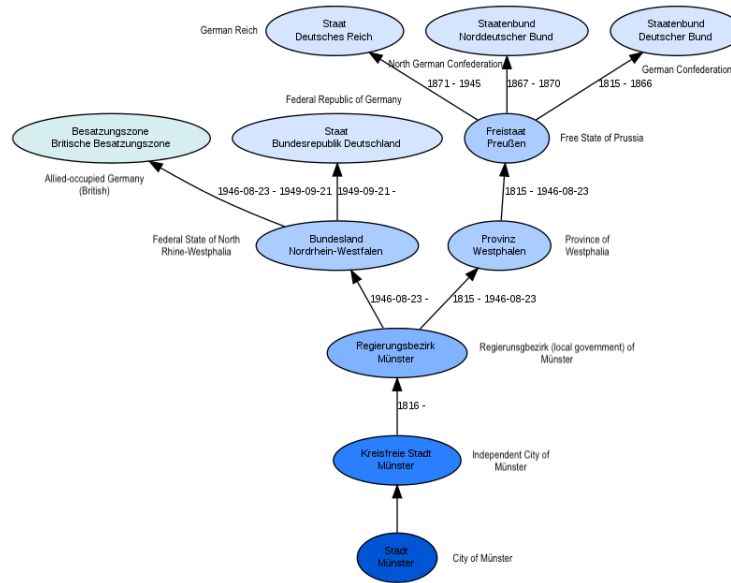
Fig. 5. The town of Münster as an example for administrative hierarchies in the genealogical gazetteer.[32]

All ontologies introduced in this paper solve the problem of multiple names. However, only *Pleiades* takes into account that names retrieved from historical sources can be flawed data and should therefore be modeled as such. To model such cases with the *CIDOC CRM*, the *CRMInf* extension can be used, as stated in section 4.1. As an authority file for names, the *GND* ontology does not allow to state when a variant of a certain name was being used. All other projects separate the name object(s) from the place object and are therefore able to describe such contextual information.

Problems of temporal disambiguation mostly remain unsolved. Except for the *CIDOC CRM*, which suggests an event-based approach, the valid time model is solely used for describing change over time.

By separating the historical source from the editor of a data set, only the *Pleiades* ontology and (with an extension of the *CRMInf* model) the *CIDOC CRM* allow in parts a model of provenance considering the needs of academic research. Different types of sources are not described by any of the data models above. The focus yet lies on web resources. Trustworthiness and completeness of sources is also ignored by others than the *CIDOC CRM* and the *Pleiades* ontology.

5. Building a gazetteer for the medieval and early modern era

5.1. Place concepts

Finally, we summarize which of the different approaches discussed above are required for a gazetteer that covers medieval and early modern places.

With regard to place concepts, the most important question is the decision between using general or specific categories. General concepts allow for more interoperability and comparability, while specific concepts can be more historically accurate and allow for deeper levels of historical analysis. We propose the use of both in combination with an inheritance structure, so that the specific concepts are specializations of the general ones. The use of multiple inheritance guarantees the specification of more complex and ambiguous historical concepts. An enrichment with context information – for example through the property `gndo:broaderTerm` – should be avoided.

As shown in the *GOV* ontology, a distinction between physical objects and fiat objects is also necessary, which makes it possible to attribute different historical or historiographic classifications to otherwise clearly defined objects. An example in use for this could be the distinction between the category of a town and an ideal type that can be assigned to the same place (e.g. *fortress town*); see [67]. Furthermore, it is important to stress the need for modelling fictitious places. In

Table 1

Challenges addressed by the ontologies discussed in this paper.

	CIDOC CRM	GND	Wikidata	Pleiades	GOV
Models multiple place names	✓	✓	✓	✓	✓
Temporal Disambiguation of names	✓	✗	✗	✓	✓
Place name data can be modelled as flawed	(✓) Possible with *CRMInf*	✗	✗	✓	✗
Distincts between fiat and physical places	✓	✗	✗	✗	✓
Distincts between general and specific concepts	✓	✓	✗	✗	(✓) does currently not include pre-modern places
Considers that place concepts are based on a certain (historic or academic) worldview	(✓) possible with *CRMInf*	✗	✗	✗	✗
Possible to implement administrative hierarchies	✓	(✗) no hierarchy model but relations between places are possible	✓	✗	(✓) currently no support of pre-modern places
Implements competing hierarchies	(✓) possible with *CRMInf*	✓	✓	✗	✗
Models territories by polygons	✓	✓	(✓) uses SVG	✗	✗
Models territories by topological relations	✓	✗	✓	✗	✗
Valid time for temporal disambiguation	✓	✓	✓	✓	✓
Models continuity between time slices	✓	✓	✓	✗	✗
Event-based temporal disambiguation	✓	✗	✗	✗	✗
Temporal disambiguation of place attributes possible	✓	✗	✓	✗	✓
Models temporal uncertainties	✓	✗	✗	(✓) places can be assigned to a *time period* instead of dates	✗
Possible to model provenance	(✓) possible with *CRMInf*	✗	✓	✓	✓
Models trustworthiness of provenance	(✓) possible with *CRMInf*	✗	✗	✓	✗
Models fictional places	✓	✓	✓	✓	✗

the examined ontologies, this was mostly disregarded due to the temporal scope of the projects (with *CIDOC CRM* and *Pleiades* as an exception).

5.2. Toponymy

The importance of modelling multiple names for each place has been demonstrated. The most common and efficient practice to resolve this issue is to distinguish between the places and its names. The names can be single (place name) objects themselves, and therefore be provided with their own temporal attribution, shown in the *GOV* and *Pleiades*. As with other properties, it is important that the provenance of each name of a place can be modelled separately as shown by the *GOV* ontology in order to ensure the traceability of the used information as necessary basis of scientific work.

5.3. Territories

To be useful for qualitative reasoning, territories can be modelled by their spatial (like topological) relations in addition to (or instead of) their geometric representation. This enables to enrich the models by inferring new inherent relations, for instance relations between current and historical places (see [68, 69]). While not offering a representation accurately e.g. by maps (as with coordinates), topological relations also better meet the requirements for representing places that are lacking clear defined boundaries. *Wikidata* as well as gazetteers based on the introduced properties of the *CIDOC CRM*, are in general capable of this feature.

If it becomes necessary to visualize the territory on a map, we propose that its extent will be approximated by choosing places whose positioning is less uncertain (like towns and villages) that belong to the respective territory. This provides a much more accurate depiction of medieval and early modern realities, in which authority was defined by power over towns, villages, farmsteads and rights to use forests or stretches of water.

In some cases, it can be useful to include the use of geometries, for example if a gazetteer serves as an object catalog for a GIS application. If territories are also capturing administrative authority in terms of dominion over spaces, it has to be considered to use a model for representing different degrees of administrative penetration. In this case, the overlap representation should also be possible.

The problem remains that most of the approaches looked at have been developed for a world mostly accessed by homogeneous administrative structures. The distinctive features of medieval and early modern dominion, which have been stated in Sections 2 and 3, are ignored by current gazetteer projects when it comes to model ruling structures on the level of towns or villages. To capture these features in a gazetteer for the medieval and early modern world, a new model that understands ruling as an interconnection between places, non-governmental institutions and people has yet to be developed. Such a model of ruling would have to be developed from scratch because of its high dependency from history as a knowledge domain. Nevertheless it should be applicable as an extension to every gazetteer that tries to take pre-modern political realities into account.

5.4. Temporal disambiguation

From the different approaches to model time introduced in Section 3.5, in practice, the *valid time* approach is used by most of the existing ontologies. Out of those approaches the *CIDOC CRM* allows a very flexible model of time. The complexity of the *valid time* model, however, is insufficient to model historical places because it does not capture, for instance, if and how two periods are connected to each other. When changes happen, marked by the transition from one valid time-interval to the next, it is unclear whether or not there is any linkage between the two time periods.

Furthermore, the valid time-interval in itself claims a continuity which is not covered by the sources. In most of the cases historical texts report events. In this case, the continuity is already an interpretation of the data from today's perspective. Only an event-based approach can account for the problem of modelling continuity between the transition from one time span to another, and the creation of time spans at all. Since historical texts in general inform us about events that resulted in changes to the world rather than predefined time spans in which there was no change, event-based models bear a closer proximity to the source material. Therefore, the aspect of temporal disambiguation has to take a more prominent place in the design and usage of ontologies for historical data.

Finally, all aspects of a place have to be distinguished by temporal properties as shown in the *GOV* ontology. This includes all attributes of a place as well

as its relation to other (place) entities, e.g. its relations in a hierarchy.

5.5. Hierarchies

The need for modelling hierarchies depends on the geographical, temporal, and conceptual scope of an ontology for pre-modern places. The *GOV* ontology shows how a model based on multiple hierarchy trees can work. This also allows for conflicting claims as well as different administrative structures to be modelled.

5.6. Provenance of data

Since the data in a gazetteer for historical places heavily relies on information drawn from historical sources, it is imperative that the provenance of all statements in such a database is made visible. Ideally, the data model used for capturing provenance is able to distinguish different concepts of historical sources. As shown with the *Pleiades* ontology, it is to be preferred to also represent the trustworthiness of a source. That way, historical data stored in a digital gazetteer becomes much more reliable and encourages for a wider use.

6. Conclusions

There is not only a trend in research concerning gazetteers to develop ontologies that are able to model places from different temporal, cultural and geopolitical reference frames on a general level, but also a need for this development in order to ensure comparability and interoperability. At the same time some design problems have to be approached not from a general but from a domain oriented perspective. In this paper, we have shown this, by integrating the domain based perspective of historical humanities with the broader approach, developed in geographical and computer sciences. We focused on the challenges arising when modelling historical places which are set in medieval and early modern times. We examined the nature of such places from different perspectives: the naming of places (toponymy), their categorization (place concepts), their relation to other places (hierarchies), their spatial extension (territories), their change over time (temporal disambiguation) and their validity in research as well as in historical sources (provenance of data). We argued, that each of these aspects comes

with it's own specific design challenges. Furthermore, especially the challenges regarding place concepts and temporal disambiguation have shown how strongly a certain historical thinking and understanding of the past can shape particular aspects of an ontology. This observation shows again, how important a domain oriented design perspective is.

We have then surveyed a number of ontologies from existing projects (none of them specializing on the medieval and/or early modern time period) and discussed how they are approaching the challenges identified by us. Thereby we have shown which existing technological solutions can meet our demands and which can not.

Derived from this discussion as well as from the challenges identified before, we have developed a catalog of design practices for the creation of a domain specific semantic gazetteer covering medieval and early modern places.

References

[1] Y.-F. Tuan, Place: An Experiental Perspektive, *Geographical Review* **65**(2) (1975), 151–165.

[2] R. Mostern and H. Southall, Gazetteers Past: Placing Names from Antiquity to the Internet, in: *Placing Names. Enriching and Integrating Gazetteers*, M.L. Berman, R. Mostern and H. Southall, eds, Bloomington / Indianapolis, 2016, pp. 15–25.

[3] R. Shaw, Gazetteers Enriched: A Conceptual Basis for Linking Gazetteers with Other Kinds of Information, in: *Placing Names. Enriching and Integrating Gazetteers*, M.L. Berman, R. Mostern and H. Southall, eds, Bloomington / Indianapolis, 2016, pp. 51–63.

[4] H. Southall, R. Mostern and M. Lex Berman, On Historical Gazetteers, *International Journal of Humanities and Arts Computing* **5**.2 (2011), 127–145, doi:10.3366/ijhac.2011.0028.

[5] M.F. Goodchild and L.L. Hill, Introduction to digital gazetteer research, *International Journal of Geographical Information Science* **22**(10) (2008), 1039–1044, doi:10.1080/13658810701850497.

[6] L.L. Hill, *Georeferencing. The Geographic Association of Information*, Digital Libraries and Electronic Publishing, Cambridge / Massachusetts / London, 2006.

[7] Points of Interest SWG | OGC. http://www.opengeospatial.org/projects/groups/poiswg.

[8] R. Singh, International Standards for Gazetteer Data Structures, in: *Placing Names. Enriching and Integrating Gazetteers*, M.L. Berman, R. Mostern and H. Southall, eds, Bloomington / Indianapolis, 2016, pp. 67–79.

[9] M.L. Berman, R. Mostern and H. Southall, Introduction, in: *Placing Names. Enriching and Integrating Gazetteers*, M.L. Berman, R. Mostern and H. Southall, eds, Bloomington / Indianapolis, 2016, pp. 1–11.

[10] A. Rutz, Doing territory. Politische Räume als Herausforderung für die Landesgeschichte nach dem "spatial turn", in: *Methoden und Wege der Landesgeschichte*, S. Hirbodian,

C. Jörg and S. Klapp, eds, Landesgeschichte, Vol. 1, Ostfildern, 2015, pp. 95–110.

[11] K.E. Grossner, Representing historical knowledge in geographic information systems., PhD thesis, University of California, 2010.

[12] R. Simon, L. Isaksen, E. Barker and P. de Soto Canamares, The Pleiades Gazetteer and the Pelagios Project, in: *Placing Names. Enriching and Integrating Gazetteers*, D.M. Berry, R. Mostern and H. Southall, eds, Bloomington / Indianapolis, 2016, pp. 97–109.

[13] K.E. Grossner, K. Janowicz and C. Keßler, Place, Period and Setting for Linked Data Gazetteers, in: *Placing Names. Enriching and Integrating Gazetteers*, M.L. Berman, R. Mostern and H. Southall, eds, Bloomington / Indianapolis, 2016, pp. 80–96.

[14] T. Berners-Lee, J. Hendler and O. Lassila, The Semantic Web, *Scientific American* **284(5)** (2001), 34–43.

[15] N. Guarino, Formal Ontology in Information Systems, in: *Formal Ontology in Information Systems. Proceedings of the 1st International Conference June 6-8*, N. Guarino, ed., IOS Press, Amsterdam / Trento, 1998, pp. 3–15.

[16] A.G. Cohn, Qualitative spatial representation and reasoning techniques, in: *KI-97: Advances in Artificial Intelligence*, G. Brewka, C. Habel and B. Nebel, eds, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 1997, pp. 1–30.

[17] J. Jones, W. Kuhn, C. Keßler and S. Scheider, Making the Web of Data Available Via Web Feature Services, in: *Connecting a Digital Europe Through Location and Place*, J. Huerta, S. Schade and C. Granell, eds, Cham, 2014, pp. 341–361, doi:10.1007/978-3-319-03611-3_20.

[18] M.J. Egenhofer, A formal definition of binary topological relationships, in: *3rd International Conference, FODO 1989 on Foundations of Data Organization and Algorithms*, W. Litwin and H.-J. Schek, eds, Lecture Notes in Computer Science, Vol. 367, New York, 1989, pp. 457–472.

[19] C.M. Korsmeier, *Die Ortsnamen der Stadt Münster und des Kreises Warendorf*, Westfälisches Ortsnamenbuch, Vol. 3, Verlag für Regionalgeschichte, Bielefeld, 2011.

[20] J. Trame, C. Keßler and W. Kuhn, Linked Data and Time – Modeling Researcher Life Lines by Events, in: *Spatial Information Theory*, T. Tenbrink, J. Stell, A. Galton and Z. Wood, eds, Springer International Publishing, Cham, 2013, pp. 205–223. ISBN 978-3-319-01790-7.

[21] A.G. Lowerre, Geospatial Technologies and the Geography of Domesday England in the Twenty-First Century, in: *Domesday now. New approaches to the inquest and the book*, D. Roffe and K.S.B. Keats-Rohan, eds, Woodbridge, 2016, pp. 219–246.

[22] Hull Domesday Project. http://www.domesdaybook.net/home.

[23] R. Mostern and I. Johnson, From named place to naming event: creating gazetteers for history, *International Journal of Geographical Information Science* **22**(10) (2008), 1091–1108, doi:10.1080/13658810701851438.

[24] E. Isenmann, *Die deutsche Stadt im Mittelalter 1150-1550. Stadtgestalt, Recht, Verfassung, Stadtregiment, Kirche, Gesellschaft, Wirtschaft*, Wien / Köln / Weimar, 2014.

[25] A. Hanschmidt, Zwischen bürgerlicher Stadtautonomie und fürstlicher Stadtherrschaft (1580-1661), in: *Geschichte der Stadt Münster*, F.-J. Jakobi, ed., Münster, 1993, pp. 249–300.

[26] A. Rutz, Grenzen im Raum – Grenzen in der Geschichte. Probleme und Perspektiven, in: *Grenzen im Raum – Grenzen in der Literatur (Zeitschrift für deutsche Philologie. Sonderheft zum Bd. 129)*, E. Geulen and S. Kraft, eds, Berlin, 2010, pp. 7–32.

[27] M. Worboys, Event-oriented approaches to geographic phenomena, *International Journal of Geographical Information Science* **19**(1) (2005), 1–28, doi:10.1080/13658810412331280167.

[28] A. Galton, Fields and Objects in Space, Time, and Spacetime, *Spatial Cognition & Computation* **4.1** (2004), 39–68, doi:10.1207/s15427633scc0401_4.

[29] E. Hyvönen, J. Väätäinen, J. Tuominen and T. Kauppinen, Representing and Utilizing Changing Historical Places as an Ontology Time Series, in: *Geospatial Semantics and Semantic Web: Foundations, Algorithms, and Applications*, N. Ashish and A. Sheth, eds, Springer-Verlag, 2011, Book chapter.

[30] LODE: An ontology for Linking Open Descriptions of Events. http://linkedevents.org/ontology/.

[31] Event Ontology. http://motools.sourceforge.net/event/event. html.

[32] PeriodO – Periods, Organized. http://perio.do/.

[33] F. Niccolucci and S. Hermon, Representing gazetteers and period thesauri in four-dimensional spacetime **17**(1) (2016), 63–69, doi:10.1007/s00799-015-0159-x.

[34] T. Kauppinen, G. Mantegari, P. Paakkarinen, H. Kuittinen, E. Hyvönen and S. Bandini, Determining relevance of imprecise temporal intervals for cultural heritage information retrieval, *International Journal of Human-Computer Studies* **68**(9) (2010), 549–560, doi:10.1016/j.ijhcs.2010.03.002.

[35] K. Grossner and E. Meeks, Topotime: Representing historical temporality, in: *Proceedings of Digital Humanities 2014*, Lausanne, 2014.

[36] Time Ontology in OWL. https://www.w3.org/TR/owl-time/.

[37] G.-A. Nys, M.V. Ruymbeke and R. Billen, Spatio-temporal reasoning in CIDOC CRM: an hybrid ontology with GeoSPARQL and OWL-Time, in: *Proceedings of the 2nd Workshop On Computing Techniques For Spatio-Temporal Data in Archaeology And Cultural Heritage co-located with 10th International Conference on Geographical Information Science (GIScience 2018)*, A. Belussi, R. Billen, P. Hallot and S. Migliorini, eds, Melbourne, 2018.

[38] P. Billion, *Graphische Zeichen auf mittelalterlichen Portolankarten. Ursprünge, Produktion und Rezeption bis 1440*, Marburg, 2011.

[39] B. Mons and J. Velterop, Nano-Publication in the e-science era, in: *International Semantic Web Conference*, Washington, D.C, 2009.

[40] P. Golden and R. Shaw, Nanopublication beyond the sciences: the PeriodO period gazetteer, *PeerJ Computer Science* **2** (2016), doi:10.7717/peerj-cs.44.

[41] The CIDOC CRM. http://cidoc-crm.org/.

[42] M. Rehbein, Ontologien, in: *Digital Humanities: Eine Einführung*, F. Jannidis, H. Kohle and M. Rehbein, eds, Stuttgart, 2017, pp. 162–176, doi:10.1007/978-3-476-05446-3_11.

[43] Versions of the CIDOC-CRM | CIDOC CRM. http://new. cidoc-crm.org/versions-of-the-cidoc-crm.

[44] Erlangen CRM / OWL - An OWL DL 1.0 implementation of the CIDOC Conceptual Reference Model (CIDOC CRM). http://erlangen-crm.org/current-version.

[45] Deutsche Nationalbibliothek - GND. http://www.dnb.de/DE/ Standardisierung/GND/gnd_node.html.

[46] T. Kluttig, Gemeinsame Normdatei und Archive - was soll das?, in: *Nachlässe - Neue Wege der Überlieferung im Verbund*, C. Rehm, M. Storm and A. Wettmann, eds, Veröffentlichungen des Sächsischen Staatsarchives. Reihe A: Archivverzeich-

nisse, Editionen und Fachbeiträge, Vol. 17, Halle (Saale), 2014, pp. 85–88.

[47] GND Ontology. http://d-nb.info/standards/elementset/gnd.

[48] Wikidata. https://www.wikidata.org/.

[49] Pleiades. http://pleiades.stoa.org/.

[50] GOV :: Das Genealogische Orts-Verzeichnis. http://gov. genealogy.net/search/index.

[51] J. Schnadt, Das genealogische Ortsverzeichnis (GOV). Familienforschung mit geodätischem Raumbezug, *Vermessung Brandenburg* **1** (2011), 35–42.

[52] Getty Thesaurus of Geographic Names (Getty Research Institute). https://www.getty.edu/research/tools/vocabularies/tgn/.

[53] GeoNames. http://www.geonames.org/.

[54] World-Historical Gazetteer. A project of the World History Center at the University of Pittsburgh. http://whgazetteer.org/.

[55] J. Zedlitz and N. Luttenberger, Modelling (Historical) Administrative Information on the Semantic Web, in: *The Second International Conference on Building and Exploring Web Based Environments WEB 2014*, Chamonix, 2014, pp. 33–39.

[56] G. Hiebel, M. Doerr and Øyvind. Eide, CRMgeo: A spatiotemporal extension of CIDOC-CRM, *International Journal on Digital Libraries* (2016), 1–9, doi:10.1007/s00799-016-0192-4.

[57] M. Doerr, A. Kritsotaki and K. Boutsika, Factual Argumentation – a Core Model for Assertions Making, *J. Comput. Cult. Herit.* **3**(3) (2011), 8:1–8:34, doi:10.1145/1921614.1921615.

[58] M. Theodoridou, M. Daskalaki, M. Doerr and G. Bruseker, Methodological tips for mappings to CIDOC CRM, in: *44th Computer Applications and Quantitative Methods in Archaeology Conference (CAA 2016) "Exploring Oceans of Data"*, Oslo, 2016-04-29.

[59] GeoSPARQL - A Geographic Query Language for RDF Data | OGC. http://www.opengeospatial.org/standards/geosparql.

[60] DNB, Katalog der Deutschen Nationalbibliothek - Hochstift Münster. https://portal.dnb.de/opac.htm?method= simpleSearch&cqlMode=true&query=idn%3D042794803.

[61] Help:Qualifiers - Wikidata. https://www.wikidata.org/wiki/ Help:Qualifiers.

[62] Vocabularies - place-types. https://pleiades.stoa.org/ vocabularies/place-types.

[63] Name Romanization Guide. https://pleiades.stoa.org/help/ name-romanization-guide.

[64] S. Gillies, Vocabularies, 2011-12-08. https://pleiades.stoa.org/ vocabularies.

[65] Vocabularies - time-periods. https://pleiades.stoa.org/ vocabularies/time-periods.

[66] GOV :: Liste aller Objekttypen. http://gov.genealogy.net/type/ list.

[67] M. Weber, Die Stadt, *Archiv für Sozialwissenschaft und Sozialpolitik* **47** (1921), 621–772.

[68] T. Kauppinen, J. Väätäinen and E. Hyvönen, Creating and Using Geospatial Ontology Time Series in a Semantic Cultural Heritage Portal, in: *S. Bechhofer et al.(Eds.): Proceedings of the 5th European Semantic Web Conference 2008 ESWC 2008, LNCS 5021*, Springer-Verlag, Tenerife, 2008, pp. 110–123, doi:10.1007/978-3-540-68234-9_11.

[69] T. Kauppinen and E. Hyvönen, *Modeling and Reasoning about Changes in Ontology Time Series*, in: *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, R. Kishore, R. Ramesh and R. Sharman, eds, Integrated Series in Information Systems, Springer-Verlag, New York (NY), New York, NY, 2007, pp. 319–338.

[70] E. Hyvönen, J. Tuominen, E. Ikkala and E. Mäkelä, Ontology Services Based on Crowdsourcing: Case National Gazetteer of Historical Places, in: *Proceedings of the International Semantic Web Conference 2015 Posters & Demonstrations Track*, CEUR-WS Proceedings, Bethlehem, PA, USA, 2015, Vol 1486.

[71] J. Stadler, Conceptualizing and Mapping Geocultural Space, *International Journal of Humanities and Arts Computing* **9.2** (2015), 133–141, doi:10.3366/ijhac.2015.0145.

[72] R. Battle and D. Kolas, GeoSPARQL: Enabling a Geospatial Semantic Web, *Semantic Web* **3.4** (2012), 355–370, doi:10.3233/SW-2012-0065.

[73] Pleiades Data Model. https://pleiades.stoa.org/help/ pleiades-data-model.

[74] GOV :: Münster. http://gov.genealogy.net/item/show/ MUNTERJO31TX.

[75] G. Flouris, I. Fundulaki, P. Pediaditis, Y. Theoharis and V. Christophides, Coloring RDF Triples to Capture Provenance, in: *The Semantic Web - ISWC 2009: 8th International Semantic Web Conference, ISWC 2009, Chantilly, VA, USA, October 25-29, 2009. Proceedings*, A. Bernstein, D.R. Karger, T. Heath, L. Feigenbaum, D. Maynard, E. Motta and K. Thirunarayan, eds, Springer, Berlin, Heidelberg, 2009, pp. 196–212, doi:10.1007/978-3-642-04930-9_13.

[76] J. Zhao and O. Hartig, Towards interoperable provenance publication on the Linked Data Web, in: *5th Linked Data on the Web (LDOW) Workshop at the World Wide Web Conference*, CEUR Workshop Proceedings, 2012.

[77] PROV-O: The PROV Ontology. https://www.w3.org/TR/ prov-o/.

[78] Irish Historic Towns Atlases Gazetteers, 2015-10-27. https:// www.ria.ie/additional-resources.

[79] R. Mostern, Historical Gazetteers: An Experiential Perspective, with Examples from Chinese History, *Historical Methods: A Journal of Quantitative and Interdisciplinary History* **41**(1) (2008), 39–46, doi:10.3200/HMTS.41.1.39-64.

[80] M. Balzer, Die Stadtwerdung - Entwicklungen und Wandlungen vom 9. bis 12. Jahrhundert, in: *Geschichte der Stadt Münster*, F.-J. Jakobi, ed., Münster, 1993, pp. 53–90.