(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

The gazetteer content model issue: Could Spatial Data Infrastructures provide it?

F.J. López-Pellicer[1], F.J. Zarazaga-Soria[1], A. Mogollón-Díaz[2], J. Nogueras-Iso[1], and P.R. Muro-Medrano[1]

[1] Computer Science and Systems Engineering Department, University of Zaragoza
Maria de Luna, 3, 50015 Zaragoza, Spain
{fjlopez,javy,jnog,prmuro}@unizar.es
http://iaaa.cps.unizar.es

[2] Computer Science Department, Pontifical University of Salamanca
Madrid campus, Paseo Juan XXIII, 3, 28040 Madrid, Spain
alxmog@upsam.net
http://www.upsam.com

Summary. The gazetteer is perceived differently in different contexts: Geographical Information Systems, Digital Libraries, Toponymic Standardization Processes and Cartotoponomy. These contexts shape the content of the SDI gazetteer. This paper analyzes the conventional core properties and the nature of the gazetteer to ease the building of SDI gazetteers.

Key words: Gazetteer, Spatial Data Infrastructures, Geographical Information Systems, Digital Libraries, Toponymic Standardization Processes, Cartography

1 Introduction

The *spatial reference system* (SRS) of *Spatial Data Infrastructure* (SDI) average users is not coordinate-based but place-name based. Users describe and retrieve data using conventional place-names whose themes range from more concrete ones (natural landscape features, populated places and localities, transportation routes and constructed features) to less tangible ones (civil and political divisions, administrative areas, cadastral parcels and postal addresses). Several SDI experts recognize that data in the SDI should also be indexed using place-name SRSs to increase the recall and/or the precision of place-names constrained queries [1, 2]. Therefore defining and supporting a reliable conventional place-name vocabulary is an issue on the SDI development. This vocabulary should deal with problems such as which place-names should be used, which is the role of synonyms, how wide is the area referred

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

by a place-name or recognizing which place the user is asking for. The *conventional wisdom* says that an specialized artifact named *gazetteer* is able to deal with this complex task. However, it seems that it does not exist an unified opinion about the gazetteer content model. Instead, there are different frameworks that identifies the most important types of information that a gazetteer should contain [3, 4, 5].

Gazetteers, places and place-names. A *gazetteer* is any artifact that contains authoritative facts about places and place-names. A *place* or *geographic feature* is any relatively permanent part of the natural or man-made landscape or seascape that has recognizable identity within a particular cultural context [6]. Different kinds of *named geographic features* may be found in gazetteers. The traditional categories are natural landscape features, populated places and localities, civil and political divisions, administrative areas, transportation routes and constructed features [7]. New ones categories range from devices to utilities service areas. A *place-name*, *toponym* or *geographic name* is an official or conventional name used in everyday language to refer or to identify a place, phenomenon or area that have a cultural recognized identity [5]. However, it must consider that the main function of place-name is to serve as a label so "*its semantic meaning, even though it is evident, is only a consequence of its role as a label*" [8] which depend on the application domain. In other words, the concept place-name, which is closely related to the concept label, might have a slightly different intension and extension than the conventional place-name concept within an application domain. For instance, the intension of the place-name concept in the geographic information context includes their use as a identifier [4] or as a metonymic substitute of a coordinate-based representation of location [9]; in the same context, its extension broadens to include codes such as "*1*" [4] and hierarchical specified place names such as "*New York City, New York State, USA*" [10].

Gazetteer architecture. The gazetteer architecture is a core subsystem of some SDI architectures. A good example is the *Open Geospatial Consortium* (OGC) *Geospatial Portal Reference Architecture* [2]. This reference model documents a core set of standard-based services that a geoportal architecture should have: portal services to provide a single point of access, catalog services to locate geospatial services and data, portrayal services to present data to the user and data services to provide data content and data processing. Its gazetteer subsystem is composed of a gazetteer client that provides users the capability to explore a spatially organized collection of named features, a gazetteer service that allow users to query well-known place-names to retrieve named features, and a gazetteer content model. OGC proposes as gazetteer a refactored ISO 19112 content model published through a *Web Feature Service* (WFS) [11]. However, some reference SDI such as United States geodata.gov[3] offer a non WFS gazetteer service[4]. Other SDIs publish their own gazetteer content model

[3] `http://gos2.geodata.gov/wps/portal/gos`
[4] `http://geonames.usgs.gov/pls/gnis/x`

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

through a WFS (e.g. the *Canadian Geospatial Data Infrastructure* (CGDI) [5] [12] and the *Spanish Spatial Data Infrastructure* [6] (IDE-E)). Even in those cases the content model is far from ISO 19112. For example, the *Spanish SDI Working Group* is developing the *Spanish Gazetteer Mode* (*Modelo Español de Nomenclátor*, MEN) based on the *Alexandria Digital Library Gazetteer* [13] (ADL) content model [14] that would become the recommended content model in Spain for both gazetteer services and sharing standard gazetteer data between administrations.

Disagreements about gazetteer role. One might ask why SDIs have not widely adopted the proposed OGC best practice. Some clues could be found in the role given to the place-name by SDI users. It has been detected some of these roles through the development of the MEN [15]. For instance, some toponymists have proposed that variant spellings must be out the national gazetteer because they could not be standardized. Other experts such as government officials responsible for cadastral management did not agree with the content of some fields as they appeared to be useful for naive exploratory search and not for expert data retrieval. Finally, local land managers have criticized that the spelling variants were considered as alternative place-names instead of attributes of the place-name which the spelling variants derived. After analyzing the arguments, these roles appear to be grouped as follow:

1. The place-name is a *geographic identifier* for Geographical Information System people. They see the gazetteer as the container of a spatial reference system made of location types and their respective location instances. A location instance record refers to a place and contains a preferred place-name written form, which acts as unique identifier in a context (e.g. "*Nowhere, Oklahoma*"), alternative place-names written forms (e.g. "*Madrid*", "*Madrit*", "*Majerit*") and a spatial footprint described using a coordinate reference system (e.g. Denmark footprint as "*West 7.9, South 54.3, East 13.2, North 57.8*") or a place-name reference system (e.g. Denmark footprint as "*Hovedstaden, Midtjylland, Nordjylland, Sjlland and Syddanmark*"). These place-names and the footprint may be applied for data retrieval as query constraint values. This is the conventional SDI gazetteer behavior [2] [1]. It was considered that a conventional gazetteer service may offer some explicit thesauri functionality [16], however these efforts seem to be discontinued.
2. The place-name as *encyclopaedia entry*. Digital Libraries people treat the gazetteer as a geographic dictionary, an encyclopaedia of places [3], thesauri of places [17] or geospatial ontology whose content could be used to georeference other resources following some guidelines (e.g. [18] describes how to use the *Getty Thesaurus of Geographic Names* (TGN) place-names

[5] `http://geoservices.cgdi.ca/cgi-bin/cgdi_gazeteer/cgdigaz.cgi?version=1.0&request=GetCapabilities`

[6] `http://www.idee.es/wfs/IDEE-WFS-Nomenclator/wfs?SERVICE=WFS&VERSION=1.0.0&REQUEST=GetCapabilities`

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

to describe the extension of a resource in absence of standard gazetteer data). These include not only typical SDI artifacts such as maps, imagery and aerial photography but also any kind of resource (e.g. archaeological records). These place-names and the footprint may be then applied for information retrieval [19].

3. The place-name as an *standardized administrative identifier*. It is widely recognized that a national geographic name standardization programme produces savings in time and money by increasing operation efficiency organizations [5]. A standardization programme also provides to an authority means to reinforce their authority and to recover and to disseminate a cultural heritage [7]. Typical standardization programme results are the standard gazetteer and the concise gazetteer [5]. This kind of datasets is often published through a SDI. Even there exist an international recommendation to establish these gazetteer as key component of a National SDI [20].
4. The place-name as *graphic symbol*. In the mapping process the written form of each place-names obtained from a gazetteer is applied to an entity shown graphically on a map. Both graphics and written forms that represent each entity used to vary depending on the context: scale, audience, output device, etc. A toponymic guideline contains the rules that would enable cartographers to apply correctly the place-names on maps [5]. These guidelines should contain or refer a toponym list or gazetteer but only as far as their content could be of benefit to the cartographic process. Therefore, this gazetteer seems to be better served by a label specialized portrayal service that provides georeferenced styled images of geographical names that works aside the SDI gazetteer service (e.g. [21] describe a service that provides place-names as both a GML/XML stream and an image georeferenced).

The real gazetteer. The above gazetteers can be found in a SDI. They share a common vision: a gazetteer is a structured collection whose core properties are geographic names, spatial footprints and feature types. Both temporal footprints and attribution to source are also core properties in some contexts [4, 13, 5]. Therefore, which is the gazetteer that should be used in the SDI context to index data? One, several or a mixture of the above gazetteers? Its obvious that identifying the *gazetteer content model behind the stage* may ease the decision.

This article is organized as follows. Section 2 analyzes the conventional core properties of gazetteer data. This analysis will ease a gazetteer description Section 3 characterizes the common gazetteer types. However, these gazetteers may be slices of a more complex reality. Section 4 abstracts core and other properties to seven gazetteer natures. Finally, some conclusions are settled about SDI and gazetteers.

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

2 What can be found in a gazetteer?

There are five core properties that can be found implicitly or explicitly in any gazetteer to describe a geographic feature: geographic names, spatial footprints, feature types, temporal footprints and attribution to sources. From existing gazetteers it is possible to trace variations to the features associated to these properties. This section shows a summary of these properties and their features (see Table 1).

Table 1. Core properties characteristics

Property	*Variants*
Geographic name	Linguistic, Cartography, Names authority, Geographic information
Spatial footprint	Assert, Outline, Abstraction, Inference, Reference, Whole
Feature type	Ontology, Thesauri, List
Temporal footprint	Assert, Outline, Abstraction, Reference
Attribution to source	Authority, Owner, Reference

Geographic names. It is an official or conventional name used in everyday language to refer or to identify a place, phenomenon or area that have a cultural recognized identity in the Earth surface. The detail of the data depends on the criteria used to compile the geographic name data. From the most to the least detailed:

1. "*Linguistic criteria*". Along with the written form, they need a presentation written form for collating lists, a division in parts (e.g. there are indivisible place names such as "Lake Placid", populated place, and divisible place names such as "Lake Placid", lake), the orthography and spelling variations (e.g. "Ejea", "Egea"), the language, the alphabet or script, the inflection, the pronunciation, the usage frequency, the source and their confidence [13].
2. "*Cartographic criteria*". Along with the presentation written form, which depends on the map context attributes such as scale, they need the language, the alphabet or script, the legal status (e.g the Pan-Canadian names that require presentation in both official languages of Canada on federal maps), the source references and map application points depending on map context and surrounding features [5].
3. "*Names authority criteria*". Along with the standardized written form, they need a presentation written form for collating lists, a division in parts, alternative names, the language, the alphabet or script, the legal status and the source references [5].

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

4. "*Geographic information criteria*". Along with the written form, maybe modified to act as a unique identifier, alternative names, the orthography and spelling variations and the source references [4].

Spatial footprints. It is an official or conventional description of the named place location without using their geographic names. A place could hold different spatial footprints, each one associated a time period. Also, user needs may cause different description encodings. And finally, the accuracy of the description depends on both the users goals and the description methodology. From the most to the least detailed:

1. An "*assert*", a semantic and spatial fit bounding geometry or a collection of its constituent features. A fuzzy assert falls within this category [4].
2. An "*An outline*", a set of points or a network of lines derived from the feature [5].
3. An "*abstraction*", an application point within or in the frontier of the feature [5].
4. An "*inference*", a bounding or generalized polygon derived from the feature geometry. Any generalization falls within this category [13].
5. A "*reference*", something such as the maps, charts or administrative area within which the feature lie. The reference might be ad-hoc[5] or semantically motivated[13].
6. A "*whole*", typically an administrative area or a natural feature landscape within which the feature is a division [4].

Feature types. It is an official or conventional category that is applied to group features. It is the holder of the cultural concept behind the feature instance. A feature type may imply a network of semantic relationships between a feature and overlapping and near features. A high correlation between the the semantic richness of the feature type and their reflection on explicit or derived data relationships allows inference processes. Therefore, from the most to the least correlated:

1. An "*ontology*", the feature type catalog is an ontology and gazetteer data relationships are consistent with their respective feature type relationships.
2. A "*thesauri*", the feature type catalog is a thesauri[4] or ontology and items in the gazetteer holds relationships consistent with hierarchical relationships between their feature types.
3. A "*list*", the feature type catalog is a list[5], a thesauri or an ontology but the gazetteer items do not have relationships or they are not related with those contained in the thesauri or ontology.

Temporal footprints. The temporal description of a gazetteer is an official or conventional description of the temporal validity of a gazetteer property. The temporal footprints form a partial ordered set that shows the gazetteer dynamics. The need to mix dates with named periods may cause vagueness

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

(e.g. "Budapest" is established not only in "1867" but also in the beginning of a named period known as "Ausgleich"). From the most precise to the most vague:

1. An "*assert*", a property value restrained by a bounded date range. Any fuzzy but bounded date range falls within this category [13].
2. An "*outline*", a property value restrained by a milestone within or in the limits of the temporal footprint [4].
3. An "*abstraction*", a property value restrained by a generic temporal concept such as historic, former, current or future relative to a known temporal pivot that belongs to the gazetteer such as the compilation date [5].
4. A "*reference*", a property value restrained by a named time concept. These named time concepts are defined independent of their use in a gazetteer [13].

Attribution to a source. The source of each gazetteer entry property value or a set of values may be known. The confidence level of the data is closely related to their source. From the most to the least confidence level:

1. An "*authority*", the property value proceed from an expert [5]. This expert may not own the data, but a definite process certify the value.
2. An "*owner*", the property value has a known owner, but their value may be not standardized or checked [4].
3. A "*reference*", the property value proceed from a known source[13].

3 The gazetteer views characterization

As it has been indicated in the introduction, there are four areas that require different kinds of gazetteer: Geographic Information Systems, Digital Libraries, Standardization Processes and Cartography. This section characterizes these gazetteers as follows: a contextual description, a definition by setting which features are mandatory for each property and an additional note to indicate relevant details that this analysis do not cover.

3.1 Geographic Information Systems: a Reference Gazetteer

A Geographic Information Systems gazetteer is the set of standards that give support to the use of geographic names as typed geographic identifiers. They are the result of the need of standardization of the indirect description of space properties using geographic names. Therefore the gazetteer contains from part of a spatial reference systems based on geographic identifiers till

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

a collection of these. An example is the gazetteer defined on the ISO 19112 [4]. These geographic identifiers must provide an unambiguous identification. That requires to rewrite many place names in the databases by adding an application in form of higher feature name (e.g. "*Paris*" become "*Paris, Texas*" to be distinguished from "*Paris, New York*" in a U.S. ISO 19112 compliant town gazetteer). These refactored geographic identifiers are applied as a replacement of complex geometries on databases for location and data retrieval purposes. For example, a land parcel geometry that identifies their location could be replaced by a compact, both machine and human readable, multipurpose representation: its address (e.g. "*2400 Jefferson Road, Paris, Texas*").

Its core properties could be summarized as follows:

1. The geographic name holds "*geographic information*".
2. The spatial footprints is an "*assert*" and optionally an "*abstraction*" or a "*whole*".
3. The feature types shape a "*thesauri*".
4. The temporal footprint is an "*outline*".
5. The source is an "*authority.*"

3.2 Digital Libraries: a Descriptive Gazetteer

A Digital Library gazetteer is a geospatial dictionary of geographic names. This kind of gazetteer, created mainly by librarians (e.g. [3]) and historians (e.g. *China Historical GIS*[7]), arise from the necessity of cataloging resources that may contain historical, contemporary and even future space references. In this context the gazetteer could appear as a geographic encyclopedia, such as the ADL, or as a thesaurus, such as the TGN. These are cited as sources of geographic names used in processes of metadata creation as the Digital Library gazetteer is applied to the indirect georeference of resources. For example, the spatial coverage of a norm of the city council of Zaragoza published in its Web portal [8] could be described as "*World; Europe; Spain; Aragón; Zaragoza; Zaragoza*" according to TGN or "*Zaragoza; Spain*" indicating that it is of type populated place according to ADL.

Its core properties could be summarized as follows:

1. The geographic name holds "*linguistic information*".

[7] `http://www.fas.harvard.edu/~chgis/`
[8] `http://www.zaragoza.es`

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

2. The spatial footprints could be expressed by a set of "*abstractions*", "*inferences*", "*references*" and "*wholes*".
3. The feature types shape a "*thesauri*" or an "*ontology*".
4. The temporal footprint could be expressed by a set of "*asserts*", "*outlines*", "*abstractions*" and "*references*".
5. The source is a "*reference*".

Other properties. It may include written presentation data.

3.3 Standardization Process: a Standard Gazetteer

The gazetteer is the sharing vehicle of the established official geographic names and their applications. The reference model that serves as base to these gazetteers has been established by the United Nations working group on geographic name standardization, the United Nations Group of Experts on Geographical Names (UNGEGN) [5]. Although UNGEGN is not a geographic names board or arbiter, their recommendations about the contains of a gazetteer are followed by place names boards of United Nations members. The goal of standard gazetteers is different from the previous ones. They aim to solve economic and political necessities. The economic one is derived from the reduction of costs caused by errors in the use of geographic names by administrations, companies and citizens. The political one is caused for the need of increasing the national self-esteem by protecting and/or recovering the linguistic heritage and by reinforcing the government authority on the territory by settling the official names. Among others, this result is seen in form of maps in which appear the official and/or standardized names.

Its core properties could be summarized as follows:

1. The geographic name follows "*names authority criteria*".
2. The spatial footprints is expressed by an "*abstraction*" and a "*reference*".
3. The feature types form a "*list*".
4. The temporal footprint could be expressed by a set of "*asserts*, "*outlines*" or "*abstractions*".
5. The source is an "*authority*".

Other properties. This data has a virus nature: changes spread through the country organizations. Applied to data retrieval may fail: usually the data uses non standardized geographic names. May include styled data to ease the production of official products such as maps.

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

3.4 Cartography: a Styled Gazetteer

This is usually a list of standardized geographic names of an administrative unit where next to each official name there is sufficient additional information that allows to locate their written form in a map and identify the graphical representation of the feature. Therefore their allocation in the map is dependent of criteria such as readability. In digital maps, the label layer have a subset of these attributes: a label, an attribute of type, an application point and style data, without additional relations with other layer. As labeling is an oriented process to show in a map the written forms of a geographic organization, to trace the relationships with other graphical representations such as a line or an area is not common.

Its core properties could be summarized as follows:

1. The geographic name provides "*cartographic information*".
2. The spatial footprints is expressed by an "*abstraction*" and a "*reference*".
3. The feature types form a "*list*".
4. The temporal footprint is an "*outline*" or a "*reference*"
5. The source is a "*reference*" or an "*authority*".

Other properties. This gazetteer possibly acts as a portrayal of other gazetteers. Its styled content is explicitly included or implicitly stated by a Toponymic Guide[5] and it may consider spatial footprint and geographic name changes if context properties such as scale change.

4 The gazetteer nature

We have described 5 gazetteer properties and there are at least 864 different ways of combining the features related to these properties. However, it appears to be obvious that different kinds of gazetteer are generated over the same gazetteer data again and again no wonder which features are expected. It appears that gazetteer data is interchangeable: one can feed a standard gazetteer from a descriptive gazetteer whose data came from a map. This is a very surprising and difficult to believe behavior if we realize that there does not exist a reference content model in the literature that founds these data transformations.

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

It seems that this model, "*the gazetteer*" (see Fig. 1) should describe the core gazetteer model and its extension points. Therefore, the sectoral gazetteers, the reference gazetteer, the descriptive gazetteer and the standard gazetteer, should be treated as specializations of this core morel. Finally, the styled gazetteer model seems to be better depicted as a different model responsible for rendering any kind of gazetteer data in different contexts.

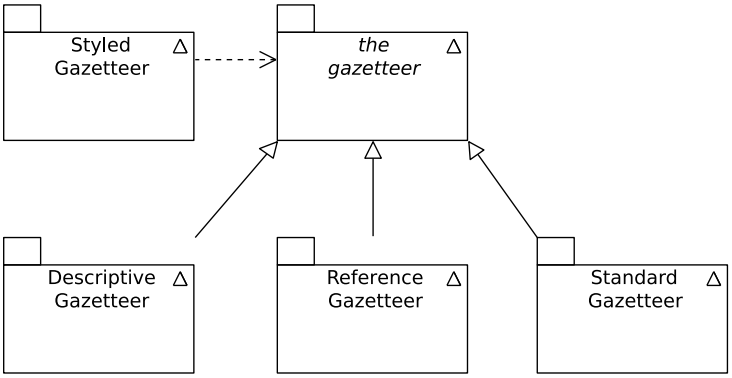

Fig. 1. The gazetteer complete picture

What are the abstract qualities of gazetteer model? It is possible to identify seven extension points.

1. *Multilingual data.* All contains explicit or implicit multilingual data. However, the gazetteer data are multilingual in a way that no previously cited gazetteer can satisfy completely. People and machines need translate place-names from a language to another (English to Spanish), from an encoding to another (Latin to Arabic), from a context to another (Environmental to Census) and even they need translate back and forth place-names from a reference system to another (Names to coordinates).
2. *Language data.* Reference and Standard Gazetteers follows people language. Their data unfolds sequentially, as the language does, to describe to describe or reference a concept (Street pre-type + Street Name , Street Name + Street Number)). This sequential way depend on the language and the context ("*calle mayor 126*" Spain "*126 calle mayor*" Puerto Rico "*126 main street*" US). Therefore, the gazetteer must be able to deal with them.

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

3. *Polyhierarchical data*. Gazetteer data are often hierarchically structured. The position in the hierarchy may affect the presentation data. Although a plain gazetteer is always possible, it is the spatial and semantic nature of gazetteer data that allow users and machines merge plain or simple gazetteers to obtain different hierarchical ones and more useful gazetteers[22].
4. *Multidimensional data*. This is a missed feature: none of these gazetteers are developed for analysis. However machines and users always use them for interpretation and analysis of the names (best name of a cluster of names), their spatial distribution (density), their evolution along time (name dynamics), their language (language coverage) and their theme (consistence) properties
5. *Clustered data*. Some information retrieval services relies that gazetteer data shows the property "*spatial proximity is matched by value similarity*" (e.g. two near peaks probably are part of the same crest and have a similar height, or a geocoding service can infer that a place name is similar and have the same properties as the queried one) although there exists discontinuities and different patterns when one move from one place to another, across temporal or geographical scales or feature types [23] that current gazetteers are not aware.
6. *Linked data*. The gazetteer data is a huge net of explicit and implicit temporal, topological and non hierarchical relationships. Property values of a gazetteer datum could be ordered by time value (Londinium, Augusta, Lundenwic, Lundenburg, London). Topological relationships based on spatial properties exists in gazetteer data. These relationships are often stable (Germany is adjacent to Poland). Finally, non hierarchical are consequence of the man-made nature of gazetteer data (Bern is the capital of Switzerland and the Canton of Bern).
7. *Presentation data*. The gazetteer often contains both business data and presentation data. Presentation data here refers to data that is applied or used to produce a non machine-readable product such as a hard copy map, electronic document or Web application. It is usually scale and neighbor feature dependent.

5 Conclusion

This article has detected something that lacks in the literature: a gazetteer reference model. *De facto* or published gazetteer content models such the Best Practice OGC WFS Gazetteer that use SDIs define only sectoral services. However, it seems that SDIs publish and use gazetteer data without bothering their respective sectoral origin. Even worse, the same data is reused in different content models, which are usually tailored, even overloaded, to serve a wide

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

range of users. This behavior is not a bad habit if SDIs developers can answer themselves whether the data model behind the stage supports:

1. Multilingual queries, and answers are consistent with the language.
2. Language queries, and answers are consistent with the language.
3. Different ways of describe the hierarchical location of a place.
4. Explicit and implicit relationships such as "*around*" or "*belongs*".
5. Inference operators and can act as a place ontology.
6. The difference between presentation data and business data.

It seems that SDI does not provide a solution for the gazetteer content model issue. It borrows available sectoral content models and mix the contained data without consider the consequences as there does not exist a general gazetteer model. Therefore, a future line of work is the development of this gazetteer model that should help the design and implementation of gazetteer databases that will feed the sectoral gazetteer views that SDIs publish and use.

Acknowledgments

This work is partially supported by the Spanish Ministry of Education and Science through the project TIN2006-00779 from the National Plan for Scientific Research, Development and Technology Innovation and by the University of Zaragoza (Spain) through the project UZ 2006-TEC-04. The work of F.J. López-Pellicer (ref. B136/2006) has been partially financed by a scholarship of the Government of Aragón (Spain). The work of A. Mogollón Díaz has been partially financed by the postgraduate program II-E "Scholarships for PHD Latin America Students MAEC-AECI 2005-06".

References

1. Nebert D (ed) (2004) Developing Spatial Data Infrastructures: The SDI Cookbook, v2.0. Global Spatial Data Infrastructure (GSDI) http://www.gsdi.org.
2. Rose L (2004) Geospatial Portal Reference Architecture: A community guide to implementing standards-based geospatial portals - 0.2. Open Geospatial Consortium Inc.
3. Hill L, Frew J, Zheng Q (1999) Geographic names: The implementation of a gazetteer in a georeferenced digital library. Technical report, Corporation for National Research Initiatives
4. International Organization for Standardization (ISO) (2003) Geographic information – spatial referencing by geographic identifiers. Published standard 19112:2003, ISO/TC 211 Geographic information/Geomatics

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

5. United Nations Group of Experts on Geographical Names (2006) Manual for the national standardization of geographical names. Department of Economic and Social Affairs, Statistics Division, United Nations, New York
6. Orth D, Payne R (1997) Principles, policies, and procedures: Domestic geography names
7. Orth D (1990) Organization and functions of a national geographic names standardization programme: A manual. World Cartography **Volume XXI**: 11–40
8. Kadmon N, Kadmon N (2000) Toponymy: The lore, laws and language of geographical names. Vantage Press
9. Markert K, Nissim M (2002) Towards a corpus annotated for metonymies: the case of location names. In: Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'2002), Las Palmas, Spain.
10. Axelrod A (2003) On building a high performance gazetteer database. In: Proceedings of the HLT-NAACL 2003 workshop on Analysis of geographic references, Morristown, NJ, USA, pp 63–68
11. Fitzke J, Atkinson R (2006) OGC Best Practices Document: Gazetteer Service - Application profile of the Web Feature Service Implementation Specification - 0.9.3. Open Geospatial Consortium
12. GeoConnections (2004) A Developers Guide to the CGDI: Developing and publishing geographic information, data and associated services, Ottawa, Ontario, Canada
13. Alexandria Digital Library Project (2004) Guide to the ADL Gazetteer Content Standard. University of California, Santa Barbara. 3.2 edn.
14. Rodríguez-Pascual A, López-Romero E, Abad-Power P, Sánchez-Maganto A (2005) Modelo de Nomencltor de España. Consejo Superior Geográfico, Ministerio de Fomento de España
15. Abad Power P (2006) Comentarios al Modelo de Nomenclátor de España - grupo de trabajo del modelo de nomenclátor de España - Consejo Superior Geográfico pabad@fomento.es.
16. Atkinson (2001) OGC Discussion Paper: Gazetteer Service Specification - 0.84. Open Geospatial Consortium
17. J. Paul Getty Trust (2007) The Getty Thesaurus of Geographic Names (TGN). http://www.getty.edu/research/conducting_research/vocabularies/tgn/about.html (last access: 19 February 2007)
18. The Training Sub-group of the Government On-Line Metadata Working Group (2005) Government of Canada Metadata Implementation Guide for Web Resources, 4th edition. Treasury Board of Canada Secretariat, Ottawa, Ontario, Canada
19. Jones CB, Alani H, Tudhope D (2001) Geographical information retrieval with ontologies of place. In: COSIT. pp 322–335
20. Nairobi UNCHS/FIG/ISK International Conference secretariat (2001) Spatial information for sustainable development. In: Recommendations from The Nairobi UNCHS/FIG/ISK International Conference.
21. MacIntosh B, Panyszak M (2004) The Canadian Geographical Names Service (CGNS). Meeting of the working group of toponymic data files and gazetteers. Twenty-second Session of United Nations Group of Expers on Geographical Names
22. Nogueras-Iso J, López-Pellicer FJ, Lacasta J, Zarazaga-Soria FJ, Muro-Medrano PR (2007) Building an Address Gazetteer on top of an Urban Network Ontology.

(Draft) Lecture Notes in Geoinformation and Cartography (LNG&C). 2007, vol. XVII, p. 187-200. ISSN 1863-2246.

In: Ontologies for Urban Development: Interfacing Urban Information Systems. Studies in Computational Intelligence. Springer In press.
23. Arpinar IB, Sheth A, Ramakrishnana C (2004). In: Geospatial Ontology Development and Semantic Analytics. Blackwell Publishing