

**United Nations Group of Experts on  
Geographical Names**

**Second session**

New York, 3 – 7 May 2021

**Item 14 of the provisional agenda \***

**Toponymic data files and gazetteers**

**Linked Data considerations for geographical names  
standardisation**

Submitted by Australia \*\*

Summary

Linked data is a set of design principles for sharing machine-readable, interlinked data on the Internet, and open data is a concept whereby data can be freely used, reused and redistributed with no or minimal restrictions. Together, considering linked open data approaches in relation to geographical names data management could further the aims of the United Nations Group of Experts on Geographical Names by making gazetteers and related information more interoperable and reusable.

The paper is prepared in response to decision 1/2019/13 of the Group of Experts, made at its 2019 session, in which it recognized technological methods such as linked data for the provision of geographical names data to be considered by the Working Group on Toponymic Data Files and Gazetteers in future. It contains a discussion on the opportunities and challenges to linked open data approaches for geographical name data management, accounts about national and regional developments and actions that could be taken by names authorities and the Group of Experts.

---

\* GEGN.2/2021/1

\*\* Prepared by Susan Birtles (Australia) on behalf of the Working Group on Geographical Names Data Management (formerly known as Working Group on Toponymic Data Files and Gazetteers)

## **Background**

Having standardized geographical names which are easily findable and accessible for national and international use has been a constant goal in UNGEGN's vision over many years, and the ways in which names can be made accessible continually evolves. The potential for standardising geographical names utilizing the internet was first recognised in resolution VII/9 of the Seventh Conference in 1998, along with a recommendation that data on toponymic websites be provided free of charge in the interests of international standardization. That positive trend continues and is also fully compliant with, and recognizes, Res VIII/6 of the Eighth Conference 2002, which recommends integrating geographical names better into national, regional and international spatial data infrastructures (SDIs). Those recommendations have been widely adopted and the internet has been imperative to the dissemination of geographical names data and information for some time.

At its first session in 2019 the Group of Experts agreed (1/2019/13) to look again into use of the internet and openly accessible and integrated data. Specifically, for the Working Group on Geographical Names Data Management (formerly known as Working Group on Toponymic Data Files and Gazetteers) to consider methods such as Linked Data for the provision of geographical names data. This paper is in response to that resolution, and provides some general considerations and explanations about Linked Data, as well as national and regional examples on how to use and benefit from Linked Data.

A note on terminology: UNGEGN's Linked Data conversations to date have noted the benefits of 'machine-readable' data. A relatively new term now receiving increasing use in its place is 'machine-actionable' as it more readily describes the full ability of machines to find, access, interpret and use a resource on the web. The country examples in the appendix (and the previously submitted summary) generally use the term 'machine-readable' as it has been the commonly used term for some time. The main section of this paper has adopted the term 'machine-actionable'. However, within the context of this paper, both terms have the same meaning i.e. data that a machine can find, access, interpret, and use in some way.

### **From world wide web to semantic web**

We are all familiar with the World Wide Web (WWW); it allows us to search and connect to disparate and homogenous sources of information stored on servers across the internet. We use browsers to interact with the information, search engines to help us find things, and hyperlinks to navigate. But we have to understand or look-up the meanings and relationships between things to interpret and act on the information we find.

Where the WWW is a source of human readable information, the Semantic Web conceptualizes a web of machine-actionable data. Linked Data refers to a set of standards, practices, and tools for publishing and interlinking structured data on the internet by making use of various web-based technologies known as the Semantic Web Stack. As part of the broader vision of the Semantic Web, Linked Data makes standardised, tagged and machine-actionable data from heterogeneous sources available online for linking. For data publishers, it aims to efficiently maximise the capacity for interoperability and correct interpretation of published data. For data consumers it aims to maximise the efficient and correct re-use of data. In this way, it facilitates inter-organisational data sharing, linking and the development of context-driven applications that make use of data from various sources.

### **Unique locator for persistent identification of a resource and data linking- URIs**

One of major principles of the Linked Data is in use of persistent and unique Uniform Resource Identifiers (URI). It allows interlinking of all Linked Data to a single global data graph, as a unified method for interrelating, reusing and finding unique relationships between various data assets. As each of the data assets is explicitly identified and described in meaning and in format, it allows data to be linked to other data published in a machine-actionable way. Similarly to how each page we view

through a browser has a http address which uniquely identifies that page (e.g. <https://unstats.un.org/unsd/ungegn/> is a URI for UNGEGN's homepage), it's possible to allocate URIs for individual pieces of data available on the web. Those URIs can then be published in various combinations and formats.

The examples provided by New Zealand, which discusses the benefits of enabling linkages between records in different databases, highlights how this allocation of URIs can be done in practice and their importance for improvement of human interaction with homogenous sources of information. Here, hyperlinks included in New Zealand's online Gazetteer serve as resources which allow users to access related information from other locations on the web, however many of them have been manually added. In one of the New Zealand examples, and well explained in the Finnish example, URIs are generated for each record in the online gazetteer and names register respectively and, as such, information is made accessible to all internet users in human readable HTML format. Each piece of information (gazetteer record or name card) is uniquely identified with it's own URI, which can be copied into other places to direct users back to the source information in the gazetteer or register.

The examples mentioned above demonstrate the use of URIs for the improvement of human understanding and linking of data. To take this concept further by making all resources linkable; every data asset, links between data assets and classifications used to describe these data assets are all assigned with URIs and, through use of those URIs, are linked to the additional information of other assets through defined relationships. This machine-actionability of Linked Data augments our abilities to find, interpret, and analyse data to generate insights and make decisions.

The example from Australia discusses how it is not only the place names records that must be published in linked data formats. Supporting data (place type categories, name status codes, lists of names authorities etc) need to be defined in code lists with attached URIs, and there also needs to be ontological models that describe the linkages between all those elements.

### **Using Linked Data**

When applied to geographical names, Linked Data can open opportunities for bringing together data to generate information and support decision making. Further integrated examples are available from the Netherlands, where Kadastre has made a number of key government registries available as Linked Data. Also the example from Germany about bringing together various Linked Data and existing standards-based data. Since the data, their interconnections and their semantics are machine-actionable, various applications using this previously disparate information can be built and a user can be presented with various hyperlinks to find related information from other registries. For instance querying on certain types of places within a defined administrative area, then clicking through to find a related piece of legislation that defines terms.

Names authorities can benefit not just as publishers by encouraging easier and broader access to data, they can also benefit as consumers. The example from the Open European Location Services Project shares an example of how linked data enables searching for a toponym root across data in different languages.

Linked Data is subject to a network effect i.e. the applications and benefits grow as more data becomes linked. As this happens, Linked Data may help to enhance understand of complex and multi-disciplinary situations by revealing insights more efficiently and more comprehensively than is possible by manual analysis. It has potential to enable evidence-based policies and augment decision making by geographical names and other authorities.

### **Conclusion and Recommendations**

Methods of geographical names data management must keep pace with technology developments. Computers, the internet, and geographical information systems have all changed the way data is managed. Linked Data is simply another stage in the evolution of data management methodologies that the Group of Experts needs to encourage to facilitate widespread use of standardised names. The examples gathered for this paper reveal that where Linked Data is being implemented, it is quite often through pilot projects or demonstrators. Geographical names are one of the most widely used location identifiers and can be used to connect various pieces of disparate data. For this reason they are very worthwhile in considering during Linked Data developments; their inclusion enabling the demonstration both of the value of Linked Data and of standardised geographical names.

Linked Data is gaining increasing use across the world and will support UNGEGN's vision of making standardized geographical names more easily and more widely accessible for national and international use and re-use. It can also make the work of names authorities and toponymists much easier, and open up opportunities for new applications or streamlining production and maintenance of existing maps and gazetteers. For names authorities, geospatial and other agencies dealing with geographical names data management who have begun to incorporate Linked Data approaches, the role of standardisation becomes increasingly clear. The need for every thing and every relationship to be described is encouraging the codifying of knowledge and leading towards useable documentation that describes many of the nuanced relationships of geographical names. Even if not being implemented, consideration can still be given to preparing for Linked Data. Standardising place type categories and code lists for attributes (e.g. official status, naming authority, language of origin), and being able to describe relationships (e.g. between different types of names, attributes of names, or how a name relates to other data) are all important precursors to Linked Data implementations. The Group of Experts is encouraged to consider Linked Data as another type of technology or method that can be adopted in relation to geographical names data management.

#### **Points for discussion**

##### **The Group of Experts is invited to:**

- (a) Note that Linked Data is another stage in the evolution of data management methodologies that the Group of Experts needs to consider to facilitate widespread use of standardised names.
- (b) Encourage collaborations between national names authorities and linked data practitioners
- (c) Continue to share case studies, as they become available, that demonstrate the combined benefits of standardised geographical names and Linked Data.

## Appendix of examples

The following examples have been shared to help demonstrate the various stages to implementing and benefits of considering Linked Data. Each example also gives Linked Data explanations in different ways by different experts.

### New Zealand

The New Zealand Geographic Board (the Board) has been adding URL hyperlinks for the Dictionary of New Zealand Biography [1] to relevant records of their Gazetteer [2] to provide connection between personal names and corresponding biographical details. The work is manual and consideration needed to be given to the likely permanence of the biography URLs and maintenance activities once the initial links had been included.

The Board has also added URL hyperlinks in the Gazetteer for James Cook related place names, to Google's Content Creator [3] platform that was used for the Tuia250 commemorations [4]. Hyperlinks in the Gazetteer have also been added to send the user to Kā Huru Manu [5], a cultural heritage atlas with original Maoro names for much of the South Island or Te Waipounamu.

Waterways are the most common feature type in the Gazetteer (c.16300 names in territorial New Zealand). In early January 2019 the first iteration of a river names dataset was released internally within Toitū Te Whenua Land Information New Zealand. The advantages for the Board are (a) time saved manually capturing extents in the Gazetteer: with the feature ID linked, extents may be added to names automatically, and (b) data maintenance: with the IDs linked, the Board's Secretariat and Toitū Te Whenua could set up automatic notifications when the Board alters a name and/or assigns an extent, or when Toitū Te Whenua updates the extent based on new data. Adding Gazetteer IDs to waterways in the Topo data also completes the same process as the NZGB Secretariat's audit. It means that official names, names missing from the Gazetteer, duplicates, and errors are readily identifiable. For example, there are at least 70 recorded names missing from the Gazetteer.

Investment in linked data approaches may provide less resource intensive methods for these and similar activities in the future.

Links:

[1] <https://teara.govt.nz/en/biographies>

[2] <https://gazetteer.linz.govt.nz/>

[3] <https://earth.google.com/web/data=Mj8KPQo7CIExMjJFZ1puZmJFV2hFRFlvSFVNVG5YU05CRWZTYTBiSWoSFgoUMDA0MUM0QzM2QjEzOTM4N0NFNjU>

[4] <https://www.linz.govt.nz/regulatory/place-names/tuia-%E2%80%93-encounters-250>

[5] <https://www.kahurumanu.co.nz/atlas>

### Finland

The Geographic Names Register of the National Land Survey of Finland contains data on about 800,000 named places and their standardized names in five languages - Finnish, Swedish, North Saami, Inari Saami, and Skolt Saami. The Land Survey has taken its first steps towards linked open data on geographical names by introducing URI-based information cards for the places and names in the register. The URI of an information card works as a web address, making the information accessible to all internet users in human readable HTML format. The permanent URI allows the linking the information card to any site, document, or service on the internet.

The data for the information cards comes from a National Land Survey's geographic names web feature service (OGC API – Features). If an external web service utilizes the National Land Survey's open query interfaces for geographic names data, the information cards can be united with the query results, because an interface reply always includes the persistent identifier of places and names, on which also the URI is based. Moreover, the information may be found via the indexing systems of global web search engines.

Below are examples of URIs for the information cards of named places. The information card of the name opens by clicking the name in the information card of the place. In the top right of the card, another name may be entered to be searched for. The asterisk (\*) can be used as a general wildcard character and the question mark (?) to replace one character. The language of the card can be selected from Finnish (*Suomeksi*), Swedish (*På svenska*), or English (*In English*).

Populated Places: A city, Helsinki, <http://paikkatiedot.fi/so/1000772/12002368>

A house, Peltolampi, <http://paikkatiedot.fi/so/1000772/10768042>

Waters: A lake, Sevettijärvi, <http://paikkatiedot.fi/so/1000772/10824500>

A river, Ivalojoiki, <http://paikkatiedot.fi/so/1000772/10889831>

Buildings: A church, Pyhän Laurin kirkko, <http://paikkatiedot.fi/so/1000772/12008586>

A shopping center, Tripla, <http://paikkatiedot.fi/so/1000772/12013467>

World heritage sites: A Struve geodetic arc station point, Mustaviiri, <http://paikkatiedot.fi/so/1000772/12005185>

A sea fortress, Suomenlinna, <http://paikkatiedot.fi/so/1000772/12005145>

The home page for the information cards – in Finnish, Swedish and English – will open during the spring of 2021 at the address <https://tietokortit.maanmittauslaitos.fi/nimisto>.

## Australia

Each Australian state and territory – including Commonwealth waters and Antarctic areas of interest – has a place name registrar, naming board or committee for approving or registering place names. Each also manages their own place name databases. The use of consistent and accurate national level place names data is extremely important across daily activities in every sector. Via collaboration across all partners a common data structure, classification and technical integration architecture was developed to build the national Composite Gazetteer of Australia portal, which currently brings together approximately 290,000 named places from the various databases.

As the next step in promotion and popularisation of this important composite dataset, a pilot project was conducted to transfer the records into a Linked Data format. As part of this project, an ontological model to define classes and relationship was developed, followed by creation of a Linked Data registry, and a vocabulary defining Place Type. Supported by the Place Names ontological model, the National Composite Gazetteer of Australia has been created as a set of dynamically built landing pages representing each individual record from the gazetteer (the registry) to make the information accessible to all internet users in human readable HTML. Each record in this registry is associated with a URI. URIs are also assigned to each of the terms classifying those records (e.g. feature type, feature category, feature group, source registry, naming authority, etc).

To enable further re-use and interoperability of the Place Names data, a set of alternate profiles was developed as a set of the OGC<sup>1</sup> compliant web services, APIs<sup>2</sup>. For example, the [landing page](#) for the capital city Canberra also has a set of [alternate profiles](#). Users can utilise the National Composite Gazetteer of Australia Registry web services for geographic names. The query results will include the persistent identifier of place names, which may be further indexed by the global web search engines.

Below are links to examples of some vocabulary terms and named places. The vocabulary terms can also be accessed by clicking on the URI represented by the corresponding hyperlink on the landing page.

Population Centre: Canberra, [https://fsdf.org.au/dataset/placenames/placenames/ACT\\_ACT000175](https://fsdf.org.au/dataset/placenames/placenames/ACT_ACT000175)

Watercourse: Murray River, [https://fsdf.org.au/dataset/placenames/placenames/NSW\\_40833](https://fsdf.org.au/dataset/placenames/placenames/NSW_40833)

Island: Heard Island, [https://fsdf.org.au/dataset/placenames/placenames/AAD\\_825](https://fsdf.org.au/dataset/placenames/placenames/AAD_825)

Mountain: Budj Bim, [https://fsdf.org.au/dataset/placenames/placenames/VIC\\_123297](https://fsdf.org.au/dataset/placenames/placenames/VIC_123297)

## **Netherlands**

### Kadaster, Netherlands – Key Registries, Linked Data and the Toponamenzoeker

Linked Data, as part of the broader vision of the semantic web, is an innovative approach to making standardised, tagged and machine-readable data from heterogenous sources available online for linking. In this way, Linked Data is able to transcend some of the limitations of relational databases and, in making use of various web-based technologies known as the Semantic Web Stack, facilitates inter-organisational data sharing, linking and the development of context-driven applications which make use of data from various sources. In the context of ever-increasing amounts of data from such heterogenous sources, government organisations in particular are beginning to adopt linked data solutions in order to meet the demands for better interoperability both for smart government initiatives and the development of applications with improved usability for citizens (Folmer et al., 2020).

#### Kadaster and Linked Data

Kadaster, the Netherlands' Cadastre, Land Registry and Mapping Agency, has a lengthy record with regards to the application and ongoing development of linked data technologies. Indeed, many of the key government registrations have been available as linked open data for a number years including, the key register large-scale topography (Dutch acronym: BGT), the key register topography (Dutch acronym: BRT), the key register addresses and buildings (Dutch acronym: BAG) as well as the digital cadastral map, part of the key register cadastre (Dutch acronym: BRK). Recent developments with regards to the use of linked data technologies within the organisation has seen an updated approach taken to the publication of the key registers as linked data; an approach which has seen the topographical and address datasets published in shorter timeframes and with less resources than previous approaches. These recent publications can be accessed through the Kadaster triple store (<https://data.labs.kadaster.nl>) where services such as SPARQL, ElasticSearch, data browsing can be instantiated for end user use (Rowland et al., 2021, submitted for publication).

#### The Toponamenzoeker

<sup>1</sup> OGC – Open Geospatial Consortium

<sup>2</sup> API – Application Programming Interface

An example application that uses these services to improve searchability of geospatial data for the end user is the Toponamenzoeker. This tool makes use of the linked data publication of the BRT to allow users to search for objects using their geographical names, the results of which are returned on a mapping interface. Clicking on the returned object displays more information about the BRT also stored in the BRT. This application can be found here: <https://labs.kadaster.nl/demonstrators/namen-app/#/>.

#### Future Visions for Linked (Open) Data at Kadaster

The recent update to the publication approach for linked data at Kadaster is central to another initiative within the organisation, the development and publication of Kadaster's Knowledge Graph (Ronzhin et al., 2019). This graph is initially centred on the theme of building and makes integrated information from various registries available as linked open data to be queried by various end user groups. Similarly to the Toponamenzoeker, an application (Pandviewer) is planned for development which will allow these end users to interact with the linked data contained in the Graph to browser and display information across various registries for a single geographical object. Kadaster's effort supports the increased availability of linked open spatial data in the Netherlands generally, and within the Dutch government more specifically, of which there are a number of other related initiatives external to Kadaster.

#### External Linked Data Initiatives

There are a number of external linked open data (LOD) initiatives of note in the Netherlands. Firstly, from 2016, various linked open data pilots have been initiated in the context of the Dutch Service for Education Implementation (Dutch acronym: DUO) in order to improve the interoperability around sources of educational data. This data includes information relating to student enrolment, diploma registers, public student counts among other metrics. Included in the heterogenous sources used by this project is the BAG registry. A second initiative, Network Digital Heritage (in Dutch Netwerk Digitaal Erfgoed) is part of a government directed initiative which seeks to improve the standardisation and linking of heritage information. This is done by supporting relevant heritage organisations with the application of LOD principles within their organisations; making heritage information available and useful to both humans and machines.

#### References:

Folmer, E., Ronzhin, S., Van Hillegersberg, J., Beek, W., Lemmens, R. Business Rationale for Linked Data at Governments: A Case Study at the Netherlands' Kadaster Data Platform. IEEE. Access 8, 70822-70835, (2020).

Ronzhin, S., Folmer, E., Maria, P., Brattinga, M., Beek, W., Lemmens, R., van't Heer, R. Kadaster Knowledge Graph: Beyond the Fifth Star of Open Data. Information, 10(10). (2019).

Rowland, A., Folmer, E., Beek, W., Wenneker, R. (2021). Interoperability and Integration: An Updated Approach to Linked Data Publication at the Dutch Land Registry. Manuscript submitted for publication.

#### **Germany**

In 2019, the Federal Agency for Cartography and Geodesy (BKG) initiated activities on linked geodata as part of a spatial data infrastructure project in Germany. The focus of the project is on the implementation of an intelligent platform that supports the integration of various linked open data sources (e.g. Wikidata, LinkedGeodata.org) but also existing standards-based data (e.g. for INSPIRE). It is a matter of integrating all available data attributes, format specifications relevant for import or export and the development of new data services that are only possible through semantic data



integration. In order to provide a prototype of linked data in the BKG, various components had to be developed that were to enable the provision of linked data (SPARQL Endpoint), the conversion of various geofomats to RDF and the exploration and reconversion of linked data into geofomats for their provision.

A basic requirement for linked open data applications is to describe each resource of each triplet with a URI. It is recommended that these URIs be made resolvable in the web browser so that users can understand the data behind the URI and its relationships to other related data resources. The following example shows the implementation of Geographical Names (GN250) from the BKG, represented as a visual graph.

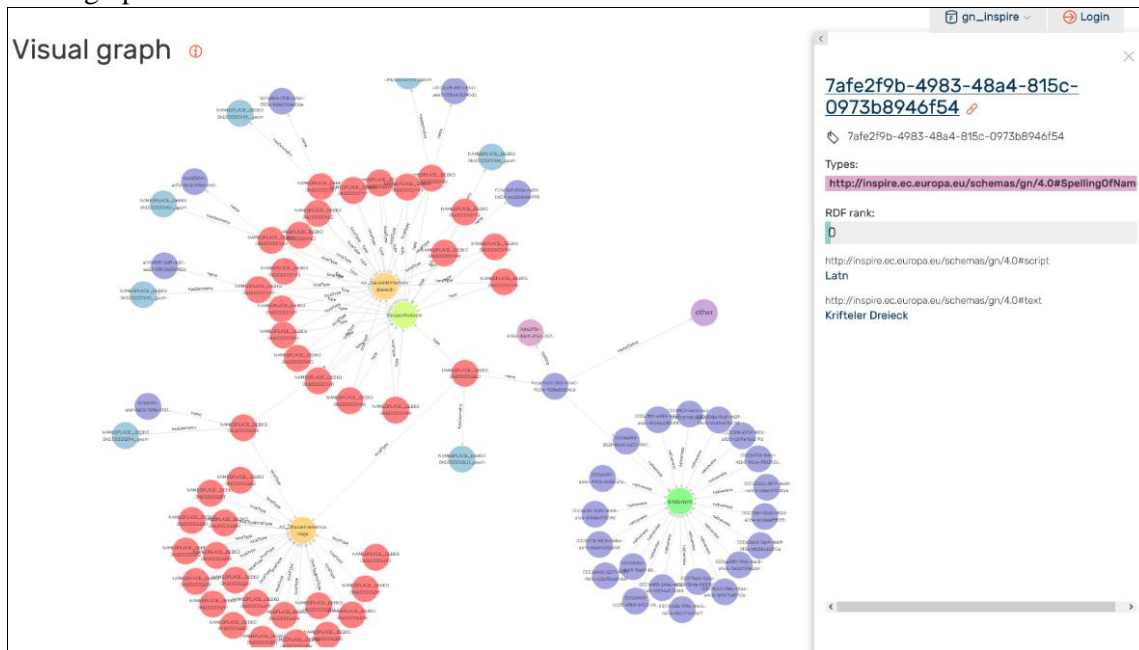


Figure 1: Visual graph of INSPIRE Geographical Names (GN250), using "Krifteler Dreieck" as an example. (Source:[https://ld.gdi-de.org/graphdb/graphs-visualizations?uri=http://ld.gdi-de.org/id/de.bund.bkg.inspire.gn250/NAMEDPLACE\\_DEBKGGND000052LU](https://ld.gdi-de.org/graphdb/graphs-visualizations?uri=http://ld.gdi-de.org/id/de.bund.bkg.inspire.gn250/NAMEDPLACE_DEBKGGND000052LU))

### European Location Services Project

The Linked Data technology as an innovative approach for publishing heterogeneous data sources on the web is able to transcend the traditional confines of separate databases, as well as the confines of separate institutions, keeping existing infrastructures intact. Moreover, exposing national data as Linked Data on the Web makes it a part of the Semantic Web. This allows to shift focus from collection and dissemination of data to meaningful data consumption [1]. The Open ELS project has shown the advantages this can have for geographical names, among others.

Open ELS was a two-year project co-financed by the European Union and coordinated by EuroGeographics to demonstrate the potential of a single access point of harmonised geospatial information in Europe. The project ended in 2019. It built on key results of the European Location Framework (ELF), a project for a practical implementation of the INSPIRE directive (Infrastructure for Spatial Information in Europe).

One goal of the Open ELS project was to make Open Data easier to use by Application Programming Interfaces (APIs) and Linked Data. Geographical names (GN) were one of the selected data for being exposed as Linked Data with published so-called RDFs and SPARQLs. Resource

Description Framework (RDF) is a directed, labelled graph data format for representing data in the Web. SPARQL is a query language for RDF. It can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware.

Four NMCAs co-operated to achieve the described goal: Norway, Finland, the Netherlands and Spain. The four countries used different approaches with the focus of achieving the common result of publishing Linked Data (in RDF) according to common ontologies, so-called Unique Resource Identifier (URI) structures and individual SPARQL endpoints. Only key choices like a common geometry were made to see if easy interoperability can be achieved in practice [1].

As a result, all four participating countries generated Linked Data versions of the two INSPIRE datasets, namely Administrative Units and Geographical Names. These data can be queried via national SPARQL endpoints. The results can be visualised on the country's own official base maps. The quantitative characteristics of the results of the transformation for geographical names can be seen in Table 1. Furthermore, a URI lookup service was created. This is a gazetteer service containing place names data from various European countries. The service response contains the URIs together with some general information on the spatial objects, such as the coordinates and the feature type classifications. A demonstration web page was built to prove the results and a Data Story (like an advertisement tool for data) is available at <https://data.labs.pdok.nl/stories/OpenELS/> [1].

Country	Total number of toponyms	Total number of triples	Size in MB (for rdf/xml serialisation)	Example URI
Finland	~1500000	8501970	1990	<a href="http://paikkatiedot.fi/so/openels/gn/GeographicalName/40472569">http://paikkatiedot.fi/so/openels/gn/GeographicalName/40472569</a>
The Netherlands	~1000000	10000000	2100	<a href="http://data.labs.pdok.nl/dataset/openels/gn/NamedPlace/NL.TOP10NL.GN.130456316">http://data.labs.pdok.nl/dataset/openels/gn/NamedPlace/NL.TOP10NL.GN.130456316</a>
Norway	949615	15817780	2763	<a href="https://data.geonorge.no/openels/gn/NamedPlace/654818">https://data.geonorge.no/openels/gn/NamedPlace/654818</a>
Spain	1132743	18491679	2400	<a href="http://datos.idee.es/recurso/openels/gn/NamedPlace/176097">http://datos.idee.es/recurso/openels/gn/NamedPlace/176097</a>

Table 1: Overview of datasets containing geographical names [1]

The Open ELS project demonstrated that Linked Data brings semantic interoperability by means of ontologies. Often, it is difficult to compare structures of administrative division between countries because they use language-specific names for levels of administrative division. With the help of a common ontology (in this case the INSPIRE Administrative Units ontology), it's possible to model national systems of administrative units and draw comparisons between them, avoiding misinterpretation caused by language specific notions. Linked Data makes it possible to link not only data at the ontological level, but also data at the instance level between data sets. This becomes important, for example, in the case of geographical names, since two different data sets contain information about the same object. Thus, they could be linked and data could be enriched with them. This makes it possible, for example, to retrieve the geographical name of a data set in another language [1].

An example for geographical names with “holy” as part of the name:  
The OpenELS project published more than 40 million place names (toponyms) and locations as Linked Data. Standardised concepts for describing the meaning of place names enables semantic interoperability between national data sets. Seamless access to such rich data allows for conducting interesting research. For example, linguists could potentially use this data to analyse the spatial distribution of common toponym roots. The following example shows locations of places that have “holy” as part of the name. Obviously, the root “holy” is spelled differently in different languages (

Table 2). The Linked Data technology makes it possible to formulate a single query that can interrogate the national endpoints in their native languages [1].

Language	Spelling
English	Holy
Dutch	Heilige
Norwegian	Hellig
Finnish	Pyhä

Table 2: Translation of "holy" into languages of project participants [1]

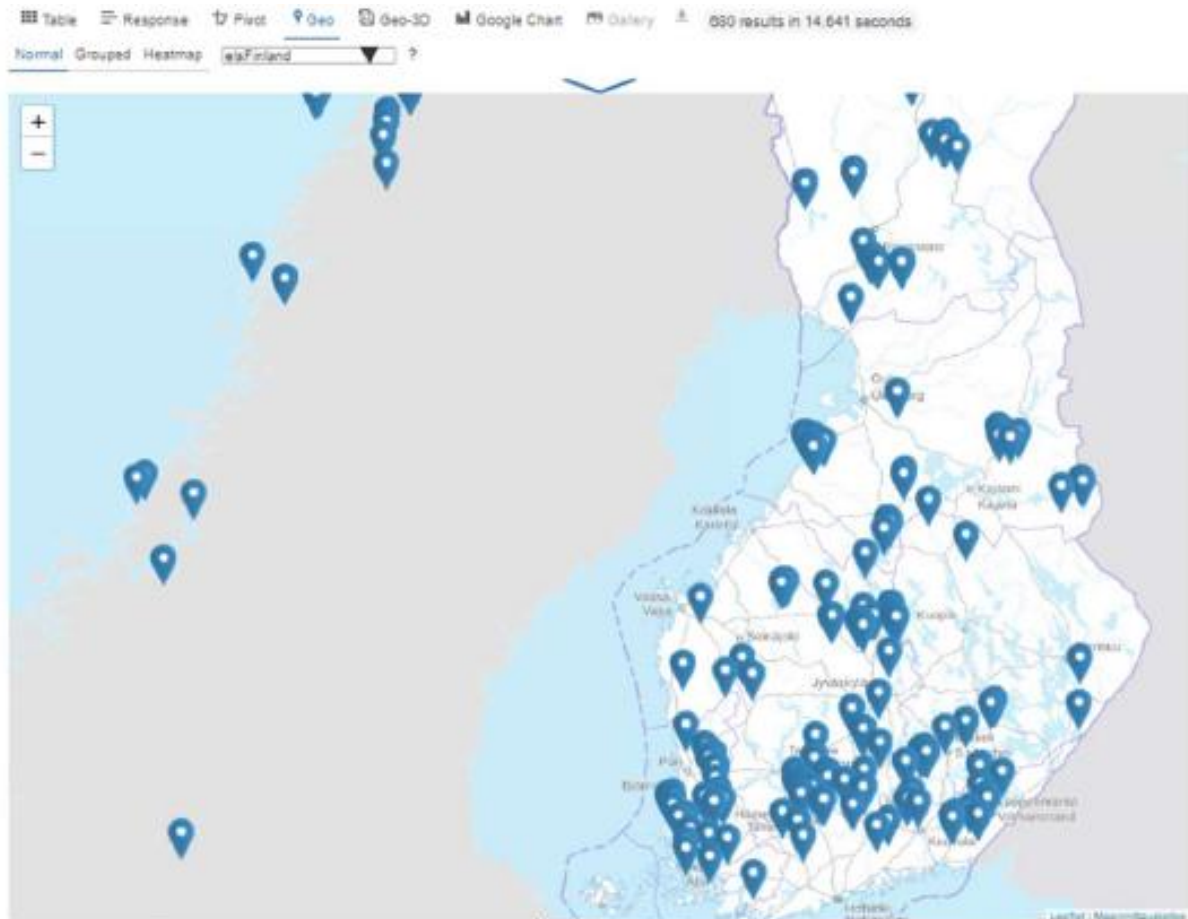


Figure 2: Visualisation of the results for the query retrieving toponyms that feature “holy” as part of the name [1]

#### References:

[1] “Next Generation of Spatial Data Infrastructure: Lessons from Linked Data implementations across Europe” from Stanislav Ronzhin et al.

#### Acknowledgements

This paper is prepared on behalf of the Working Group on Toponymic Data Files and Gazetteers with contributions and assistance from both working group members and their colleagues: Irina Bastrakova (Australia), Teemu Leskinen (Finland), Katja Hilgert and Pier-Giorgio Zaccheddu (Germany), Jasper Hogerwerf and Lexi Rowland (The Netherlands), Wendy Shaw (New Zealand).