A Data Repository for Named Places and Their Standardised Names Integrated With the Production of National Map Series

Teemu Leskinen

National Land Survey of Finland

Abstract. The Geographic Names Register (GNR) of the National Land Survey of Finland (NLS) is the authoritative geographic names data repository of Finland and one of the elements included in the national and international spatial data infrastructure. The GNR, initially developed in the late 1990s, was entirely renewed in an NLS development project during 2010–2014. The reform covered the names data model and database, the names data management and integration with the production of national maps, the production application as well as the production lines and organisation of work. The GNR named feature classification was also renewed, as were the GNR information services and data products.

Keywords: Geographic names data modelling, Names data management, Names data integration, Names data dissemination

1. Introduction

Finland is a multilingual country. Finnish and Swedish are the official state languages, and three indigenous Saami languages – Inari Saami, North Saami and Skolt Saami – have an official status in the four northernmost municipalities of Finland. In Finland, a specific geographic feature may have several parallel names in one or different languages. In bilingual Finnish–Swedish and Swedish–Finnish areas, a feature typically has two names, and in Lapland's Finnish–Saami language areas, a feature may have up to five names.

The National Land Survey of Finland (NLS) produces, in print, two national map series based on the NLS Topographic Database (TDB): the Basic Map 1:25 000 and the Topographic Map 1:50 000. In addition, seven national small scale map datasets – all incrementally generalised from the TDB data – are being produced at scales 1:100 000, 1:250 000, 1:500 000, 1:1 million,

1:2 million, 1:4.5 million and 1:8 million. In national maps all standardised names of the feature are presented, if any. Names in bilingual Finnish–Swedish and Swedish–Finnish areas occur in two languages, the name in the municipal majority language placed first. In Lapland, all existing Saami names of a given feature are presented, in a standard language order, together with the (possible) Finnish and Swedish name (*Figure 5*). The amounts of geographic names presented in national map series are:

- Basic Map 1:25 000 ca. 810 000 names;
- Topographic Map 1:50 000 ca. 360 000 names;
- Topographic Map 1:100 000 ca. 170 000 names;
- Topographic Map 1:250 000 ca. 45 000 names;
- Topographic Map 1:500 000 ca. 13 500 names;
- General Map 1:1 million ca. 2 500 names;
- General Map 1:2 million ca. 600 names;
- General Map 1:4.5 million ca. 200 names;
- General Map 1:8 million ca. 25 names.

The multilingual and multi-names circumstances in Finland and NLS national map production environment have guided the development of the NLS Geographic Names Register (GNR), initially established in the late 1990s. The GNR was entirely renewed in an NLS development project during 2010–2014. The reform covered the data model and database, the names data management and integration with the production of national maps, the production application as well as the production lines and organisation of work. The GNR named feature classification was also renewed, as were the GNR information services and data products.

2. Geographic Names Register

2.1. General

The Geographic Names Register of the National Land Survey is the authoritative geographic names data repository of Finland and one of the elements included in the national and international spatial data infrastructure. The GNR comprises the primary Place Name Register (PNR), containing information on over 800 000 named places and their standardised names, and the Map Name Register (MNR), an operational dataset holding the cartographic attributes (placement, typography) related to the PNR names selected to be presented in the NLS cartographic product in question.

2.2. Data Model

The GNR data model is feature-oriented and includes the objects Place, Place name and Map name. All objects are interconnected and provided with external persistent unique identifiers. A named feature (Place) has at least one name (Place name) – in Finland, commonly several names in different languages – that may have zero, one or multiple occurrences (Map names) in different cartographic products (*Figures 1 and 2*).

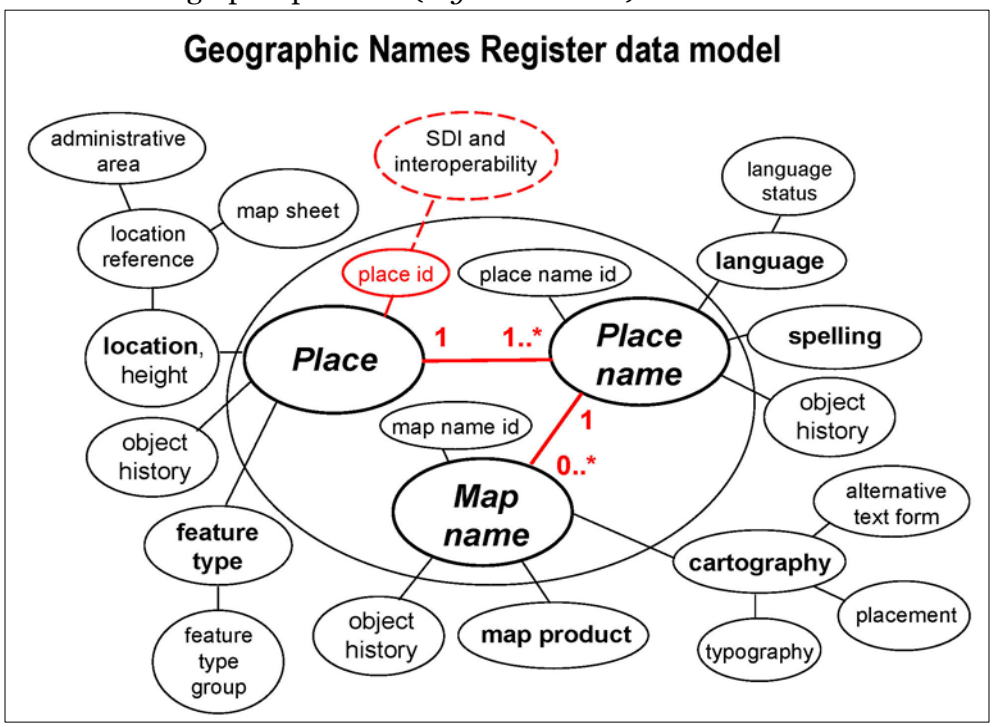

Figure 1. Geographic Names Register data model.

2.3. Place Name Register

The objects Place and Place name build up the Place Name Register. The PNR is not based on any map scale and includes no cartographic information.

The object Place carries the following information:

- Place id, a unique identifier as the link to the related Place name(s) and as a possible link to related spatial objects maintained in external datasets;
- The feature type, and the feature type group via code tables;
- The object co-ordinates, centre point (mouth of a watercourse), in ETRS-TM35FIN co-ordinate system;
- The height above sea level, derived automatically from a digital elevation model;
- The municipality, and other administrative areas via code tables, in which the object is situated, derived automatically;

- The map sheet in different map series in which the object is situated, derived automatically;
- Timestamps for the creation, latest modification and possible deletion of the object (see also *Section 2.5, Object version management*).

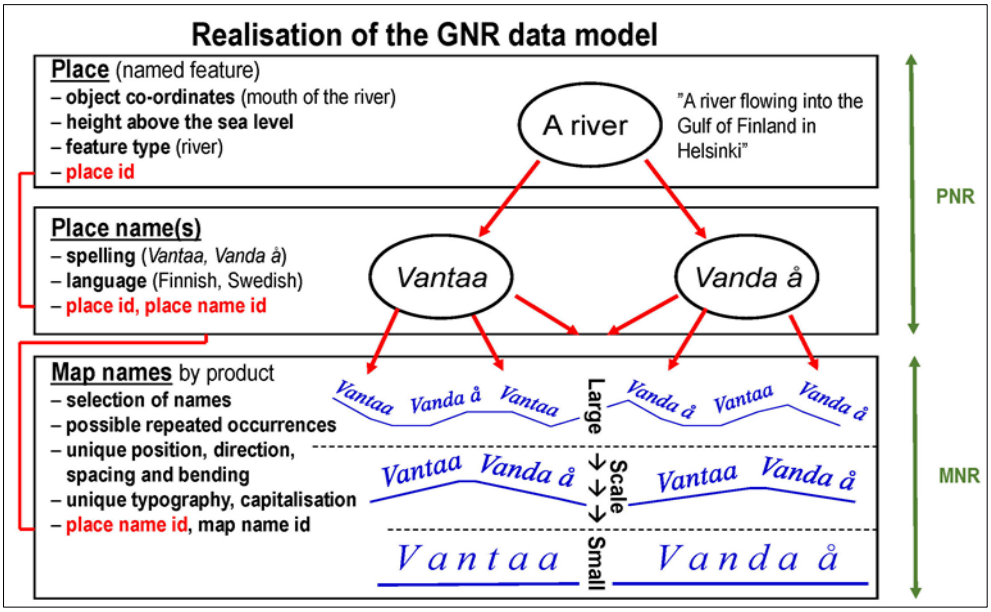

Figure 2. Realisation of the GNR data model.

The PNR feature classification was fully renewed in the GNR reform (*Table 1*). While the reclassification of the Places is in progress, both the old and the new classification are being maintained as attributes.

The object Place name carries the following information:

- Place name id, a unique identifier as the link to the Map names and as a possible link to related linguistic data elements maintained in external datasets;
- Place id as the link to the Place;
- The spelling of the name, unabbreviated, correct capitalisation;
- The language of the name (Finnish, Swedish, Inari Saami, North Saami or Skolt Saami);
- The official status of the language of the name in the municipality in which the Place is situated (monolingual Finnish, monolingual Swedish, bilingual Finnish–Swedish, bilingual Swedish–Finnish, multilingual Finnish–Saami municipality), derived automatically;
- The source of the name;
- The status of the name, e.g. official, standardised, other;
- Timestamps for the creation, latest modification and possible deletion of the object (see also *Section 2.5, Object version management*).

TERRAIN FEATURES

Landforms
- Group of islands or islets
- Island or islet
- Boulder in water or boulders in water
- Cape
- Isthmus
- Shore or waterside
- Group of elevations
- Elevation
- Depression
- Shallows
- Deep

Terrain areas
- Forested area
- Marsh or paludified area
- Field
- Meadow or pasture
- Blockfield or rocky area

Detailed terrain features
- Boulder
- Tree

Other terrain features
- Former dwelling place
- Other terrain feature

HYDROGRAPHIC FEATURES

Water areas
- Baltic sea and its main parts
- Part of sea
- Group of lakes or ponds
- Lake or pond
- Part of lake
- Reservoir
- Part of reservoir
- Canal

Watercourses
- Watercourse
- Part of watercourse
- Rapids

Detailed hydrographic features
- Spring
- Water pit

Other hydrographic features
- Other hydrographic feature

PROTECTED SITES

Nature reserves
(3 feature types)

Other protected sites
(6 feature types)

POPULATED PLACES

Population centres and villages
- City or town or population centre
- Village or neighbourhood

Parts of population centres
- Part of city or town or population centre
- Park
- Market place or square

Houses
- House

Other populated places
- Other populated place

ADMINISTRATIVE UNITS AND PUBLIC SERVICES

Administrative units
(5 feature types)

Administrative buildings and offices
(5 feature types)

Religious buildings
(3 feature types)

Burial services
(2 feature types)

Social and health services
(4 feature types)

Safety and defence units
(14 feature types)

Educational and research units
(6 feature types)

Waste management services
(3 feature types)

Other administrative units and public services
(1 feature type)

TRANSPORT SITES
(17 feature types)

PRODUCTION AND INDUSTRIAL UNITS
(13 feature types)

CULTURAL AND RECREATIONAL SERVICES

Cultural services
(5 feature types)

Sports and exercise sites
(10 feature types)

Recreational services
(11 feature types)

Table 1. Renewed Place Name Register feature classification.

2.4. Map Name Register

The Map Name Register object Map name carries the following information:

- Map name id, a unique identifier;
- Place name id as the link to the Place name (and Place);
- Map product, one of the nine NLS map series;
- Typography code, interpreted via code tables and carrying the following information:

    - Text font;
    - Text size, graphic size in mm/100;
    - Text slant, angle in degrees;
    - Text colour;
- Text position, co-ordinates of the lower left corner of the text, in ETRS-TM35FIN co-ordinate system;
- Text direction, expressed as relative co-ordinates (dx,dy) measured from the text position;
- Text bending, up to 32 pairs of relative co-ordinates conducting a curved text;
- Spacing indicator, whether the text direction or text bending parameters also indicate the length of the text;
- Capitals indicator, whether the spelling of the PNR name is turned into upper case during the plotting;
- Parallel name preference indicator, for cases when feature's two parallel names in one language are presented as map names. If the indicator is activated, a space followed by the Finnish, Swedish, Inari Saami, North Saami or Skolt Saami word for the English "or", according to the language of the respective Place name, is appended to the text during the plotting;
- Alternative text form, for cases when the spelling of the name in the map differs from the spelling of the name in the PNR, e.g. for texts split into two lines;
- Timestamps for the creation, latest modification and possible deletion of the object (see also *Section 2.5, Object version management*).

2.5. Object Version Management

In the GNR reform, a complete version management for Places, Place names and Map names was implemented, based on timestamps for the creation, modification and deletion of the versions and separate database tables for the modified object versions (*Figure 3*). Before any modification of a Place, Place name or Map name, the unaltered object version is inserted into the respective version table, i.e. Place version, Place name version or Map name version, and the version is identified by an incremental version number. The object version management enables temporal cross-sections of the GNR data and complete change-only updates for the GNR data dissemination.

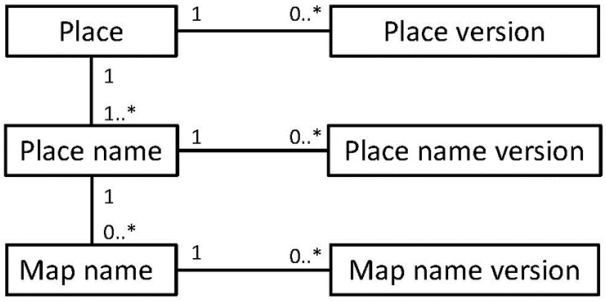

Figure 3. GNR object version management.

2.6. Database and Transaction Management

The Geographic Names Register is an Oracle Spatial database enabling efficient spatial operations. In addition to the dynamic database tables Place, Place name and Map name, and the respective object version tables, the database includes some 50 code tables for e.g. managing the hierarchical feature type, administrative area and map sheet index structures and interpreting and explaining the codes used in the database for the user interface and data dissemination.

The management of absolutely necessary transactions of long duration, i.e. isolation of data changes during the production and the detection and automatic and semi-automatic resolution of possible object conflicts during the merge, is based on Oracle Workspace technology. The GNR hierarchical workspace structure may have one to three temporary workspace levels each of which consisting of one or several workspaces for different kinds of national, regional, local and object-wise production tasks (*Figure 4*).

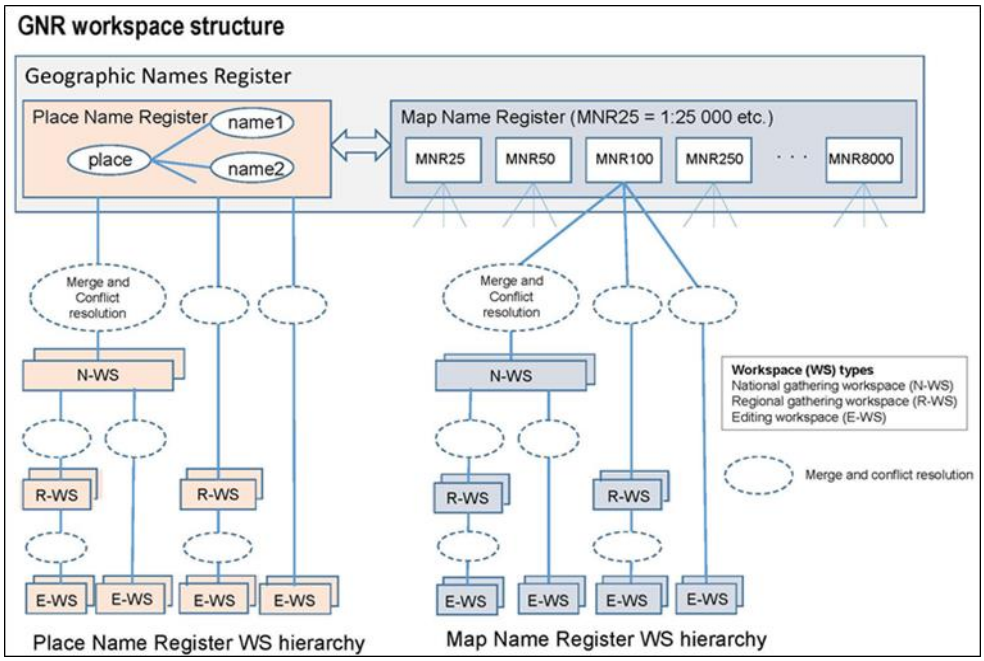

Figure 4. GNR workspace structure.

3. Data Maintenance

3.1. Production Application

The Geographic Names Register production functionality is integrated into the NLS Topographic Database production application based on Smallworld

technology. Thanks to the integration, all the versatile TDB production functionality developed in the NLS for decades is available for the GNR production. Recently added GNR production functionality covers the GNR workspace management, the Place Name Register data maintenance and the Map Name Register data maintenance. The GNR production application user interface consists of specialised tools for the maintenance of the GNR – forms, lists and the map interface for displaying and altering the data. In total there are over 200 daily NLS users using the application and performing GNR transactions.

3.2. Workspace Management

The GNR workspace management user interface is common to PNR and MNR production and includes e.g. the functionality for browsing the workspace structure (*Section 2.6*), creating, refreshing and merging the workspaces as well as the detection and automatic and semi-automatic resolution of possible object conflicts during the merge.

3.3. Place Name Register Maintenance

The Place Name Register production is maintenance of Places and their attributes described in *Section 2.3*. Place names are maintained as attributes of Places, with attributes of their own.

To be able to edit Places and Place names, the user creates a new or opens an existing PNR workspace (*Sections 2.6 and 3.2*). The existing Places to be edited are fetched from the PNR by using the Place Name Register search form. The form allows the user to combine different search terms freely. The search terms for Places and Place names include e.g. the Place id, the location (a polygon, map sheet or administrative area) and the height of the feature, the feature type, the Place name id, the spelling of the name, the language of the name and the status of the name. A time period for the latest modification of Places and related Place names can also be included as a search term.

A PNR search is search for Places and returns a sortable list of Place names with the essential information on both Places and Place names arranged as columns. All parallel Place names are included in the list even if the search terms would match only some of them. For example, a query for the Finnish name and spelling “Inari” returns all parallel names of the municipality i.e. Inari, Enare, Anàr, Aanaar and Aanar as separate rows, with their parallel names as columns by language in turn.

The map interface of the PNR production application includes the background maps, the portrayal of PNR data on the screen, and the geometry tool for maintaining Places’ locations. As to the background maps, both the TDB vector map data and a complete set of NLS raster maps in different

scales are available. In the portrayal of PNR data, the locations of selected Places are displayed as red symbols and the Place names are automatically placed around Places according to the language of the name: Finnish name(s) appear in upper right, Swedish name(s) in upper left, North Saami name(s) in lower right, Inari Saami name(s) in lower left position and Skolt Saami name(s) under the symbol of the Place (*Figure 5*).

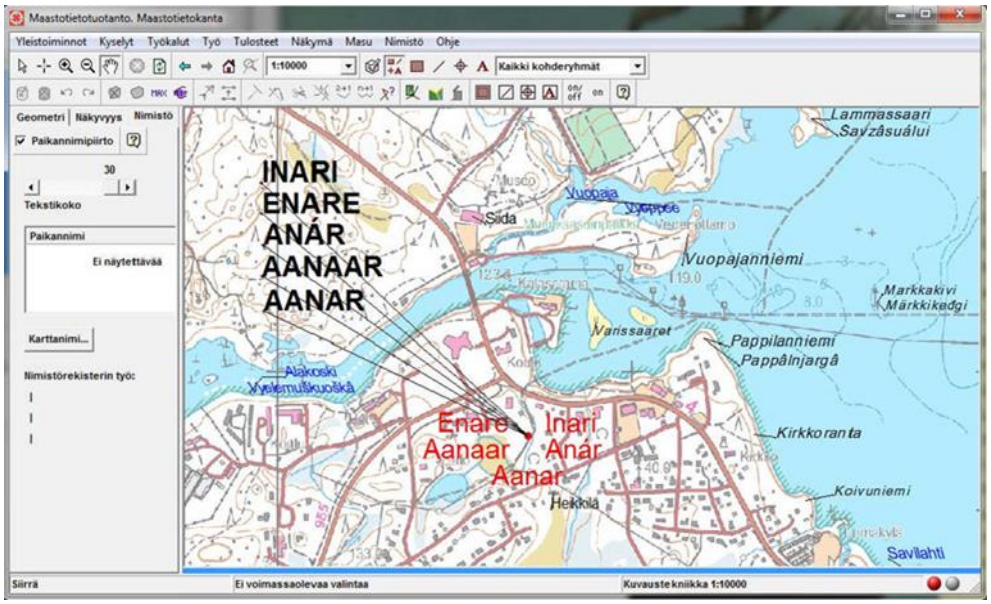

Figure 5. GNR production application map interface. A Place and related Place names are displayed in red, and Map names in 1:25 000 (Basic Map) in black and blue.

The Places and Place names are managed by using the respective object forms and the geometry tool. An existing Place to be edited is selected from the search list or the screen. A Place may be added, modified or deleted, and a Place name may be added to a Place, modified, or withdrawn from a Place by using the Place name form opening from the Place form.

In addition to the object-wise tools, there are three inter-object tools for managing the PNR Places. Two Places with their Place names may be merged as one Place, a Place and its Place names may be split into two Places, and a Place name may be moved from one Place to another. These special operations are necessary in preserving the consistency within the PNR and between the PNR and MNR. There are also special group tools for setting attribute values for several Places at one time.

The version management of the database objects Place and Place name (*Section 2.5*) enables queries for the changes made to the PNR during a given period of time. In addition to the time period, the user may combine different

criteria for the object changes. The search results may include e.g. added or deleted Place names, changes in the spelling, language or status of the Place names, or they may include changes in the location or feature type of the Places. The query can also be restricted by spatial terms for the location of the Places, such as a polygon or map sheet. The query returns a list of current (or deleted) Place names and Places, similar to the search list in normal PNR production. By selecting a row in the list, another list opens containing all combinations of the respective Place and Place name versions in chronological order to be analysed.

3.4. Map Name Register Maintenance

The Map Name Register production is maintenance of Map names and their spatial and other attributes described in *Section 2.4*. Map names are always related to a certain cartographic product, e.g. the Basic Map 1:25 000 or the General Map 1:1 million. The map product to be managed is chosen from a pull-down menu opening from the GNR production application main window.

To be able to edit Map names, the user creates a new or opens an existing MNR workspace (*Sections 2.6 and 3.2*). The existing Map names are fetched for editing by using the Map Name Register search form allowing the user to define the map sheet or map sheets as the search term. As the result of the search, the Map names appear in a list and may be selected to be plotted on the screen. The background map data may be the editable vector data, or a raster map of the cartographic product in question. As to the raster background maps, special versions with no cartographic names in them are available in order to avoid disturbing textual information on the screen.

A Map name may be added, modified or deleted. Names to be added as Map names are always selected from the Place names stored in the Place Name Register, i.e. a Place name must be accepted to the PNR before it can be presented as a Map name in any cartographic product. The Place names to be added as Map names are fetched from the PNR by using the normal PNR search tools described in *Section 3.3*.

Map names are managed by using the respective object form, the geometry tool and the map interface. A Map name to be edited can be selected from the search list or the screen. In addition to the editable cartographic attributes, the Map name object form includes the necessary read-only information of the Place and Place name in question.

The GNR production application geometry tool allows the user to control the placement of the Map name in different ways (*Figure 6*). A Map name may be placed horizontally by using one geometry point for the lower left corner of the text, directed according to two points, or bended according to three or

more points. The text may also be stretched to reach the last point of the given geometry. By using the object form, the name can be split into two lines, or the spelling of the name turned into capitals during the plotting. In the user interface, only a pre-defined set of all existing typography codes (*Section 2.4*) is available for a given Map name: the codes allowed for each combination of a cartographic product and PNR feature type are introduced in the GNR code tables.

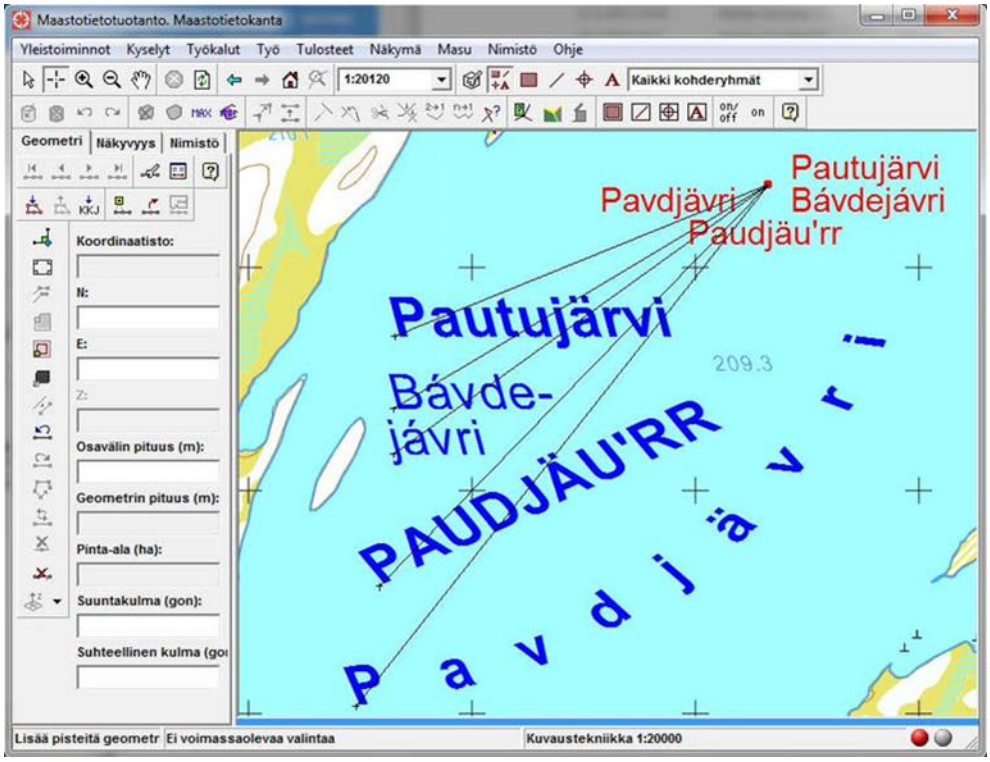

Figure 6. Managing the placement and other attributes of Map names (fictional).

To ensure the consistency between the PNR and MNR, and between the MNR Map name layers of different cartographic products, the user interface offers 13 tools for making sure, for example, that recent changes in the PNR (additions, deletions, certain modifications of Places and Place names) are taken into consideration in map compilation, that all parallel Place names of a PNR Place are presented as map names according to NLS cartographic rules and that Place names of all PNR Places of a given feature type are presented as Map names.

4. Production Lines

4.1. General

The GNR reform covered the PNR and MNR production lines and organisation of work within the NLS, and the procedures for the linguistic treatment of the GNR names data carried out by the Institute for the Languages of Finland and University of Oulu.

4.2. Place Name Register Production

The PNR production is carried out by an NLS team devoted to the task. At first, the team concentrates on the reclassification of the PNR Places according to the renewed feature classification (*Section 2.3 and Table 1*) and complementing the PNR data contents by using different names data sources, including the Names Archive of the Institute for the Languages of Finland and official city and town plans sanctioned and maintained by municipalities.

The quality management of the PNR data and data production covers both linguistic and technical aspects. An overall process for the resolution of reclamations, feedback and change requests concerning the PNR data and the names in national maps has been implemented, with appropriate feedback channels and defined duties, roles and designated persons in charge of the decisions.

4.3. Map Name Register Production

The MNR, containing all Map names presented in different NLS cartographic products, is maintained as a part of normal production and compilation of Basic Maps 1:25 000, Topographic Maps 1:50 000 and small scale maps 1:100 000–1:8 million. The map name compilation may include requests to the PNR production team for the necessary PNR operations, such as the addition of a new Place and/or Place name to be included as a Map name.

4.4. Linguistic Treatment

The linguistic treatment of the GNR names is carried out by the Institute for the Languages of Finland (Finnish and Swedish names) and University of Oulu experts (Saami names). These experts can access and alter the PNR data directly online by using the NLS GNR production application.

The names treatment means e.g. the checking of all additions and changes made to the PNR Place names and possible corrections of the spelling of the names. The locations and feature types of PNR Places are checked on request. The most important information sources for the treatment are the Names Archive of the Institute, recent names field collections, and all the treatment

material accumulated during the continuous and successful co-operation between the NLS and the Institute since the 1950s.

5. Services and Products

5.1. WFS Interfaces

The Geographic Names Register datasets are available through NLS Web Feature Service (WFS) interfaces (National Land Survey of Finland 2015). At present, the national GNR WFS products include two XML schemas for the Place Name Register – with equal data contents but different data structure, serving a little different use cases – and one XML schema for the Map Name Register (*Figure 7*).

The feature-oriented PNR XML profile (Places) has the Place as the GML Feature Member, with the Place name(s) as attributes, while the name-oriented PNR XML profile (Place names) has the Place name as the GML Feature Member, with the Place data and possible parallel Place names as attributes. The MNR XML profile (Map names) has the Map name as the GML Feature Member, with corresponding Place and Place name information as attributes. The XML schemas also include the explanations for the codes used in the service in Finnish, Swedish and English.

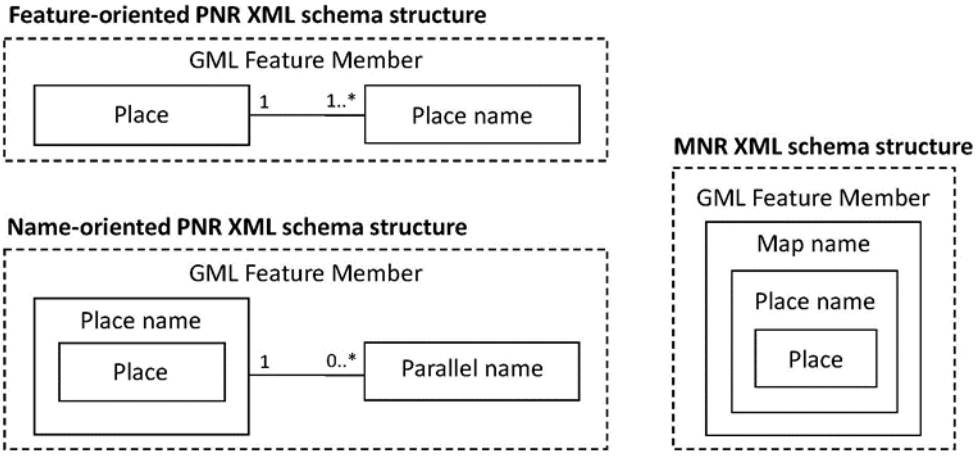

Figure 7. GNR Web Feature Service (WFS) XML schema structures.

A Place Name Register WFS query can be filtered e.g. by:

- Place identifier or Place name identifier;
- The location of the Place, a polygon, buffer area, administrative area or map sheet;
- The height of the Place;
- The feature type or feature type group of the Place;

- The spelling of the Place name;
- The language of the Place name;
- The municipal status of the language of the Place name;
- The 'Relevance at Scale' indicator depicting the size or importance of the Place by means of the smallest scale in which the Place name(s) of the Place have been selected as Map name(s) in national map series;
- Places' and Place names' database lifespan information (timestamps for the creation, latest modification and deletion of the objects).

A Map Name Register WFS query can be filtered e.g. by:

- Map product (e.g. Basic Map 1:25 000 or General Map 1:1 million);
- The location of the Map name (e.g. a bounding box or a polygon).

In addition to the national XML schemas for PNR and MNR WFS products, the geographic names XML schema profile sanctioned by the EU INSPIRE directive (Infrastructure for Spatial Information in the European Community) has been implemented for international use. For object-wise interoperability purposes, the GNR data will soon be available as linked data (Uniform Resource Identifiers (URI) for both PNR Places and Place names), according to the guidelines provided by the Finnish Public Administration Recommendations.

5.2. GML Products

In addition to the WFS, the GNR datasets are available as downloadable GML files (National Land Survey of Finland 2015). The XML schemas of GNR GML data products follow the corresponding GNR WFS XML schemas. In total there are 11 GML data products (downloadable files) with nationwide data coverage – two for the PNR (Places, Place names) and nine for the MNR (Map names 1:25 000 to Map names 1:8 million).

5.3. National Land Survey Data Policy

The National Land Survey of Finland has made its topographic datasets available to the public and to companies to be used freely and free of charge. The open data products are available without compensation and with extensive and permanent rights of use. The datasets can be used for private and commercial purposes, be published, further licensed and attached to other products and services. The NLS open data products and services include e.g. the Topographic Database, GNR names data, elevation models, orthophotos, laser scanning data, small scale map datasets and raster maps in scales 1:5000–1:8 million.

6. Fields of Application

Examples of fields of application of the GNR services and data products are:

- National and international standardisation of geographic names; clear and consistent use of nationally standardised and accurate geographic names in any type of communication;
- Search; finding named places and geographic names by using their attributes (e.g. spelling); map browsing applications; gazetteer services;
- Geocoding; automatic positioning and navigation;
- Interoperability within the national and international spatial data infrastructure, based on unique and persistent Place and Place name identifiers; linked data; geoparsing; ontology;
- Visualisation; map production; geographic names as an information layer in viewing services;
- Research, e.g. onomastics, historical research, natural science, archaeology, genealogy;
- Cultural heritage promotion; safeguarding of the cultural heritage related to inherited geographic names and respective places.

References

National Land Survey of Finland (2015) Data and services, Digital products, Product descriptions, Geographic names http://www.maanmittauslaitos.fi/en/digituotteet/geographic-names. Accessed 10 April 2015