30 June 2017

Original: English

**Eleventh United Nations Conference on the
Standardization of Geographical Names**
New York, 8-17 August 2017
Item 9 b) of the provisional agenda*
National standardization: Office treatment of names

Toponymic data and map production in the Netherlands: from field work to crowd sourcing

Submitted by the Netherlands**

Toponymic data and map production in the Netherlands:
from field work to crowdsourcing

Introduction
Geographical names are an indispensable part of a topographic database or map. Still, in the past decades the collection of toponymic data received less and less attention within the National Mapping Agency of the Netherlands, formerly the Topographic Survey and now Kadaster. As the maintenance of the Dutch topography has been given a legal basis and the topographic databases and maps have been acknowledged as a Key Register, fulfilling the legal task of keeping the geographical names in this register up to date has been increasingly recognized as a problem. Nevertheless, it took some time to find an acceptable and workable solution for this issue. Recently, by means of crowdsourcing most of the costly and time-consuming data collection and verification activities, a new update and maintenance process for toponymic data has started. A pilot project to test the cooperation with volunteers from local historical societies was very successful and is now implemented nationwide. However, it also pointed to the need for good guidelines and regulations for the registration of geographical names. Since the Netherlands have no national names authority, these new activities may give rise to greater interest in the subject of toponymic data standardization within Kadaster and the Netherlands.

How toponymic data collection used to be
At the time when topographic maps were still drawn by hand, necessitating several years to prepare and publish a new map edition, the Topographic Survey, part of the Ministry of Defense, paid much attention to a careful data collection process. This consisted of two phases: the interpretation of aerial images of a map sheet at the office and the conversion of this image to the map model, followed by a field work visit to the area of the map sheet to collect all the information that was not or insufficiently visible on aerial images. It goes without saying that most information regarding geographical names originated from field work activities. Until the 1980s map sheets were completely revised for every new edition, which meant for toponymic data that topographers were obliged to verify all the existing names on the map with the corresponding source. This could be, among others, a municipal employee at a town hall (for the names and numbers of inhabitants of populated places), nameplates on the front of farms or houses (for building names), or local inhabitants (for commonly used field names). Nautical maps of the Hydrographic Service (*Hydrografische Dienst*) were used for offshore water names, as well as waterway and water management maps of Rijkswaterstaat for inland water and polder names. On a 'names model', these sources were registered and for area names the approximate boundaries were indicated (see figure 1).

In the early 1990s, the digitization of the maps led to a new process of updating existing map sheets. This allowed for a shortening in the update cycle, from 8 to 12 years until the 1980s to 4 to 8 years in the 1990s, the exact updating interval of a sheet depending on the importance of the area and the preferences of the army. Names models show less activity on the geographical names from this time onwards, as only changes in the current names were to be registered. In practice, processing changes only was easier for updating the topography, with aerial images as a clear and available source. When it comes to toponymic data, the completeness of the updating process depended more on the insight and quality of the topographer.
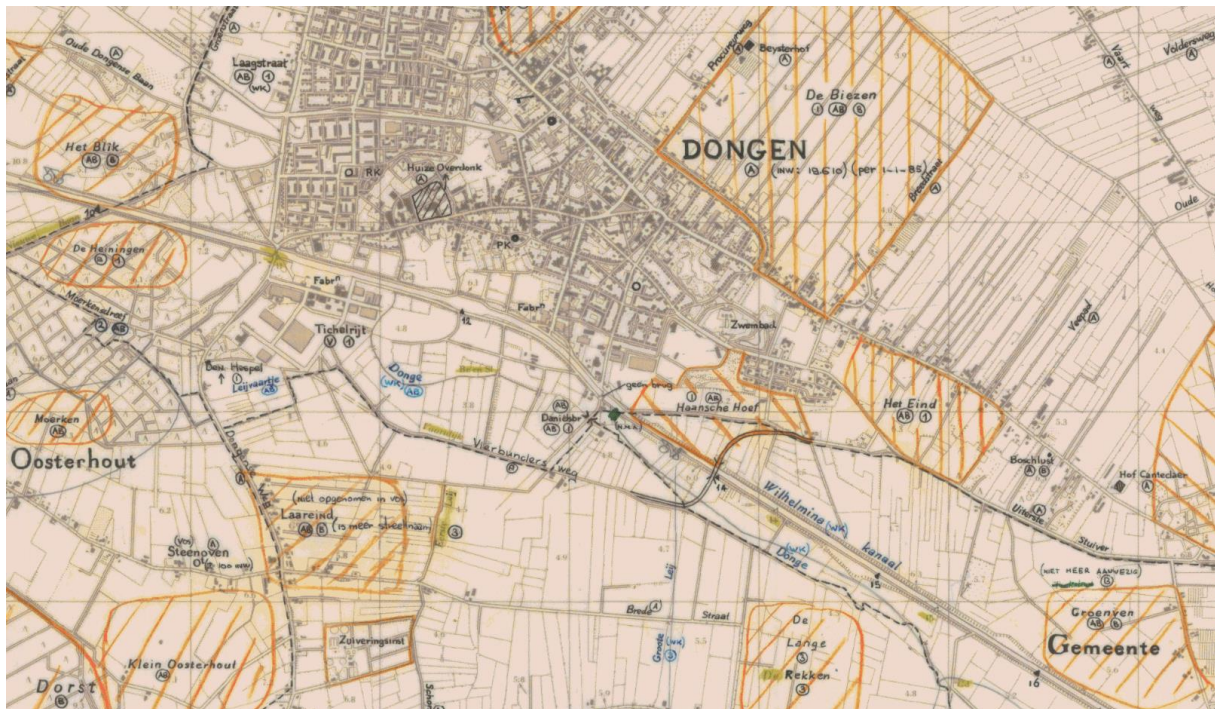
Figure 1 – Detail of a 'names model' from 1985, with circled letter codes or numbers for the sources used and an indication of the approximate boundaries for area names.

New standards and procedures

In 2004 the Topographic Survey was transferred from the Ministry of Defense to Kadaster, the cadastral and land registry agency under the then Ministry of Public Housing, Spatial Planning and Environment. To emphasize the important role of the topographic maps for civil society it was decided to include the topographic maps in a new system of key registers as the Key Register for Topography (*Basisregistratie Topografie*, BRT), with the newly developed feature-oriented vector database TOP10NL as its foundation. The key register system comprises all essential databases for governmental and other public service tasks, such as the registration of persons, companies, buildings, vehicles, cadastral parcels and incomes. In 2008 the Kadaster Act gave the BRT its legal status and made Kadaster responsible for the maintenance of all registered topographic and toponymic data.

As the update frequency of the maps, already in a uniform 4-year cycle for all of the Netherlands after 2000, had to be shortened to a newly agreed 2-year cycle, the costly and time-consuming field work activities came under attack. In 2009 Kadaster signed a contract which made yearly updated 360 degree street view images available for the production process. This allowed topographers to survey the map sheet from behind their desks. Subsequently, all field work was removed from the production process.

For the collection of toponymic data and other attribute information, street view images are no solution, as only part of the names can be found on signs and many of them – e.g. building names on facades – are not legible from the street location where street view images are taken.

Instead, more emphasis has been placed on the use of external data sources for enriching the TOP10NL database. For functional area features, such as cemeteries, industrial areas, allotment complexes, sporting grounds, camping sites, nature reserves and many more, the use of external data in combination with aerial images helped in giving complete coverage for all locations, transferring

geometry from points to polygons, and adding names and type names. For populated places, new built-up area polygons were drawn, based on the location of place-name signs and a set of additional rules, while information on the number of inhabitants within these polygons is retrieved through direct cooperation with the Central Bureau of Statistics (see figure 2). For the names of most populated places, and also for street names, the BRT makes use of the Key Register for Addresses and Buildings, which is maintained by the municipalities. However, many geographical names, such as building names and field names, could not be updated with the existing external data.

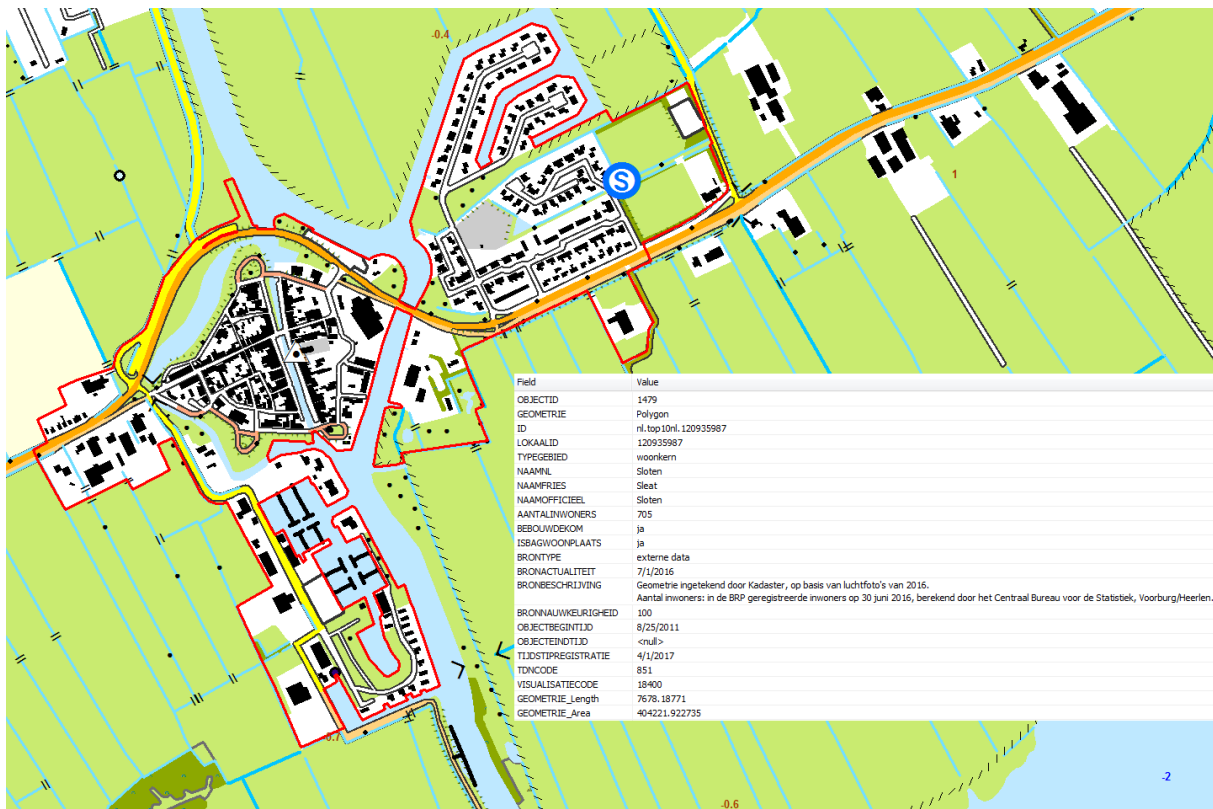| Field	Value
OBJECTID	1479
GEOMETRIE	Polygon
ID	nl.top10nl.120935987
LOKAALID	120935987
TYPEGEBIED	woonkern
NAAMNL	Sloten
NAAMFRIES	Sleat
NAAMOFFICIEEL	Sloten
AANTALINWONERS	705
BEBOUWDEKOM	ja
ISBAGWOONPLAATS	ja
BRONTYPE	externe data
BRONACTUALITEIT	7/1/2016
BRONBESCHRIJVING	Geometrie ingetekend door Kadaster, op basis van luchtfoto's van 2016.
	Aantal inwoners: in de BRP geregistreerde inwoners op 30 juni 2016, berekend door het Centraal Bureau voor de Statistiek, Voorburg/Heerlen.
BRONNAUWKEURIGHEID	100
OBJECTBEGINTIJD	8/25/2011
OBJECTEINDTIJD	<null>
TIJDSTIPREGISTRATIE	4/1/2017
TDNCODE	851
VISUALISATIECODE	18400
GEOMETRIE_Length	7678.18771
GEOMETRIE_Area	404221.922735

Figure 2: Detail of TOP10NL with a visualization of the topography as well as the built-up area polygon and corresponding attribute information for the town of Sloten, in Frisian Sleat.

In 2014, important efficiency improvements were achieved by introducing the LEAN principles in the maintenance process of TOP10NL. 'Wastes' such as waiting times were eliminated and the production process was redesigned, with the marking of changes between the previous and current aerial images and the processing of these changes as key steps. This made it possible, within one year only, to shorten the update cycle from two to one year for all topography in the Netherlands. At the same time, the introduction of automated generalization for the small-scale maps (1:50,000 up to 1:1,000,000) and an automated change detection process for the still manual update of all texts on these maps, including the 1:25,000 map visualization of TOP10NL, made it possible to deliver all BRT products as open data 5 times a year with the same temporal accuracy.

The new LEAN process allowed a reduction in processing time for a single map sheet by more than 70% on average, where 50% was required for a 1-year cycle, thus creating time and capacity for quality improvements and the elimination of the 'blind spots' in the process: data that despite legal obligation are not properly maintained, of which geographical names form the lion's share.

Pilot projects to collect toponymic data

Early 2016, a pilot project was run to seek cooperation with municipalities for checking the names on the 1:25,000 map and delivering relevant changes. Via existing contacts, 9 random municipalities across the Netherlands were approached to participate. They were provided with a special map with all names accentuated, on which the changes in geographical names could be registered, as well as an explanatory letter and a short instruction manual. In the end, only 2 maps were returned and just 1 was quite completely dealt with and useful for processing. One municipality simply responded that there were no changes or corrections for the names on the map for its area, others did not want to cooperate or never responded.

Although this led to some skepticism about the chances to successfully maintain all toponymic data without the meanwhile irrevocable field work activities, a new pilot project was set up in the fall of 2016, in which the role of the municipalities was replaced by local historical societies. Kadaster already gained good experiences in crowdsourcing with the user feedback system of the BRT and it was suggested a 'specialized' crowd of volunteers interested in history and their own living environment would have more knowledge, time and willingness to contribute than the average municipal employee.

This time seven historical societies across the Netherlands were approached by e-mail and later by telephone with the request to cooperate in the maintenance of geographical names on the topographic maps. Six of them responded positive, most even with enthusiasm. One urban historical society refused, considering itself insufficiently knowledgeable about the local toponymy. Instead, the local historical information center was found willing to cooperate.

The societies received a paper map of their working area, with a nameless 1:25,000 background and a selection of the names from TOP10NL as well as a selection of features without a name in TOP10NL visualized over it (see figure 3). Names for which an official source exists, such as street names, were not included. Two societies operating in the bilingual province of Frisia received an additional map with the official, Dutch and Frisian names of all populated places and waters. As in the first pilot project, an explanatory letter and a short instruction manual were added, with the request to check the existing names, add missing names, indicate the boundaries of named area features, and return the map to Kadaster within two months. Four out of seven societies did so, two responded with some delay. The remaining society did not return anything, although it indicated it was still planning to work on the map and return it sometime later.

All maps received contain a large number of changes and additions in the names. One society even added more maps with additional water names and polder names in its working area. An experienced topographer evaluated and processed the changes in a testing environment and considered the vast majority of the contributions useful and – as far as could be observed – flawless. The pilot project was thus considered a success and in early 2017 it was decided to set up a new maintenance process for geographical names based on cooperation with historical societies.

Figure 3: Example of a toponymic map for the former island of Sint-Philipsland, with the named features from TOP10NL in green or black and unnamed features in red, on a nameless 1:25,000 background map

Crowdsourcing toponymic data – a new maintenance process

At the end of May 2017, the process for the crowdsourcing of geographical names was launched. Until the end of 2018, for all of the Netherlands historical societies will be approached to cooperate in checking and updating toponymic data in the BRT. As a preparatory step, an inventory of all historical societies and their working areas has been made, resulting in a nationwide map of historical societies and required alternative sources as well as a corresponding table with their addresses and contact details. It turns out the coverage of the associations across the country is quite complete, leaving only a few gaps. In many areas, there are even multiple societies with an overlapping working area, e.g. a society for a region or municipality next to several societies for individual villages. The latter form a backup in case the first does not cooperate. In villages or areas where there are no historical societies, alternative sources have been registered and will be approached, such as local historical centers, provincial historical societies or village associations.

For about a year, every month a contiguous area of around 60 societies starts to run through the new process, which is maintained and visualized on a process board and a status map. As a first step, an e-mail is sent to the contact person of a society with the request to cooperate, including a short explanation of the purposes and expectations. About a week later, this is further explained and

discussed in a phone call. When a society is willing to cooperate, contact details are checked and registered. The society can choose to receive a digital toponymic map of its working area only or a paper map as well. The layout and contents of the map is similar to the one used in the pilot project. The map is then prepared and sent to the society. After confirmation of receipt, a period of about 2 months is observed in which the volunteers of the historical society would be able to add their comments to the map. Any possible questions are answered and at the request of the society an appointment can be made to discuss possible issues. After a 2-month period, the society is asked for the status of the map and the expected time of completion of the work.

All received maps are registered and made ready for further processing. As a next step, a short sample check is done on the changes and additions marked on the map, insofar this is possible with the available sources, to get an indication of their quality and usefulness. In the rather exceptional case that multiple changes on a map are questionable, further processing is discussed internally. Experienced topographers then use the map to implement the changes in the TOP10NL database. Any questions arising from the comments are registered. To solve these questions and to provide insight in what has been done with their contributions, feedback to the societies is given by sending an updated digital toponymic map of the area as well as any additional questions. Answers to the questions and any last additional comments are then implemented and the updating of names in TOP10NL is completed. Subsequently, with an automated change detection script all relevant changes for the texts on the 1:25,000 map are obtained and considered for placement. In case a society provides more names in an area than can be accounted for in view of a desirable names density, only a selection of the names will be placed. Excess names will be stored in the annotation database for possible use in the future. On the smaller scaled maps, 1:50,000 up to 1:1,000,000, the TOP25 names are stepwise further diluted as usual.

After the updating of all BRT products is completed, in a separate step the quality of the toponymic data in TOP10NL is assessed. Next to the toponymic maps with the comments provided by historical societies, other sources are used as well to check the quality of specific categories of names, such as the BAG for street names and nautical maps of the Hydrographic Service for offshore water names. If the results of the quality assessment give cause for this, additional actions may be scheduled to improve the quality of one or more toponymic data categories.

The current intention is to continue the maintenance process in cooperation with historical societies after the update cycle for geographical names is completed at the end of 2018, although the number of changes and the amount of time and effort needed to implement them will be much smaller.

Need for geographical names standardization
Like every map production process, an update and maintenance process for toponymic data requires a well-defined and up to date set of rules and regulations. The internal document with regulations concerning geographical names was last updated in 2007, even before the BRT with the TOP10NL database was implemented, so it had to be urgently adjusted to the changed circumstances and the newly developed process.

In the absence of a national names authority in the Netherlands, the Topographic Survey developed its own regulations for a uniform registration and visualization of geographical names. These include rules on the use of capital letters, the combination of words, the use of hyphens, the construction of adjectives from geographical names, as well as rules for specific types of names such as building names. Concerning the orthography of the names, the 1947 Dutch Orthography Act prescribed the

continuous use of the 19th century De Vries and Te Winkel spelling for domestic geographical names, pending governmental regulations dealing with these names to be developed later. This old orthography thus remained dominant on the topographic maps.

The governmental regulations announced in 1947 were never promulgated and in 2006 a new Spelling Act replaced the previous act from 1947, this time without excluding its validity for the orthography of geographical names. This means that the current orthography rules now also apply to geographical names in the Netherlands.

Meanwhile, the Key Register for Addresses and Buildings (BAG) which is maintained by the municipalities, gives legal status to the names of streets and roads - called 'public spaces' in the BAG - and the names of populated places, as determined by municipal councils. These names are necessary in addresses. The BAG data model provides the opportunity to register other types of 'public spaces' as well, although few municipalities use this option. Other registers in the key register system, such as the BRT, re-use the names from the BAG.

Unfortunately, at the introduction of the BAG in 2009, no standardization rules had been established on the registration of the names. This has led to many different registration practices between the municipalities. In the case of populated place names that occur more than once in the Netherlands, the postal services used to add an abbreviation to the name to make the name unique, e.g. "Bergen NH" and "Bergen L" for two villages with the name of Bergen in the province of North Holland or Limburg, respectively. Some municipalities decided to register these additional abbreviations in the BAG, others did not. This resulted in registered names such as "Spaarndam" next to "Spaarndam gem. Haarlem" (for the section of the village in the municipality of Haarlem) and "Bergen (NH)" with brackets next to "Bergen L" without brackets. In the BRT it was decided to leave these abbreviations out. Even larger differences occur in the registration of street names. For example, title names such as "Burgemeester" (mayor) are sometimes abbreviated and sometimes not, while the abbreviation is sometimes indicated with a full stop ("Burg.") and sometimes not ("Burg "). In some street names even the space between words is absent or a name strangely ends with a superfluous full stop.

These examples indicate the urgent need for generally accepted and applied standardization rules for geographical names in the Netherlands. This can best be done through cooperation between the various registry holders, such as Kadaster and the municipalities. Ideally, this cooperation results in mutually agreed standards for geographical names and the creation of a commission responsible for the maintenance of these standards.

Doing so, the Netherlands would finally comply with the 4th resolution of UNGEGN adopted in 1967 recommending to establish a national geographical names authority and standardize the geographical names.