CROWDSOURCING AND VGI IN NATIONAL MAPPING AGENCY'S DATA COLLECTION

Daphne Bol, Magdalena Grus, Mari Laakso

Msc, Daphne Bol;

The Netherlands' Cadastre, Land Registry and Mapping Agency (Kadaster);
Hofstraat 110 7311 KZ Apeldoorn;
telephone: 00316 52 48 19 02;

e-mail: daphne.bol@kadaster.nl

Msc, Magdalena Grus;

The Netherlands' Cadastre, Land Registry and Mapping Agency (Kadaster);
Hofstraat 110 7311 KZ Apeldoorn;
telephone: 00316 52 06 23 65;

e-mail: magdalena.grus@kadaster.nl

D.Sc.(Tech.) Mari Laakso;

Finnish Geospatial Research Institute, National Land Survey;
Geodeetinrinne 2, 02430 Masala;
telephone: 00358 40 01 00 707;
e-mail: mari.laakso@nls.fi

Abstract
In the recent years several Nationaal Mapping Agencies have initiated projects to explore citizens' participation to improve the quality of the topographic data. One of the conclusions was that the data coming from crowdsourcing initiatives is not always uniform and countrywide. The question is: how to collect uniform and complete countrywide data that can be used to enrich the Topographical Key Registers in a sustainable way? In a new project 'Crowdsourcing at school!' the dutch Kadaster explores the possibility to collect uniform data with a national coverage by means of curriculum for children in primary schools. At the same time children practice their geographical orientation and their role in the society as an active contributor to the governmental data. The results of this project help finding answer for the optimal usability of the crowdsourced data and its applicability to Topographical Key Register in different countires.

Keywords: *crowdsourcing, children, data, national coverage, Topgrapgical Key Register*

TOPOGRAPHICAL KEY REGISTER (BRT) AND CUSTOMER VALUE

The Dutch Cadastre, Land Registry and Mapping Agency (in short Kadaster) is a non-departmental public body, operating under the political responsibility of the Minister of Infrastructure and Environment. One of its statutory tasks is to maintain the key registers. One of the Key Registers is the Topographical Key Register (BRT), which consists of digital topographic data sets at different map scales (1:10k, base data set and derived / generalized data sets at scale 1:50k, 1:100k, 1:250k, 1:500k and 1:1.000k). The data quality of the Topographical Key Register is internally and externally controlled according to ISO19113 standards. The controlled elements are: logical consistency, positional accuracy, thematic accuracy, actuality, completeness and the feedback system. In 2012 Topographical Key Register became an open data under the CC-BY licence. Since then all the products of the BRT family have became freely available to everyone for use and republishing without restrictions such as copyright and patents (te Winkel, 2015). Some signifficant effects of making the BRT data open concerning the customers are increasingly visible. One of those effects is a clear shift in user groups of BRT. The percentage of business and private users has increased significantly

in the first few years (Fig 1). Another visible effect is a grow of the BRT user community. Since BRT has been made available as open data the number of BRT LinkedIn user group members has increased substantially. Also the group activity represented by discussions and comments has increased significantly (Bregt et al.,2015). Next to making BRT products open, the implementation of LEAN method in BRT production process helps to increase the role and inmportance of the users. The core element of the LEAN methodology is to maximize customer value while minimizing waste. Thanks to LEAN method the focus has shifted to the direct creation of value for the BRT customer. In the Dutch Kadaster next to the exclusion of the waste in the production process also the customer was asked to participate actively in improving the production process. Every change in the specifications of the BRT product family or in the production process has been evaluated by the customers from their perspective. From the BRT user group consultation it has been found out that the actuality of the map is the most important and desired quality element of Topographical Key Register BRT.

Figure 1

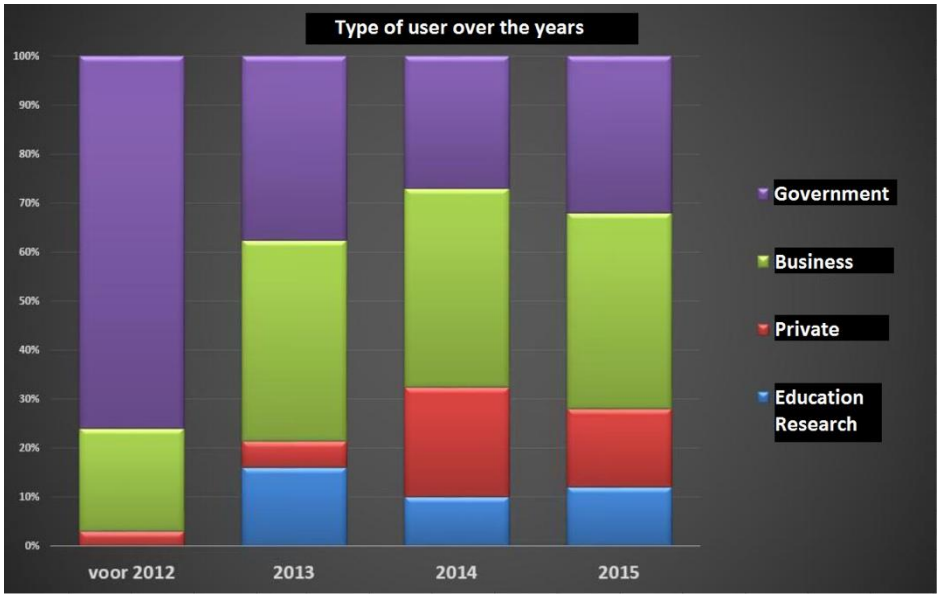

Figure 1. The users of the BRT separated in user groups
(source:"De effecten van een open basisregistratie topografie (BRT)", Bregt et al.,2015)

CROWDSOURCING AND VOLUNTARY GEOGRAFIC INFORMATION (VGI) IN THE NETHERLANDS

Since 2012 Kadster has been facing a growing need for the higher actuality of topographic data. Next to the LEAN method implementation and giving the customer a prominent role, Kadaster has initiated several projects to explore citizens' participation (crowdsourcing en voluntary geographic information) to improve the actuality of the topographical objects. Next to the high quality standards the Dutch topographic datasets has to be uniform, nationwide and deliverable each year. An important user group of our topographical maps is the emergency service sector. Unfortunately the current situation does not meet the quality and visualisation requirements of this group. The emergency services demand the highest actuality and uniform data with a national coverage. The question arised whether crowdsourcing would be a suitable method to collect those data. Althouh the crowdsourced data have very high actuality, they are very often not uniform nor countrywide. Moreover there is no guarantee about the sustainability of such initiatives.

CROWDSOURCING AND VOLUNTARY GEOGRAFIC INFORMATION (VGI) IN FINLAND

Crowdsourcing and VGI concepts are being implemented and tested also by the Finnish Geospatial Research Institute (FGI), part of National Land Survey (NLS) in Masala. As part of the universal renewal of the National Topographic Database schema a project has been initiated to investigate to possibilities of crowdsourcing in data collection. The main objectives of the project are to create a concept for the usage of the volunteered geographic information and to test the concept with a pilot. The plans are to offer citizens and for other partner bodies as well a special layer on top of the authoritative data where they can contribute their data on. This layer is called as Citizen Topographic Database. These

data would be visible for all the users. The main challenge in data quality is to create an interface which guides the contributor to enter as complete data as possible. The current pilot design will focus on selected data objects to which complete data models are created and the required attributes are decided. The pilot environment will be offered to the audience in web browser based and as a mobile version. In addition to the technical challenges in collecting and storing the data the conceptualizing of promoting, engaging and motivating people is highly important. Whilst campaigns on special topics work often well the continuation and sustainability of the activity is difficult to gain.

Based on the results from the previous crowdsourcing projects at the Kadaster and NLS but also experiences from other National Mapping Agencies it can be concluded that maintaining continous participants' involvement is challenging. The question which arise was: how to collect actual, uniform and countrywide data, that can be used to enrich the Topographical Key Register users, in a sustainable way?
The two organizations – the Dutch Kadaster and Finnish FGI – are a collaborating partners in the field of crowdsourcing and VGI within the Short-term scientific mission (STSM) for European COST Action TD1202 research network.

CROWDSOURCING AS A PART OF EDUCATIONAL PROGRAM

To find the answer for the question how to collect actual, uniform and countrywide data, that can be used to enrich the Topographical Key Register users in a sustainable way, the Dutch Kadaster started a new project 'Crowdsourcing at school!'. In this project Kadaster has been exploring the possibility to collect data with a national coverage by means of curriculum for school children. Involving the Dutch educational system seems to be a suitable approach because of the amount of and location of primary and secondary schools in the Netherlands which are spread evenly throughout the whole country (Figure 2). At the same time children get a chance to practice their geographical orientation, acquaintance with topographic maps and their role in the society as an active contributor to governmental data.

Figure 2

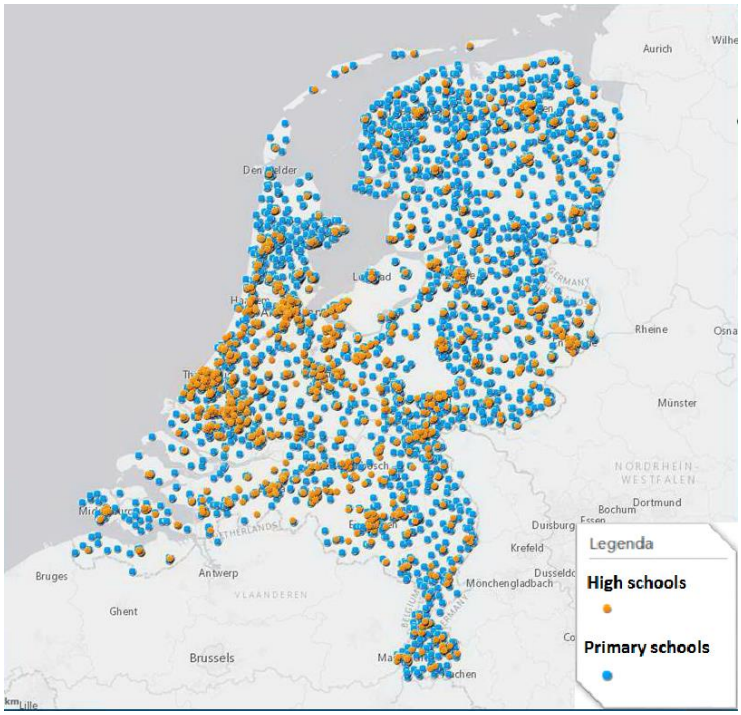

Figure 2. The spread of the primary and secondary schools (high schools) in the Netherlands.

In the crowdsourcing curriculum presented by Kadaster to schools in a pilot project, children get to know what is crowdsourcing and why their contribution can be important for the whole society. The curriculum is based on an onlineprogram and involves several parts.The first part is about the Kadaster as organization. Through several assignments, children gain knowledge about the history of land registration, and the role of Kadaster. The second part is about crowdsourcing. In the theorethical part children learn about the concept of crowdsourcing and get several examples of crowdsourcing initiatievs such as Wikipedia, OpenStreetMap etc. In the practical part of this curriculum children learn to collected data for the emergency services (police, ambulance and fire services) through an online application (Figure 3).

Figure 3

Figure 3. Online application to collect data for emergency services

In the future, the curriculum can be extended or adjusted to the needs of other target groups. By integrating this crowdsourcing method into the educational system, data collection is ensured for a longer period. By means of application, children work on their topographical knowledge by using different kinds of maps (topographical maps and aerial photos). By collecting data for emergency services they learn how they can contribute in order to help other people.

Currently, this curriculum is in the pilot phase. We are testing the online program at primary schools (children in the age of 11-12). In this phase we want to gain as much as possible experiences and feedback from pupils and from the teachers about the content, format and the user friendliness of the method.

Further we are also curious about the quality of the collected data and the usability of the new objects by the target group which is the emergency services. However further research will be made in cooperation with Finnish NLS and the French National Mapping Agency (*IGN-FRANCE*).

School visits

Until now the curriculum has been tested at four primary schools in the Netherlands. The lesson have been prepared, coordinated and given by two kadaster collegues. The lesson took one hour and consisted of two assignments. The first assignment was about the Kadaster organisation and its role in general. The second part was about crowdsourcing and data collection for the emergency services. In an online application the data were collected such as: obstacles on the road, bus stations, petrolstations, etc.. Three different methods were tested for datacolletion:

- In the first method, the whole class/group children were looking for the new objects. There was only one child at a time physically interacting with the online application. The other children and the teacher were actively helping and supervising.
- In de second approach, the children were divided in groups of 2-3 children. Each group had the same application running on the computer. Each new added object by one group was directly visible for the rest of the groups.
- The third approach was very similar to the second one except that each group had its own version of the application running on the screen. The new added objects were not visible by all groups at the same time.

Observations and first results

The first results which are based on all three methods seem to be very promising. The children enthousiastically and with involvement were looking for new objects. We have noticed that the idea of helping others make the pupils very motivated. Children were asking questions about the possibility to do this exercise also at home.

In this stadium it is too early to say which of those three methods is the most suitable to collect new objects for the emergency services. Each of the three presented approaches has advantages and disadvantages. For example by means of the first method, the accuracy and correctness of the new added object was very high however there were (only) five new objects added (in the class of 32 pupils). We have found out that the supervision of the teacher and the rest of the group helps in adding correct objects and eliminate potential errors. In this approach the group gets unintentionally also the role of validator.

The second method resulted in many new objects (more than in method one and less than in method three). Around 60% of them were correct. There were also "joke" – objects added. For example someone has put a school building in the lake. During the test there were serveral questions from the children about the possibility to erase an incorrect object added by another group. The second approach did provide this possibility.

The third method resulted in the highest amount of the objects added comparing with the other two methods, however many of the added objects related to the same location. This is because children did not see the object added by other groups. On the other hand the high amount of objects pointing to the same location can be treated as a sort of validation. For example the objects mentioned by at least 8 of 10 groups can be named as very highly reliable, 6 - 8 of 10 high reliable, 2-6 of 10 medium reliable (where extra control/validation is needed), less than 2 not reliable and the object can be neglected.

The choice of the method can depend on many factors. It can depend on the target group for which we plan to collect data and the aim we want to achieve. It can also depend on internal process and possibilities of the organisation. If the aim is to make the children familiar with the phenomenon of crowdsourcing, the sort of method to collect data is of less importance. If the goal is to collect high quality data than the method can have a crucial role. In case we need a high quality data without extra validation procedure than the first method seems to be the most suitable. If the aim is to collect as many data possible than the method 2 and 3 seems to be a better choice.

The first results are based on data collected from 5 groups (together around 120 pupils). The research focused on investigating the children's experience with the subject: crowdsourcing and the diferent method of data collection. In the second phase we would like to concentrate more on the quality of the data collected and its usability for the emergency services.

CONCLUSIONS

Crowdsourcing and the ongoing search for new methods for collecting uniform and actual data with a national coverage is an important issue in several Mapping Agencies in many countries. The crowdsourcing as a part of the educational program seems to be a very promising development. By means of curriculum, children gain topographical knowledge and learn how to became an important member of de society. Bringing topographic mapping to the school curriculum would guarantee a wide spared of the knowledge in the society and build a solid ground for further mapping activities among the citizens in future.

Based on the first results of the research it seems that school children can play an important role in data collection and can become active contributors to governmental data. Considering the quality and usability of the data delivered by the children, there is more research needed.

We would like to implement this approach in other countries as well and further explore the chances and obstacles of the crowdsourced data collected by pupils.

REFERENCES

Bregt, A.K., Castelein, W., Grus, Ł., 2015. De effecten van een open basisregistratie topografie (BRT). Wageningen University, Wageningen, The Netherlands.

Bregt, A.K., Grus, Ł., Eertink, D., 2014. Wat zijn de effecten van een open basisregistratie topografie na twee jaar. Wageningen University, Wageningen, The Netherlands.

Te Winkel, D., 2015. International Workshop on Spatial Data and Map Quality "The Dutch Key register Topography as open data", Malta, January 2015.

BIOGRAPHY

Daphne Bol

Daphne is enthusiastic, driven, innovator, interested in geo related issues and environmentally concerned. She is originally from the Netherlands where she also lives and works. Daphne studied at the University of Amsterdam in the field of Urban Planning and at Wageningen University in the field of Urban Environmental Management and Geoinformation Sciences. After her graduation in 2014 she started working for the Dutch Kadaster as a trainee. In this position she is involved in several projects throughout the organisation. Especially geo related projects are of Daphnes' interest. In the project 'crowdsourcing for schools!' she could combine all her creativity, geospatial knowledge and enthusiasm in one project. The project characteristics together with the great partnership with Magdalena Grus worked really well for her.

Magdalena Grus

Magdalena Grus is a consultant and innovator for the topographical products and process in The Netherlands' Cadastre, Land Registry and Mapping Agency (Kadaster). For the last years her passion and subject of the intrest has been crowdsourcing and its applications for the governmental datasets. Magdalena studied Land Use Planning and Management at Warsaw University of Life Sciences and Social Spatial analyses at Wageningen University in the Netherlands. After graduation Magdalena worked for geo related companies and for the last 6 years for the Dutch Kadaster. She has been involved in different VGI- related projects such as assessment of the volunteered geographic information feedback system and the new ongoing project: 'crowdsourcing for schools!'.

Mari Laakso

D.Sc. (Tech) Mari Laakso acts as a Senior Research Scientist at the Dept. of Geoinformatics and Cartography in Finnish Geospatial Research Institute (FGI) by the National Land Survey of Finland. She has been involved in studies on how authorative mapping agencies as well as other governemental bodies could benefit from crowdsourced geographic information. Her current research interests are in usage of volunteered geographic information and location-based services to enhance the sense of community and increase the physical and social activity of people, the new ways of communication geospatial information and engaging school children to crowdsourcing and topographic mapping.