

VGI FOR LAND ADMINISTRATION – A QUALITY PERSPECTIVE

G. Navratil^{*}, A.U. Frank

Department of Geodesy and Geoinformation, Vienna University of Technology, Gusshausstr. 27-29, Vienna, Austria -
(navratil, frank)@geoinfo.tuwien.ac.at

KEY WORDS: VGI, Land Administration, Quality

ABSTRACT:

Since the use of volunteered geographic information (VGI) or crowd sourced data (Goodchild, 2007) became more common, several people proposed the use of such methods of data collection for various fields. Success stories were Wikipedia encyclopaedia and OpenStreetMap (OSM), but also using VGI in land administration has been proposed. Robin McLaren proposed crowd sourcing as a way to get a new citizen collaboration model in land administration to enhance transparency and decrease costs (McLaren, 2011). Keenja et al. discussed the perception of VGI within the Dutch cadastre (Keenja, De Vries, Bennet, & Laarakker, 2012). Basiouka and Potsiou even discuss how crowd sourcing can be used to identify errors in the Hellenic cadastre (Basiouka & Potsiou, 2012). One problem of VGI is the quality control (compare Goodchild & Li, 2012). The problem with most data in a land administration system is that there is only a small group of people that can verify the correctness of information. The correct location of a boundary, for example, can only be assessed by the owners of the pieces of land touching at the boundary (and surveyors after investigation and measurement). How shall VGI then provide reliable data? Boundaries between areas of different use may be visible but land administration is often interested in ownership boundaries. In the paper we discuss the types of data used in land administration as discussed by Dale and McLaughlin (1999). These categories are then analyzed to identify the areas where VGI can actually provide reliable input. What we hope to learn from such an analysis is how to use the methodology of crowd sourcing for land administration, even if the data collection authoritatively.

1. INTRODUCTION

Volunteered Geographic Information (VGI) is a new phenomenon (Goodchild, 2007), enabled by the increasing capabilities of Web 2.0. There are success stories for the application of VGI, e.g., the Wikipedia encyclopaedia or OpenStreetMap (OSM). Encouraged by these examples, additional fields of application have been proposed. One of them is land administration (e.g., Basiouka & Potsiou, 2012; Keenja et al., 2012; McLaren, 2011). However, there is a fundamental difference between an encyclopaedia like Wikipedia and a subsystem of public administration: Errors in the encyclopaedia are annoying but can be corrected quite easily when detected. Errors in public administration, however, may have dramatic consequences. A misspelled name in a land register may provide the legal basis for a cheater to sell land he does not own. As a result the rightful owner loses his land or the faithful buyer loses his money even if none of them did anything wrong. Thus, data in public administration must be as reliable as possible and must pass strict checks. Goodchild and Li presented methods to check the quality of VGI (Goodchild & Li, 2012). In this paper we discuss, which of these methods are applicable to land administration and which for parts of land administration VGI may be an option.

2. LAND ADMINISTRATION

Land administration is an integral part of public administration. "The careful management of land and property is fundamental to economic development and the sustainability of the environment ..." (Dale & McLaughlin, 1999, p. 2) Land administration is related to various public activities like land

tax, land market, land use planning, the protection of rights on land etc. It usually consists of legal, environmental, and geometrical information. Legal information is centred on ownership and other land related rights including mortgages or rights of way. Environmental information may contain aspects like quality of soil, pollution, land cover, noise etc. Geometrical information is the subdivision of land into small pieces (called parcels) as a basis to store legal information. For each parcel, at least size, shape, and extent are defined.

The topical separation is usually reflected by an organizational separation as well. Civil law rules are applied to most legal information. Public law restrictions, e.g., development rights, are the exception since they are granted by administrative bodies. Non-legal information, however, is always maintained under public law regulations. Thus, in many jurisdictions the legal information is handled by courts and the remainder by one or more administrative bodies.

Legal information on rights to land has two important aspects:

1. It always describes a sharp boundary. The right of ownership of one person ends where the right of ownership of his neighbour begins. It may not be clear, however, where the boundary line is located exactly.
2. The rights described are invisible in the real world. When identifying the area owned by one person we usually have to refer to visual demarcations like walls, fences, or differences in the use of the land. However, these do not necessarily exist. Fig. 1 shows an example for such a situation. The orthophoto on the left shows a large, homogeneous area. It could be assumed that, since it seems to be used as one piece, it

* Corresponding author.

is owned by one person. However, the cadastral data on the right reveals, that the area consists of four different parcels. Each of them may have a different owner and one person rented them and now uses them collectively.



Figure 1: Observation of an area (left, orthophoto) and cadastral data of the same area (right)

The geometrical information depends heavily on the legal information. A description of a boundary line can only be created if the spatial extent of the ownership right is known. Once the boundary is described, the size of the parcel is easy to define although different methods must be used depending on the type of boundary description (compare Navratil, 2011).

Environmental information is necessary to use land efficiently and to define necessary restrictions for environmental sustainability. It is used in administrative processes, for example:

- Land tax: The taxation of land ownership should be based on fair and transparent rules. There are two different approaches used. One approach assesses land tax based on an approximate market value of the parcel. This approach requires a system of mass valuation to keep the values updated. Mass valuation uses a variety of data including sales prices, object descriptions, and data describing the nearby infrastructure (Brunauer, Feilmayr, & Wagner, 2012). Another possibility is using the benefits produced by the land. This is well-suited for agricultural areas where the quality of the soil determines the production capability of the parcel.
- Land use planning: Making optimal use of land in densely populated areas is necessary in order to restrict urban sprawl and protect areas used for food production. The typical instrument for this is land use planning. The result of land use planning is a map defining the spatial distribution of processes. The creation of residential building, for example, is restricted to residential areas thereby protecting agricultural areas and securing the food production capability. The basis for land use planning is information on existing objects and their use.
- (Agricultural) subsidies: Some parts of a country may provide more difficult living conditions than other parts. Farming, for example, is easier in flat terrain with fertile soils than in mountainous regions. However, mountainous regions should not remain unused. The productivity will be lower, though, than in better areas and subsidies can provide an incentive to utilize such areas as well. These subsidies should be fair and thus the local conditions have to be taken under consideration. This requires information on the

relevant factors for productivity, i.e. soil quality, gradient, amount of precipitation, etc.

3. VOLUNTEERED GEOGRAPHIC INFORMATION

VGI are collected by a large number of private citizens (Goodchild, 2007). These citizens did not have specific training for the collection process. As a result, the tools used for the data collection and integration must be easy to use, error resistant, and transparent. A simple example for this is the Wikipedia encyclopaedia. Once the text for a specific entry is formulated, the majority of the work is done. Uploading the text to the encyclopaedia and linking the text to other entries is done quickly. Thus it is the topical knowledge of the citizens that is collected.

Other VGI initiatives concentrate on collecting observations. OSM collects the location and the type of transportation infrastructure while platforms like Flickr and Picasa provide access to georeferenced images and videos.

VGI initiatives are only successful if several components work together. Some important components are

- a large group of people providing new data, check, extend, and (if necessary) correct existing data
- tools, which allow people to perform their editing tasks with as little additional knowledge as possible,
- methods to check the quality of the entered data, and
- simple methods to reuse the data.

Goodchild lists a number of possible reasons why people contribute to VGI (Goodchild, 2007): self-promotion, personal satisfaction, and altruism. Currently the number of users, e.g., of OSM is still growing rapidly. However, the percentage of actively contributing users decreased from an all-time high of almost 14% in 2007 and 2008 to 2% in December 2012 (OpenStreetMap Foundation, 2013); the interpretation should take into account that the number of users has naturally grown much faster than the number of contributors and the observation is not necessarily a reason for concern.

4. CHECKING ENTERED DATA

Quality of VGI is a widely discussed topic. Numerous examples of quality checks for different kinds of data sets can be found in the literature. Goodchild and Li tried to categorize the approaches. They distinguish between the following methods (Goodchild & Li, 2012):

1. The crowd-sourcing approach: Data may be correct if a large group of people agrees on it. The principle has been successfully applied in the Wikipedia encyclopaedia. It does not work equally well on all kinds of geographic data collection, since "it will apply best to geographic facts that are prominent" (Goodchild & Li, 2012, p. 113) and therefore many can observe and report.
2. The social approach: Trusted users act as gatekeepers for information entered by other users. These trusted users can, for example, be users that contribute heavily and thus have more experience than others. These users can easily be distinguished as Mooney and Corcoran, for example, reported that 11% of the contributors were responsible for 87% of the changes on heavily edited (over 15 times) data in OSM (Mooney & Corcoran, 2012). Nevertheless, users that

initiate many changes may have vested interests to do so and are therefore not automatically the most objective observers.

3. The geographic approach: The provided information is compared to existing geographic knowledge. Motels, for example, should be close to a road and water does not flow uphill. Data which are not plausible must be flagged and investigated.

How does this work for data used in land administration? It is possible for laypersons to use a GNSS receiver and determine the position of boundary indicators like walls, fences, etc. They can upload the data to a database if there is a suitable web-interface that helps them through this process. OSM is an example for such a system. However, there are two major differences between OSM and land administration data:

- Precision of the coordinates
- Relevance of the visible objects

OSM aims at providing a map of the road network. The scale for such a map is usually at least 1:10000 and there is a strong relation between the quality of the observation and the scale of the data (Frank, 2009). A positional uncertainty of 5m is acceptable for OSM because in most cases the polygon will still be located within the relevant object it represents (the road). The same cannot be said about boundaries needed for land ownership conflicts. In densely populated areas and uncertainty in the range of metres is not acceptable because it would either lead to unused space or legal insecurity. High precision in the range of centimetres can be reached using GNSS devices. However, this requires sophisticated observation strategies and these may be difficult for untrained people. Although providing training would be possible, the time investment could limit the possible number of contributors. In addition, the equipment providing the necessary quality is also more expensive, which also restricts the number of available sensors. In the following, possibilities to test the quality of VGI for the different kinds of data required for land administration will be analyzed.

4.1 VGI on Rights

Laypersons can only document visible objects. Walls, fences, and other visible objects are indicators for (ownership) boundaries, but they do not necessarily coincide with them. Rights themselves are not observable. What does the observation mean that somebody is living in a house? The person may be owner of the land (including the house), may have rented the house, or be the guest of the owner or the renter. These results in two different legal situations: Right of ownership and right of use. It is impossible to distinguish between the cases by observation of behaviour only. Thus additional knowledge must be used to separate these cases. There are two possibilities how to get the information: The easiest way is asking the person living in the house. This has several shortcomings: everybody who wants to check the registered information would have to ask again. How long would it take until the person living in the house gets angry? The information obtained from the occupant is not reliable; the person found on the premises may have an interest to pretend that she is owner even though she only rents. If the data is marked by the reliability with which the informant believes the data to have; high reliability comes from persons adding the data with closer connections to the occupant (family or friend); but then: How trustworthy is this kind of test?

Misconceptions about land ownership and the process of selling and buying land further confuse the data collected by VGI: the occupant may have a contract with the previous owner and therefore believes that he is the legal owner. Transfer of ownership of land, unlike transfer of ownership of other goods, is requiring a contract with additional qualifications. In most legal systems, it must be in writing and later registered to cause a valid transfer of ownership. Cases may occur, where all parties believe that A is owner because he bought from B, but legally B is still the owner of the land, because the contract is not valid or not registered.

The classification of the rights is difficult, but the spatial delimitation is even more difficult. From the observation that a person uses a specific piece of land it can neither be concluded that he has the right to use the whole area, nor can it be concluded that this is all of the that he could use.

4.2 VGI on Physical Objects

Observation of physical objects is simple. A fence, a wall, or a building is usually visible at a distance of tens of metres. However, laypersons can only take pictures of these objects. If they want to model the exact position of these objects, they usually have to refer to technology like GNSS. The application of such technology requires physical presence of the person at the location of measurement. This is not a problem for a fence separating private property from public space but what shall be done with a fence separating two properties? Surveyors are typically allowed to walk on privately owned land if this is required by their work. Laypeople producing VGI usually do not have this right and may even get sued for trespassing. So the only people allowed to measure the fence are the two neighbours and people having their permission to enter.

A way out is providing orthophotos. The creation can be organized by public bodies and the VGI-contribution is digitizing and modelling of physical objects (compare Konecny, 2012). However, this method cannot guarantee that the objects registered are used to delimit a right.

Another method could be evaluation of amateur photographs. Objects that are visible from several angles can be reconstructed at least semi-automatically. The advantage of this method would be that the huge database of already existing, georeferenced photographs can be used. Badly georeferenced photographs may be a problem, though. Photographs that are shown in completely wrong places will typically have no others to form a group that can be evaluated. However, if one image is slightly shifted, it will have an influence on the positional quality of the result. This problem may be solvable when combining several objects.

Something that is usually easier to observe is land use. It is usually widely visible if somebody is creating a house, planting trees, or growing wine on his land. This information can even be extracted from images. The advantage of VGI on land use is that land use is a classification of a (probably sporadic) process. There definition of a forest may be based on the process of forestry, not on the existence of trees. Thus if somebody observes a forest official working on a piece of land with trees it is a strong evidence that this is a forest. However, most of the time there will be no forest official and so a lot of observations may be necessary to detect the presence. The same is true for other types of land use, e.g., meadows. They are used regularly

but there may be periods where they are not used in order to let the grass grow again. But the area is still considered a meadow.

4.3 VGI on the Environment

Environmental information can also be collected by VGI. At least in theory, laypersons can collect soil or water samples for testing their quality. Fraud can be detected if enough samples are collected. However, the same problem of trespassing still exists. Taking soil or water samples requires physical presence and that may be prohibited by the land owner. However, soil and water quality have a significant spatial autocorrelation. A high soil quality on a specific parcel suggests that the neighbouring parcels have similar quality. This would allow automated plausibility checks for the data. However, the equipment necessary to do the sample testing may be quite expensive or difficult to use. It is questionable if the group of possible contributors is large enough to provide sufficient data. Therefore, land administration is based on a variety of different data:

- Rights and their spatial extent
 - Ownership
 - Use rights
 - Mortgages
 - ...
- Physical objects and their properties
 - Buildings
 - Land use
 - ...
- Environmental information
 - Soil
 - Water
 - Noise
 - ...

Some of these data rights are used in legal (civil) proceedings for land ownership and other private rights, taxation and land use planning, or the determination of land use subsidies for agriculture. As land administration is constantly under development and new challenges lead to additional tasks and data, such a list can never be complete and final. The applications of land administration data and services started with land tax and property registration only and today include support of the land market, environmental protection, and agricultural subsidies (compare Muggenhuber, Navratil, Twaroch, & Mansberger, 2011). Thus the above list can only be a snapshot of data currently used and the list may be extended in the future. This is a challenge for public data collection.

5. CONCLUSIONS

Quality is a problem for information in general and geographical information in particular. This problem is not automatically solved by using VGI. The Wikipedia encyclopaedia as one of the most successful examples of crowd sourced data is sometimes facing major edits (so-called “edit wars”) from followers and opponents of politically active persons (Garber, 2012). Can this problem also occur for land administration data? Business opportunities may be closely related to data stored in a land administration database. Thus it may be tempting for competitors to prevent some of these business opportunities by falsifying land administration data. Although it may be possible to correct these falsifications, delays may result in serious business disadvantages. What does this mean for any liability the public administration holds?

Should liability charges be passed down to the original contributor? A clear solution is necessary but may have an impact of the willingness of citizens to contribute data.

VGI approaches can be used but only for collecting specific types of information. Citizens can only provide information they have. This is in general information derived from observation or communication. Thus invisible facts, e.g., on the ownership can only be provided by a very limited number of persons and this provides problems for quality control in VGI. An approach to solve this could be a comparison between VGI and data from other sources like public registers and insurance records. It may be problematic, though, to handle contradictions between public registers and VGI because both sources may be wrong.

VGI can provide information on topics where direct observation is sufficient. For example, occupation and use of land can be observed and thus VGI processes can collect the information. It may be difficult to verify the data using VGI only since there may only be a limited number of citizens able to make the observations. However, these obstacles could be overcome. The result could be data sets supporting information from public sources. These additional data sets can provide a different perspective, e.g., land use data in contrast to land cover data provided by remote sensing.

The discussion showed that it is not possible to base land administration completely on VGI. Rights on land, which constitute a large portion of data used in land administration, cannot be observed directly by citizens with local knowledge because they are not observable. Thus the quality of VGI is not sufficient for the currently used land administration procedures, e.g., for land taxation. However, VGI can support traditional data collections mechanisms in situations where the observation is possible but heavily time-dependent.

6. REFERENCES

- Basiouka, S., & Potsiou, C., 2012. VGI in Cadastre: a Greek experiment to investigate the potential of crowd sourcing techniques in Cadastral Mapping. *Survey Review*, 44(325), pp. 153-161.
- Brunauer, W., Feilmayr, W., & Wagner, K., 2012. A New Residential Property Price Index for Austria. *Statistiken DAten & Analysen*, 12(Q3), pp. 90-102.
- Dale, P., & McLaughlin, J., 1999. *Land Administration*. Oxford University Press, Oxford, 169 p.
- Frank, A. U., 2009. Scale is Introduced by Observation Processes. In R. Devillers & M. Goodchild (Eds.), *Proc. 6th Intern. Symp. on Spatial Data Quality (ISSDQ)*, CRC Press, St. John's, Newfoundland and Labrador, pp. 17-29.
- Garber, M., 2012. Paul Ryan, “Brown Noser”? The Wikipedia Edit Wars Begin for Romney’s Running Mate. *The Atlantic*. August 11.
- Goodchild, M. F., 2007. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal*, 69(4), pp. 211-221.

- Goodchild, M. F., & Li, L., 2012. Assuring the quality of volunteered geographic information. *Spatial Statistics*, 1(1), pp. 110-120.
- Keenja, E., De Vries, W., Bennet, R., & Laarakker, P., 2012. Crowd Sourcing for Land Administration: Perceptions within Netherlands Kadaster. In *FIG Working Week 2012, TS03B*, 12 p.
- Konecny, G., 2012. Alternatives for Economic Boundary Determination in the Establishment of a Cadastral System. In *FIG Working Week 2012*, FIG, Rome, TS07E, 12 p.
- McLaren, R., 2011. Crowdsourcing Support of Land Administration – A Partnership Approach. *VGI (Österr. Zeitschrift für Vermessung und Geoinformation)*, 99, pp. 1-12.
- Mooney, P., & Corcoran, P., 2012. Characteristics of Heavily Edited Objects in OpenStreetMap. *Future Internet*, 4(4), pp. 285-305.
- Muggenhuber, G., Navratil, G., Twaroch, C., & Mansberger, R., 2011. Development and Potentials for Improvements of the Austrian Land Administration System. *FIG Working Week 2011*, FIG, Marrakech, TS07A, 15 p.
- Navratil, G., 2011. Cadastral Boundaries: Benefits of Complexity. *URISA*, 23(1), pp. 19-27.
- OpenStreetMap Foundation, 2013. Stats - OpenStreetMap Wiki., <http://wiki.openstreetmap.org/wiki/Stats> (April 8, 2013).