From SDI to an OGD Platform – Challenges and Opportunities

Jürg H. LÜTHY and Christian KAUL, Switzerland

Key words: Digital Cadastre, GSDI, OpenGovernmentData; traceability; Web-services

SUMMARY

In order to support the development and improvement of a society maps and spatial data play a core role. After running a SDI for the Canton of Zurich for several years the next generation of Information Infrastructure is under development. Its goal is to facilitate the use of spatial and non-spatial information to the citizens. The FIG-report on Spatially enabled Society demands under the key element "Data and Information" that sources of geospatial data shall be made available to support the location revolution. The way to do this in a sustainable manner is a consistent focus on Open Government Data (OGD), Information and Dialogue. Based on the federal strategy of OGD the government council of the Canton of Zurich confirmed the approach in the current government program. In order to accomplish these goals different aspects have to be considered.

First and most important the legal base has to be adjusted to publish governmental data. Three major issues have been identified to be fixed by legal ordinances. Shift from "proprietary" to a cost-free policy, freedom of reusing and altering dataset and handling the questions of limitations by personal rights.

For the provision of the data itself the technical infrastructure must be amended. The architecture can be differentiated between data storage, data provision, data processing and data delivery. Regarding the data storage the change will mainly be regarding safety issues (from internal use to external accessible sources). With respect to data provision the principles as described by Lüthy (2016) have to be implemented: data product specifications with a strong focus on algorithmic quality, contingency of data services also under adverse conditions and traceability requirements. To foster the use of spatial data by non-professionals the provision of data processing services can support users to create helpful applications. On the front end the services have to be highly standardised and the delivered results have to be well documented. The permanence of these services is a key factor for successful OGD-usage.

The paper describes the requirement for the new SDI strategy and provides information from practical experiences in the implementation.

From SDI to an OGD Platform – Challenges and Opportunities

Jürg H. LÜTHY and Christian KAUL, Switzerland

1. INTRODUCTION

The digital era of spatial data started in our environment exactly 30 years ago with the initialisation of the modernisation of the Swiss Cadastral system, but also with first data sets on water supply. Although at that time the map was the one mean to document the built environment. The pioneers in the transformation of maps to spatial data were not even able to plot the captured data as legibly as needed. Still these forerunners did not anticipate the technical progress for data use and distribution and neither the paradigm shift with respect to the use of data. Maps, cadastres and registries have played for centuries a central role in the society. The computation of some taxes has been based on maps on land ownership which were aligned with the proprietary registry. When the Swiss Cadastral System has begun to be modernised in the mid-eighties, they realised that the homogeneity of the captured data and the interchangeability between computer systems must become a mapping axiom. From this visionary approach many advantages have been derived since then.

Whilst on the technical side a major change was initiated, the accessibility of the data was still limited. In particular cadastral information has always been declared as public register but was not publically accessible. This didn't change with the digitalisation. The general opinion was that the significant costs for the modernisation should be refinanced through the users. Eventually the fast changes in digital mapping technologies, the growing importance of web-technologies – also thanks to the increasing band-with – and the financial power of large IT-companies fostered the development of parallel spatial data infrastructures. These infrastructures (commercials but also non-profit like OpenStreetMap) do not provide the identical information as the governmental ones but are enriched with information from other directories (restaurants, stores etc.) and its use is free of charge.

In the first decade of this century governmental agencies started to realise that return on invest will not be achievable since the request for expensive data was constantly decreasing. In 2007 the principles for releasing governmental data to the public were elaborated and published by the Sunlight Foundation[1]. The principles are the base for most definitions of Open Government Data (OGD).

1. Completeness: Datasets released by the government should be as complete as possible, reflecting the entirety of what is recorded about a particular subject.
2. Primacy: Datasets released by the government should be primary source data.
3. Timeliness: Datasets released by the government should be available to the public in a timely fashion.

[1] https://sunlightfoundation.com/policy/documents/ten-open-data-principles/

4. Ease of Physical and Electronic Access: Datasets released by the government should as accessible as possible, with accessibility defined as the ease with which information can be obtained, whether through physical or electronic means.
5. Machine readability: Machines can handle certain kinds of inputs much better than others, data should be provided in a widely-known and simple to process format.
6. Non-discrimination: Barriers to use of data can include registration or membership requirements.
7. Use of Commonly Owned Standards: Freely available alternative formats often exist by which stored data can be accessed without the need for a software license.
8. Licensing: Maximal openness includes clearly labelling public information as a work of the government and available without restrictions on use as part of the public domain.
9. Permanence: Information released by the government online should be sticky: It should be available online in archives in perpetuity.
10. Usage Costs: Most government information is collected for governmental purposes, and the existence of user fees has little to no effect on whether the government gathers the data in the first place.

2. MOTIVATION FOR OGD IN THE CANTON OF ZURICH

The FIG-report "Spatially enabled Society" focus on six key elements for a spatially enabled society. The element of "Data and Information" demands "sources of geospatial data to support the location revolution". The increasing demand for spatial data and the general trend towards OGD stipulated the Canton of Zurich to do this in a sustainable manner with a consistent focus on OpenGovenment-Data, -Information and – Dialogue.

Based on the federal strategy of OGD the government council of the Canton confirmed the approach to OGD in a decision (Hodel et. al. 2015). The goals of the legislation 2015-2019 declares the following guideline: *Goal of the legislation no 10: New technologies fosters the responsible use of data to simplify administration, reduces the burden for commercial organisation and provide a better transparency which is favourable for the society.* [2]"

Based on the current SDI allocating data and billing data to the users requires quite a lot of resources and the revenues don't cover all the costs. With the approach of OGD the automation of these processes is easier and can be done more consequently. So the loss of revenues after implementing OGD is not considered as significant financial challenge.

3. FROM SDI TO OGD - STRATEGICAL ISSUES

3.1 Changing mind-set regarding responsibility of data use

The data owner regards himself traditionally as responsible for the correct use of his data down to the end product. With the given barriers for data deliveries the data owner is theoretically aware

[2] cf https://www.zh.ch/internet/de/aktuell/news/medienmitteilungen/2015/legislaturziele_2015-2019.html, not available in English

how the intended use of data at the external organisations. In practical terms the usage of data is not controllable one the data has been shipped to the customer.

Under the concept of OGD the influence of the data owner on the data usage ends irrevocably with the allocation of the data sets on the OGD platform. It's out of the data owner's interest what third parties do with the data. It is their responsibility to make proper products out of the data, even when using the data in a quite different way the original data has been compiled for.
The change of the mind-set and the shift of responsibilities are for many data owner difficult to understand. To achieve their support for OGD it must be revealed how the current rules are not enforceable, especially because many customers have already processed and refined data sets so that the trace to the original data set often get lost. The shift of responsibility must be clearly stated. Furthermore the potential and advantages of OGD have to be made visible.

3.2 Legal basis

Within the Open Government Data initiatives the free and open access to cadastral data is always at the top of the wish-list which proves the relevance for spatial data in general and cadastral in particular. To strengthen the role of AAA-Datasets (Accurate, Authoritative and Assured) it is important to include this type of data (often cadastral data) in this cost-free-policy. Due to the use of these data sets as base for taxation the legal base for them is usually quite exhaustive. But most of spatial data governing laws have been written in the analogue era and are not adequate neiter to today's technical capabilities nor to the demands of OGD. A amendment of the legal base is needed for setting up and governing the spatial data infrastructure in the OGD-era.
Three major issues have to be fixed by legal ordinances.

1. A service on a base level for spatial data must be accessible without costs. To strengthen the role of AAA-Datasets (Accurate, Authoritative and Assured) it is important to include this type of data (often cadastral data) in this cost-free-policy.
2. The data-owner has to declare his data free for reuse in a comprehensive way. This can be done by individual decision for each dataset or more efficiently with a legal based principle for free reuse for all public spatial data.
3. A legal concept is needed to handle the questions of limitations by personal rights.

In Switzerland the personal information are protected under a privacy rules. In most cases it will not be possible to publish personal information in any of the OGD data sets. Once published data sets may be combined and enriched with other information. Therefore, the law on OGD must contain principals that the individual spatial data sets are declared and published as data free of personal information - independently from the potential combination with other data set. It is the responsibility of the data subscriber that no conclusion on an individual person can be made even when data sets are refinement or combined.

3.3 Interfaces

The range of users will be broadened from specialists to every citizen utilizing a variety of devices for spatial data presentation when more and more data sets will be published as OGD. The transition between the highly technical production environment at the Canton and the potentially

simple data retreiving systems on a citizens computer must be carefully designed. Following the principles 5 and 7 of the Sunlight Foundation declaration the interface shall be provided in a machine-readable form and in formats which are non-proprietary. The provision through kind of a common shop-solution with a lot of interaction needed from the end user are not regarded as future-oriented.

In many countries the provision of spatial data in a non-proprietary format and in a documented data model is quite challenging. In Switzerland the situation is much less critical: When starting the initiative for the reformation of cadastral surveying in the Mid-Eighties a standardised data model, a data model description and a data interchange format have been introduced. The Geo-language Interlis (cf. www.interlis.ch) provided the standard for describing and interchanging spatial data. The early commencement of standardisation of spatial data in Switzerland was a pre-requisite for building up spatial data infrastructures and support the fast technical implementation of OGD portals. Nowadays for a large amount of spatial data sets the data model is documented using Interlis and the data can be retreived as Interlis exchange files[3]. For all data sets falling under the law on spatial information a comprehensive documentation (similar to the data product specification published in the ISO standard 19132) and corresponding metadata (according to ISO 19115) are available. This allows the end users to building up the understanding for which purpose a data set has been produced, quality characteristics like timeliness of the data and potential limitations of the data.

3.4 System Quality

The abovementioned quality principles with data product specification can be regarded as a sound base for data provision. To ensure a proper usage of data through the end users of the data the traditional data set quality must be enhanced by technical aspects towards system quality (Lüthy et. al. 2015). The data quality can be described with the well-established quality elements form ISO 19157 (ISO 19157:2013). It must be considered that the requirements for data quality may be different between the data owner and end user, leaving the responsibility at the data owner. In order to support a proper use of the data it is required that the actual data quality is continuously measured and provided as metadata. The fear of uncovering deficiencies in data quality through third parties is often a barrier for not publishing data to the public. But this effect can also be regarded as major benefit: the more the data is used and the more users work with a data set, providing valuable feedback to the data owner will have a very positive impact on the data quality.

On the technical level it must be ensured, that quality of service (cf Lüthy 2016) and the logical consistency (data is in line with the formal requirements of the data model and data format) are achieved. When the two aspects are not fulfilled the automatic allocation and retrieval of data lead to erroneous data sets. Whilst data set quality can only be monitored partially by automation, the technical aspects can be controlled fully automatically. Again, the concept of Interlis as data modelling language and data set transfer format are established as well as many tools which can check all formal aspects of a data set.

[3] The format and structure of Interlis exchange files can be regarded as dialect of XML and is therefore readable by many applications.

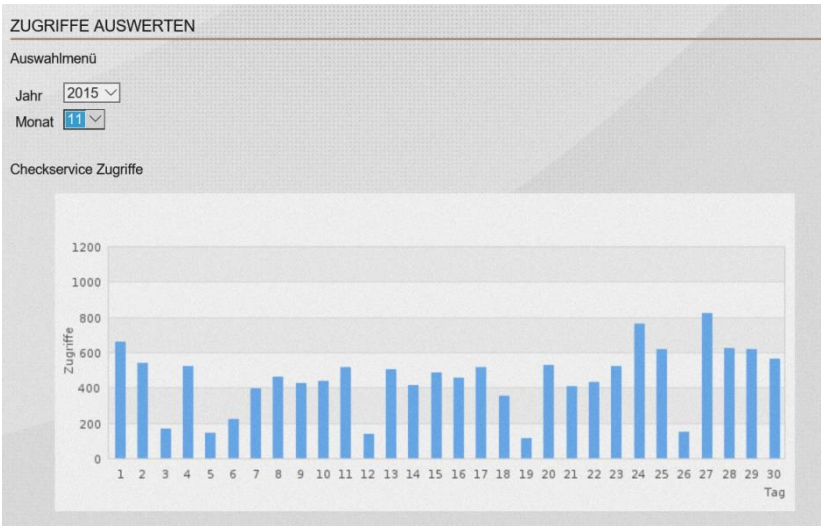

Figure 1 Established data validation for logical consistency - # of data set checks per day (courtesy of InfoGrips)

4. IMPLEMENTATION ISSUES

4.1 Back-bone

In the previous chapter the strategical issues and conventional ideas for setting up the OGD infrastructure in the Canton of Zurich have been described. The back-bone for the OGD is compound by preparative and operation aspect. The preparative aspects are all – mostly already running – initiatives for data harmonisation, data modelling, feature capture rules, portrayal rules and continuous data quality evaluation.

On the operational level the current spatial data infrastructure must be adjusted to achieve the desired quality of processes. It is assumed that the increased number of data sets falling under OGD will also result in higher loads on the server. When the digital elevation and digital surface model have been made available as OGD data sets, almost 100 users downloaded the complete data set within the first week, with a total of more than 10 TB data being transferred.

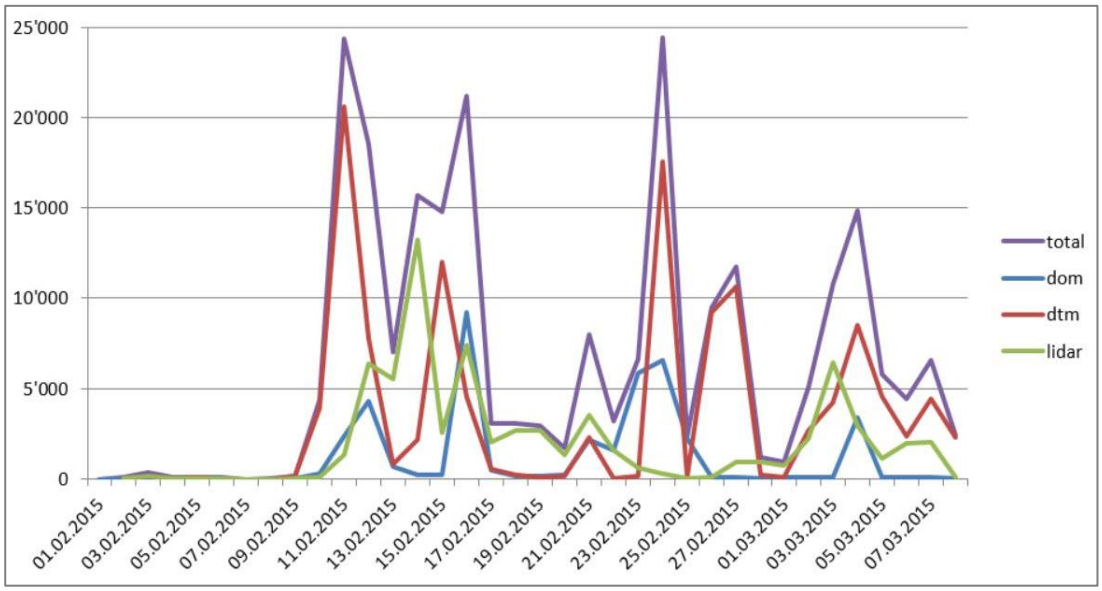

Figure 2 Number tiles downloaded per day after the release of the DTM / DSM as OGD

The access to OGD data is currently realised through the metadata catalogue for any spatial data set within the Canton. This portal must be optimised to simplify the search for data sets under OGD but also for improved documentation of the available formats and services (see also the following chapters).

4.2 Data services

The data services are the basic method for data allocation and provisioning under OGD and are regarded as core service in the Canton of Zurich. These services have to be highly standardised and the delivered results have to be well documented. According to the ten principals for OGD of the Sunlight Foundation the permanence of these services is a key factor for successful OGD-usage.

Data services can implemented as classic web-services like WMS, WFS or WCS or as download services. For the latter data from the production environment are periodically extracted and transformed to the desired format. All products and formats are usually processed at the same time, independently of the frequency of their use. This make-to-stock production has serveral disadvantages:

- Lack of efficiency: many data sets are produced which are never used;
- Lack up timeliness: the significant processing time for the to-stock production leads to periodical extracts (once per month);
- Lack of scalability: the launch of new formats or products is time-consuming.

The increased numbers of data sets which will be provided under OGD in future require a different approach: instead of holding available data sets the system will provide intelligence. This means that tools for data extraction, transformation and delivery will be made available. When an end user requires a data set, the current data will be extracted from the system, process on-the-fly to the desired format and model and automatically delivered to the end user. With the new approach the

data from both, the classic web-service and the new dynamic data service will have the same timeliness but may be used for different purposes.

4.3 Processing services

Public agencies may foster the usage of the data by provide not only bare data services but additional services like geo-coding, co-ordinate transformations. In the context of OGD it is considered as inadequate when public agencies invest in tools which at the end compete against initiatives by private companies. For the Canton of Zurich it is envisaged that the very simple processing services, which almost every data consumer may use, will be provided as so called enabler-services. More complex analysis like casting of shadow, suitability for solar cells, houses with view of the lake will bring a major added valued to a house owner and should be developed and marketed by private companies. Following basic services are considered of being not of interest for the private sector but can help to overcome the restrictions in accessing spatial data for non-spatial experts and will be part of the OGD:

- Geolocator (provided through the Federal OGD);
- Format transformation (form Interlis to dfx, shp etc);
- Co-ordinate transformation.

5. SUMMARY AND OUTLOOK

During the initialisation and setting up of the spatial data infrastructure most of the challenges were regarding conceptual and technical questions. With the stable operation of the SDI of the Canton the focus from data owners, provider and users is shifted toward the optimisation of the data usage. The principles of OGD can radically change the usage of spatial data in general. This is even more significant for cadastral data where the credibility of the information was for a long period directly associated with the personal representative of the licensed surveyor as the trustee of the cadastral data. A new conceptual approach for the strategic challenges and a careful dealing with the sensitivity regarding the provision of cadastral data are the sustainable base for the necessary developments and changes of the mind-sets. New and innovative ideas are required for many challenges like data quality, (split of) responsibilities, understanding of processes, data flows and last but not least also the technical implementation within the SDI.

The Canton of Zurich considers the implementation of the OGD strategy as an ideal opportunity to strengthen the relevance of governmental (spatial) data including cadastral data and to broaden the user domain. The first technical release of the new platform is scheduled for fall 2017 followed by a beta test phase. The go live of the OGD as part of the cantonal SDI platform is planned for January 2018.

REFERENCES

Hodel D., Lüthy J.H. (2015) Legitimierung von Open Government Data – welchen Einfluss kann die Politik nehmen. In eGov Präsenz, Fachzeitschrift des eGov Instituts, Bern. Ed. 2.2015, p.71-21.

ISO 19157 (2013). ISO 19157:2013, Geographic information -- Data quality. ISO International Organization for Standardization, Geneva.

Lüthy J.H. (2016). Expanded data quality model for increased reliability in mashed-up environments FIG Working week 2016, Recovery from Disaster, Christchurch, New Zealand, 2-6 May 2016

Lüthy J.H., Kaul, C. (2015). Demands for a spatial information infrastructure fit for Cadastre 2034. FIG Working week 2015, From the Wisdom of the Ages to the Challenges of the Modern World. Sofia, Bulgaria,17-21 May 2015
Sunlight Foundation (2010): Ten Principles for Opening Up Government Information, Sunlight Foundation, Washington DC 2010. https://sunlightfoundation.com/policy/documents/ten-open-data-principles/ last accessed 02 February 2017.

BIOGRAPHICAL NOTES

Jürg H. Lüthy is member of the Management Board at Achr Grad Ost AG, one of the largest geomatics companies in Switzerland. He obtained a master's degree in 1996 from Federal Institute of Technology Zurich (Switzerland) in Rural Engineering and Survey. From the same institution he holds a PhD (2007). He has many years of experience in spatial data management, transition from paper maps to data centric systems and the operation of Spatial Data Infrastructures. His current focus lies in the provision of holistic information using modern web-technologies like building the technical infrastructure for Cadastre of Public-law Restrictions on landownership. He is the Swiss delegate to FIG Commission 3. Since 2016 he is president of SLM Swiss Landmanagement Foundation.

Christian Kaul is head of Department for Geoinformation at the Office for Spatial Development in Canton of Zurich (Switzerland). He obtained a master's degree in 1992 from Federal Institute of Technology Zurich (Switzerland) in Rural Engineering and Survey. After ten years of experience in different domains like communal infrastructure, land management and SDI-Projects he worked a consultant in cadastral issues and procurement processes. Back in an engineering company he completed his experience in land use planning and spatial development. Since 2013 he focuses as head for Department on building modern cadastre systems and holistic spatial information infrastructure.

CONTACTS

Jürg H. Lüthy
Acht Grad Ost AG
Wagistrasse 6
8952 Schlieren
SWITZERLAND
Tel. +41 43 500 44 48
Email: juerg.luethy@achtgradost.com
Web site: www.achtgradost.com

Christian Kaul
Office for Spatial Development
Canton of Zurich
Stampfenbachstrasse 12
CH-8090 Zürich
SWITZERLAND
Tel. +41 43 259 27 67
Email: christian.kaul@bd.zh.ch
Web site: www.are.zh.ch / www.geoinformation.zh.ch