

Metadata Management – an Analysis from Creation till Use

Manuela Pötzsch, Wolfgang Reinhardt

GIS Lab – Arbeitsgemeinschaft GIS (AGIS), University of the Bundeswehr Munich,
Germany
{manuela.poetzsch, wolfgang.reinhardt}@unibw.de

Abstract. In times of distributed spatial data in complex networks and manifold use of spatial data, metadata as a reliable documentation and description of data, becomes even more essential for efficient utilization of resources. In this paper we report a research project which is carried out for the Bundeswehr GeoInformation Office (BGIO). The purpose of this project is to analyse the overall process of metadata capturing, storing and application. Therefore the importance of metadata with regard to interoperability for example in spatial data infrastructures (SDI) is taken into consideration and the current situation of metadata standards and standard conform production of metadata is reviewed. Subsequently, an overall view on processes involved from creation till the final usage of metadata is introduced. This holistic analysis considers that metadata is not required just because it is “nice to have”, but that metadata needs to fulfil user needs and even more, allow for a purposive usage of geodata.

1 INTRODUCTION

With regard to the implementation of INSPIRE, an European SDI, a survey of the European Commission aimed to evaluate current metadata availability and characteristics of online metadata for data sets held by the INSPIRE member states (Nowak and Craglia 2006). This survey demonstrates that organizations understand the necessity of providing and organizing metadata in a way that is compliant with international standards and is made accessible through network services.

2 METADATA SITUATION

“If there is a lack of metadata, it is certainly not due to a lack of standards” is remarked by Craglia et al. (2007) which is striking with regard to various initiatives that were launched to standardize the documentation of geospatial data. In 2003, the ISO 19115 geographic information - metadata was published which compiles expertise from different previous initiatives. Now that this international metadata standard is implemented in several commercial GI-software products, it is expected that coming products will

support this standard as well. Furthermore many commercial and non-commercial tools exist to produce metadata, as categorized in Behrends et al. (2006). Usually those editors support the acquisition of metadata by providing a template following a certain metadata standard. But consequently, since this ISO 19115 was introduced in 2003, a lot of older metadata catalogues still exist in other standards.

Another problem is that the ISO 19115 metadata standard is quite complex and very difficult to handle in practical implementation. For non-experts it is hard to deal with the high number of metadata elements and choose the “right” ones. Education work needs to be done, to make clear to communities that ISO 19115 just offers a list of elements with their according structure. Core metadata elements as often applied in accessing portals of spatial data certainly address different needs as descriptive elements for documentation of data lineage and usability.

3 METADATA MANAGEMENT

This paper focuses at an overall analysis of metadata existence from capturing through the final use. Regarding the point of usability, one can notice this lacks on existing metadata. Education on the side of producers is required in order to sharpen the awareness of the importance of data documentation, not only for their own purpose, but also for different kind of users.

An overview of different steps and parties involved in metadata life is shown in figure 1. The capturing of metadata should be done simultaneously with data capturing to assure best consistency. Most efficient would be automated production and extraction of metadata, which is stated in Gould et al. (2006) as a near term challenge. After the automated metadata extraction during data capturing, editing and respectively complementing needs to be done, since only part of the metadata can be extracted automatically. Different editors provide templates for editing metadata; those tools can be used for validation of the metadata according to a chosen standard or user requirements. From the experiences gained in the application of a metadata editor, developed in the framework of this project, this is of high importance especially for interoperability purposes.

Once the metadata exist the question arises, how to manage this information most efficiently. The big conflict in metadata management is that on one side metadata should be stored centrally for data queries, but on the other side metadata should be available at user sites together with the geodata to explain how the data can be used. The integration of data and its already existing metadata by a model-driven approach has been proved to

be possible in Najar (2006). Metadata management is considered as the whole process of transferring metadata from the producer to an appropriate storage system, and to provide user access. Therefore, a spatial data search engine that deals with filtering and rendering could be applicable for an authorized user. Metadata management needs to be aware of user requirements. Therefore, producers are in demand of user-specific metadata and this information needs to be made accessible through user-friendly interfaces. Particularly concerning the data sharing capabilities of the web, an interface could allow users to contribute to metadata by sharing their experiences gained through the application of a dataset.

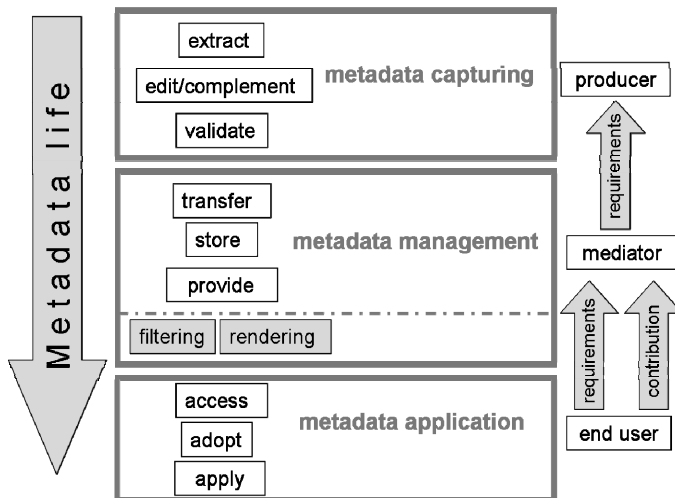


Figure 1: Overview of metadata life

Metadata can be categorized at different levels of granularity. For discovery of a dataset only a minimum amount of information is required to enable the user to find out what data exists, where and who the point-of-contact is. In order to evaluate the identified data, descriptive metadata is needed to judge the data for its fitness of use e.g. detailed information about the content. The difference between discovery and descriptive metadata needs to be pointed out, since discovery metadata should be similar for all spatial data products and organisations, but descriptive metadata depends on the different needs and guidelines of communities.

Our ongoing work focuses on an optimization of the whole metadata process and will define an adopted concept for handling metadata from creation to the final usage of metadata by different kinds of user groups. This involves analysing the whole combined process with steps like captur-

ing, editing, validating, transferring, storing, cataloguing, displaying, also updating or translating metadata. Especially the issue of application and usability is of high importance, as mentioned before. Consequently this study also aims to close the gap between users and producers as stated in (Comber et al. 2007).

4 CONCLUSION

Metadata standards as well as many tools to capture metadata exist, but more research work is required in the field of metadata management. The integrated workflow from creation, transfer, update through utilization of metadata in connection with data and user types needs to be analysed. In conformance with quality management guidelines of data producers, more attention should be drawn on documenting quality aspects and translating them in an “understandable” form for users.

REFERENCES

- Behrends, K., W. Czekga and S. Braune (2006). Metadateneditoren für geowissenschaftliche Daten. in *Angewandte Geoinformatik*, Wichmann, Heidelberg, 2006
- Comber, A.J., P.F. Fisher and R.A. Wadsworth (2007). User-focused metadata for spatial data, geographical information and data quality assessments. *Proceedings of AGILE*, 8.-11.05.2007 Aalborg, Denmark
- Craglia, M., I. Kanellopoulos and P. Smiths (2007). Metadata: where we are now, and where we should be going. *Proceedings of AGILE*, 8.-11.05.2007 Aalborg, Denmark
- Gould, M., J. Rocha, et al (2006). Near-term metadata challenges, 12th EC GIS Workshop, Innsbruck, Austria, 21.-23. June 2006
- Najar, C. (2006). A model-driven approach to management of integrated metadata – spatial data in the context of spatial data infrastructures, Dissertation, ETH Zürich, Nr. 16474, 2006
- Nowak, J. and M. Craglia (2006). *INSPIRE Metadata Survey Results*, 2006, European Commission, Joint Research Centre