Tasks and Functions of the Bavarian Environmental Object Catalogue in Meeting the Requirements Imposed by Environmental Legislation

Erich Weihs

Abstract

The claim for opening the information bases of the administration for general use has been underpinned by federal and state laws in the past few years. Many of the relevant regulations are based on EU directives and international law and are facts on the utilization of the internet that has become a matter of course. The briefly described legal bases point the way towards a provision of information by the administration. In addition, there are the requirements of eGovernment and eBusiness for cross-sector and inter-departmental services of the state. The services can only be realized as a joint task. This requires a change of paradigm in the administration. Authorities will only very slowly part with their way of thinking oriented by administrative services. In the economy, acting oriented by the market is a matter of course. Only the eGovernment initiatives are pointing towards market-oriented services, such as the Bund Online 2005 initiative.

Within the scope of the eGovernment initiatives, some inter-departmental applications begin to emerge. An initial step was made by including the health department of the city of Munich as a client. The service of the UOK is the identification and/or provision of data for information. It is a condition for the required systematic and comparable provision of environmental data that they will be actually found. The Environmental Object Catalogue (UOK) takes account of the requirement for systematics.

The basis of the UOK is a native XML internet database that consistently uses the current standard of XML technology. Native XML databases store the XML data without conversion into relational databases. That way, the XML data that are based on respective technical and professional standards and interfaces are processed directly and without any losses (a conversion of complex XML structures into relational forms is not exactly reversible).

1. Legal Requirements on the Environmental Object Catalogue (UOK)

The Federal Environmental Information Act (UIG) that entered into effect in 2005 clearly expands the citizens' right to free access to environmental information. Within the meaning of the UIG, environment is not only defined as technical and ecological environmental protection, but includes also areas of health and culture. The status of human health and safety, the living conditions of the people as well as sites of cultural interest and buildings as far as they are respectively affected by the conditions of the environmental components. This includes also the contamination of the food chain. The right to information does not only cover scientific data, but also data from the enforcement, e.g. of environmental impact assessment or other spatially relevant processes. For environmental information systems, this 1

Bavarian Ministry for Environment, Health and Consumer Protection, Munich (Bayerisches Staatsministerium für Umwelt, Gesundheit und Verbraucherschutz), erich.weihs@web.de

http://www.UOK.bayern.de

The Bavarian UIG (BayUIG) will correspond to the Federal UIG to a very large extent (also its wording). This paper cites the Federal UIG because the Bavarian UIG has not been adopted by the State Parliament yet.
means that information from all fields of sovereign acting must be made available in an active and comparable manner, i.e. the duty to provide information in certain areas cannot be complied with passively, i.e. only upon request. For the UOK, this means an extension in terms of identification of current data from administrative processes. 

Apart from the Environmental Information Act, there exist further legal provisions that follow the international trend (e.g.: USA, Canada and the majority of the EU member countries) and will lead to an opening of the administration and to a release of information. Like the Environmental Information Act, the Federal Act on Freedom of Information will entail respective state laws, too, as far as that has not already occurred like in Berlin, Brandenburg, North Rhine-Westphalia and Schleswig-Holstein. Under that law, the right to access to information exists in relation to all federal authorities (Kahler, 2005 12 ff.). That way, the act follows the international trend to guarantee access to the information of the public administration. According to Art. 1 Subsec. 1 of the law, everyone can procure, under the law, access to official information from federal authorities (and thus directly or directly from the individual states) or directly under respective state laws. A direct overlapping exists for example with regard to the obligation to provide information on processes of environmental impact assessments according to UIG Art. 10 (2) and others No. 6.

Another obligation to release information on data at the disposal of public authorities can be found in the articles of the Århus Convention that has already been ratified by Germany. Article 5 (2) reads as follows: "Each Party shall endeavour to ensure that officials and authorities assist and provide guidance to the public in seeking access to information, in facilitating participation in decision-making and in seeking access to justice in environmental matters.." and (9) reads as follows: "Within the scope of the relevant provisions of this Convention, the public shall have access to information, have the possibility to participate in decision-making .......

And also INSPIRE and de facto standards of other catalogue standards are to be taken into account here. A key objective of INSPIRE is to make more and better spatial data available for Community policy-making and implementation of Community policies in the Member States at all levels. INSPIRE focuses on environmental policy, but is open for use by and future extension to other sectors such as agriculture, transport and energy. The proposal focuses specifically on information needed in order to monitor and improve the state of the environment, including air, water, soil and the natural landscape. The altogether 34 articles regulate to a very large extent the coordinated collection and passing-on of "geographics" data. The annexes enlist all types of data to be collected within the scope of a comprehensive environmental concept - similar to the UIG - and Article 8 defines how they are to be collected as meta data. Meta data must also be published for geographic data even if they are subject to limited access. The reasons for a limitation are to be stated. Article 27 finally cites indirectly the standards that the concrete implementation is based on: "The standards adopted by European standardisation bodies in accordance with the procedure laid down in Directive 98/34/EC may support the implementation of this Directive. One of the standards is ISO 17904-2."

According to Art. 24 (6) of the Act on Freedom of Information as published in BGBl. I p. 1359, and risks assessments with regard to the environmental components according to Art. 2 Subsec. 3 No. 1. In the cases of sentences 1 No. 5 and 6 it is sufficient for spreading to indicate where such information is accessible or can be found. The published environmental information is updated at appropriate intervals.

**7** Support of the access to environmental information, UIG

**8** Infrastructure for Spatial Information in Europe. HYPERLINK "http://inspire.jrc.it/" INSPIRE.

Geographical Information - Meta Data that is also adopted by the UOK as far as they can be defined by ISO Meta File standards. Further obligations result from the publication duties of the EU for the fulfillment of which the UOK is very well suited due to its flexibility. Here, the border (that is artificial anyway) between meta data, facts and factual information is becoming blurred. Currently, »Violent Organic Compounds - VOC data« according to the Federal Anti-Pollution Act the spatial reference of which, according to ISO 19115, is related to the exact address of the source of emission are included in the UOK. Further types of data will follow.

2. Subject-Related Consequences of the Legal Provisions for the UOK

The tasks of an information provider are not fulfilled by the fact that it collects observation data alone, but only by making them available for utilization in compliance with the UIG and the other above-mentioned requirements. Factual data and meta data must be kept current and harmonized. Harmonization means comparability of the data and meta data in terms of space, time and contents - and the uniform assignment of keywords. INSPIRE even demands harmonization throughout the EU. The decentralized and scattered identification in the various internet offers of the division through internet search engines, such as Google etc., alone is not sufficient because the required harmonization of the data is not possible that way.

This means for the UOK that information from all areas of autonomous acting must be systematically made available or at least identified. Since most of the data are available in the division in a spatially and organizationally decentralized manner, their space- and time-related identification is of particular importance. The UOK meets these requirements as follows (Weihs 2005):

1. Unrestricted search in the UOK in the data accessible to "guests" by means of public search engines, such as google®, as far as they allow for a specific search;
2. Cross-sectional space-, time, and subject-related exploitation of the data in the UOK as such and via interfaces with other databases and search engines in the sector of authorities (e.g. PortalU);
3. Uniform keyword-assignment according to specialized thesauruses and fine-grade gazetteer (Weihs 2001, 2002);
4. User-friendly internet access to the environmental information for the general public and specialized users;
5. Additional subject-related integration of certain classes (e.g. research projects, VOC data) into the internet portal of the sector via specific roles;
6. Data maintenance and search in the UOK according to a uniform authorization and role concept in order to distinguish between guests and specialized users according to the respective needs;
7. Systematic classification of objects related in terms of subject and contents (e.g., data catalogue, measuring stations, fenland archive, media data) in specific thematic classes;
8. Subject-related integration into administrative processes (e.g. support of the planning of research projects, budgeting, approval, tendering, project course);
9. Harmonization of the meta data according to ISO in the division of the state government (as we are hoping) and within the scope of PortalU10 (www.PortalU.de), with the data catalogue of the federal country and of the individual states11 and GEOMIS Bund.

We assume that the use of the UOK starts with the search in google® with the identification of certain meta information. Currently, the UOK is represented in google® with 90,000 hits that in part are

---

10 German Environmental Information Network. Upon introduction of a new software, conversion to PortalU
11 Formerly, the virtual Environmental Data Catalogue (vUDK) that did not contain data itself, but searched in other catalogues, including the UOK
externally further linked. As far as the identification is sufficient for the user, the search with Google® is complete. Otherwise, the user will continue to make further detailed searches in the UOK. That way, it is possible, starting from Google® & co., to continue the search in the UOK. With the search in the UOK, the specialist access of authorities, experts, planners, etc. is facilitated under professional aspects (Fig. 3). The search in the UOK distinguishes between simple "Google utilization" and a subject-related utilization that can also be made in a space-related manner with a map-service functionality. Here, compatibility of the search results and their sufficient content are required. Only then the required assessment knowledge for sustainability strategies information bases will be supported. (Knetzsch, Rosenkranz 2003). It is aimed at providing, through identification in the UOK, a direct access to the data for the desired space, time and contents. The meta data a spatially related to the community level (e.g. soil information system, land map level, fenland archive, surveying aerial pictures), coordinates or, in individual cases, also to plants and factories (measuring stations, landfills). In individual cases, identification is still broken down to personal addresses (VOC data). The research projects included in the UOK for the division are normally without a spatial relation, but contain very detailed project information for the specialized user.


3.1 The UOK as a Service-Oriented Component of the I&C Strategy of Bavaria

The environmental data catalogue as meta information goes back to the core system designed in 1978 as an integration layer connecting the specialized information systems (Weihs 1978, LABO 1994). At that time already, integration was not meant to be the installation of a central data stock, but rather the idea of coupling the specialized information systems in specific problem-related cases. This requirement is met today by the "State I&C Strategy for the Bavarian Administration" adopted by the Council of Ministers. In this connection, we consciously speak of "coupling" and "joint task" in order to stress that the individual specialized systems are independent, but can use functions or data of other systems as services for tasks requiring such use. It is clear that this can only be achieved when there is a common view of the tasks of the administration that parts with a way of thinking in terms of branches, departments or silos. That requires a change of paradigms that cannot be achieved overnight. Understanding the integration or coupling layer in the above-mentioned sense therefore is the actual core of implementing service orientation. It connects - apart from the direct network access required for the specialized applications - the subject-related and administrative data with the individual applications via protocols of the internet technologies according to Fig. 1.

---

12 Google® is used as a synonym of public internet search engines
13 Duty to report to the European Commission according to Art. 11 Subsec. 1 of VOC-RL 1999/03/EC
14 The existing and future I&C applications are to be linked to form production processes for eGovernment services on the basis of loose systems coupled via standardized interfaces and internet technologies (so-called service-oriented architecture - SOA)
15 "Internet technologies" is used as a synonym for a number of TCP/IP-based techniques, interfaces and standards (such as W3C, SOAP, ISO, Catalogue Services, UDDI, XML, etc.) that of course apply also to the intranet, the net of the authorities, etc.
The UOK meets the requirements - to the extent they can already be met - of the state I&C strategy as a component of the communication and infrastructure services for the identification of facts and data in the environmental sector as shown in Fig. 2. The individual functionalities (thesaurus, search, application procedure for research projects, etc.) can be used as services by other applications. That is necessary because, according to the Environmental Information Act (UIG), »environment« is understood in a very broad sense in the UIG as described above and is independent from the respective departmental organization. Only this way, continuity of the fulfillment of the tasks in the I&C sector can be guaranteed without considerable additional expenditure e.g. in case of a change of departments. This is technically based on the identification where which data can be found, independently from the organizational structure of the departments. The reference is described in the UOK so exactly that the information from a search is already sufficient (e.g. media database, VOC data, research projects) or a data switching becomes possible. Since the practice-related meta information is quite detailed in many cases, the factual information found in the UOK will already be sufficient in various cases. Thus, web services descriptions form a standardized agreement that enables an automatic coupling of the services. According to Fig. 2, 3, this service connects the communication services with the coupling layer. The UOK has been designed with regard to these agreements and fulfills them to the extent it requires them for its services. The described architecture concept is designated as SOA (Service Oriented Architecture) that forms a component within the UOK.

3.2 The Technical Basis of the UOK

Originally, the UOK was based on a development for data identification that technically corresponded to the Environmental Data Catalogue (UDK) of the federal government and of the individual states. However, there did soon occur application cases that exceeded the rather relative term »meta data«.

---

Figs. 1, 2: Diagram of the service-oriented I&C architecture and its consideration in the UOK

---

279
Depending on a deductive or inductive point of view, these are in every individual case data or - in a narrower sense - meta data. In the UOK, all objects/entities are defined in XML according to one uniform methodology and grouped in classes that again form collections in the database: Objects referring to the same subject are grouped in classes.

As a consistent web service, the UOK is exclusively programmed on the basis of XML and in JAVA and is platform-independent. Except for the database subject to licence, there are only used JAVA and open source products. In the UOK, XML accommodates the realization of the above-mentioned requirements by its flexibility of the architectural concept without having to do without the advantages of relational or hierarchical models required here (Weihs 1999). A transfer of the XML data into a relational table structure is not necessary (and not desired here). Considerable advantages are the performance and the simplicity of programming because the tree (XML) structure of the object can be used any time (e.g. update of partial trees, alterations of/additions to the data model/schema). Standards according to ISO, de facto standards, such as Dublin Core, therefore are not only taken into account in exogenous interfaces, but are also imaged in the XML database model 1:1.

Access to the UOK is available via internet or intranet in every functionality (e.g. search, presenting, collecting, output, import, export, rights management, documentation). The individual functions (=services) can be realized as web services as the basis of an SOA concept for the access from other applications. The keyword assignment of the media database is already done as UOK thesaurus service. One of the following applications will be the utilization of the UOK thesaurus as a service in the eGovernment portal of the state government. Since recently, the collection of meta data has been possible by means of Word. A Word document will be stored in the database as an XML object. By means of respective XML queries, the document is analyzed and the meta information is created. Export into the database is effected upon “click” by a macro stored in the form. The other way round, a Word document can be provided on the basis of the XML object changed and stored in the database. That way, we hope to extend the circle of clients who are ready to enter data because only the customary office environment is required for collecting data.

The consistent separation of the XML data layer from the visualization realized by means of Stylessheets (XSLT) offers, among other things, another advantage: Since many data are referenced spatially in the UOK by coordinates or place names (the latter are converted into coordinates by the Gazetteer), that information can be used under Stylessheets at the moment of generation of the html output by integrating map services for a quite simple graphical representation.

For PortalU, the g2k interface and the index of the PortalU search engine are operated. The index has the advantage that the UOK hits appear on the ranking list of the PortalU representation. The same result, however, could also be achieved by defining the virtual directory that is provided for the public search engines. In that case, the index would be completely created at PortalU. The g2k interface is not taken into account in the ranking since no search-engine index is created here, but it is current with regard to the search because it directly represents the results of the database search.

The UOK in its technical realization as Veritas Cluster is installed at the computer centre ‘Rechenzentrum Süd’ of the state administration according to Figures 3 and 4 in order to guarantee 7/24 h operation. The failsafe operation is required in order to guarantee the services in the division and e.g. for the PortalU or the thesaurus. The data base is located on the RAID system of the computer centre. For further securing the operation, business-critical applications (e.g. research projects) are filed in an own database (instance) (DB 1). The publicly accessible data (via various guest accesses) are filed in the Public Database that is automatically matched with DB 1. The thesaurus is in an own database, too. That separation is necessary, among other things, in order to guarantee secure access of the guests (exclusion of possible manipulation from the network) and of search engines.
Apart from the databases and their respective software, the software component of PortalU is installed there, too. It comprises the software for the creation of the local index that regularly generates the data for the DSC interface of PortalU out of the XML database as well as the respective query software.

Data maintenance and subject-related applications of the UOK are realized in the subject-related database. The data of the individual classes (e.g., data catalogue, research projects, references in the media database, conservation areas, etc.) are grouped into so-called collections. There, the data supplier decides which data and with what view shall be made available to the public sector.

4. Résumé

The UOK tries to meet the relevant requirements of the UIG and other laws concerning the identification of environmental data by means of an inter-departmental approach. Technology and organization form the basis for its integration into the developing SOA architecture of the Bavarian state administration. The consistent XML architecture with its separation of data and visualization layer has proven to be successful. Consistent application does also mean that the "XML trees" of the DOM model are stored in the database without conversion. That way, complex x-query (= XML sql-queries) searches logically referring to the DOM model just like thesaurus and text operations are possible with good performance.
Studying this paper may contribute to a change of paradigms with regard to XML applications at the environmental authorities in one case or the other.

Bibliography

INSPIRE: Infrastructure for Spatial Information in Europe http://inspire.jrc.it/


Kahler, J (2005): Auskunft und Recht bei Daten GeoBIT 10/2005


