The Communication of the WWW-UDK with other Information Systems - Abstract -

Oliver Karschnick[1], Fred Kruse[1], Anja Spöringer[1], Sven Behrens[2], Marco Eichler[3]

Introduction

In order to fulfill the requirements of the Directive on Access to Environmental Information (90/313/EEC) the German and Austrian authorities have established the Environmental Data Catalog (German: Umweltdatenkatalog; UDK). The UDK is an appropriate tool to make the public access to environmental data more efficient and easy. Introduced in 1991, it can now be considered as a quasi standard of environmental metadata in both of these countries.

However, considering the necessity of presenting more data from different sources in a homogeneous way in order to get the best performing management system of environmental data, the UDK needs normative interfaces to other databases.

In this paper XML interfaces based on SOAP technology enabling the WWW-UDK to exchange data with other metadata systems are presented.

The UDK Data Model

The UDK has been developed to collect and retrieve environmental metadata. This data should answer the question: *Who* holds *where what* kind of data in *which* format?

The metadata records are stored as so-called UDK objects, the attributes of which are e.g. object identification number, object name, general descriptions, key words as well as time and spatial coverage.

---

[1] Niedersächsisches Umweltministerium, Koordinierungsstelle UDK, Archivstr. 2, D-30169 Hannover, udk@numis.niedersachsen.de, http://www.umweltdatenkatalog.de

[2] disy Informationssysteme GmbH, Stephanienstr. 30, D-76133 Karlsruhe, behrens@disy.net, http://www.disy.net

[3] BVG mbH, Menkestr. 7, D-26419 Schortens, udk@bvg-fri.de, http://www.bvg-fri.de

The objects are assigned to 6 different UDK object classes. These classes are labeled as:

1. dataset / database
2. service / application / information system
3. document / report / literature
4. geographical information / map
5. organizational unit / task
6. project / programmes

In principle, addresses of the administrations and objects being assigned to one catalogue are distinguished. Based on the identification number objects and addresses can be linked hierarchically (parent-children relation) and in parallel (object-object relation), which is a special feature of this data model. However, it is possible of course to include more catalogues into one database.

Collection and retrieval of metadata with the UDK

For the collection of metadata the Windows-based UDK (Win-UDK) is used. This programme is built for the retrieval and management of the database, whereas the web-based WWW-UDK was developed for public retrieval of data only.

Windows based collection and retrieval of metadata

The Win-UDK, actual Version 4.3, is used to collect and retrieve the environmental metadata. The input of data is usually done by the experts, who are responsable for the data themselves, so that a high level of data quality is achieved. To realize this decentralized input of data a decentralized software is necessary also.

The Win-UDK 4.2 is available in a stand alone ('E-Version') or a client server version ('CS-Version'), for both of which a three-layer architecture is used. For the communication between the client and the server through the local area network (LAN) or the internet the SOAP technique (Simple Access Object Protocol) is used in this version. A special update service guarantees that all databases concerned obtain to the actualized data. Furthermore an update service for the client software is integrated to provide up to date databases through the LAN and the internet.

Both, the E and CS version, are developed with Visual Basic 6 and are being running on all Windows operating systems (Windows 9x, NT 4.0, 2000, XP). The databases Oracle, MS-SQL-Server, Access, Informix, and Ingres are supported. In

addition, the Win-UDK contains a XML-based interface to provide an import of data from other information systems.

Web-based retrieval of the metadata

The WWW-UDK allows every internet user to retrieve the data which was being collected by the use of the Win-UDK. Actually, the Virtual HTML-UDK (V-UDK, can be accessed on http:\\www.umweltdatenkatalog.de) allows retrieval accessibility to the databases of the complete UDK-data collected in Germany. At the moment there are about 25000 objects included, a number which is rapidly growing.

The WWW-UDK is a multi layer client server application developed in Java. The WWW-UDK provides a browser based HTML-UDK as a client frontend which is developed as a Java Servlet and can be used immediately by any browser. The UDK Server manages client requests, queries the underlying database and maps the result to a UDK business object, which is then printed for the user. The Virtual HTML-UDK has been realized by adding an extra layer for distributing the queries and collecting the results. The V-UDK Server is communicating with the distributed UDK Servers by the means of Java RMI.

The UDK has become a de facto standard for environmental metadata. However, since there are a lot of other information systems available in the internet, there is a great interest to integrate this data as well. And of course in the new ages of Web Services there is on the other hand a need to access the UDK not only interactively via the UDK-homepage but by other (web) services as well.

The next paragraph is dealing with the implementation of an interface enabling the UDK to communicate with other information systems.

The SOAP Interface of the WWW-UDK

Operating in a steadily growing heap of metadata, it is inevitable to ask for an increasing functionality of the V-UDK in terms of accessing other environmental metadata information systems, for example like the Bavarian catalogue UOK. In addition it is useful to extend the search facility to other specific information systems like for example the information system of Baden-Württemberg. To achieve this functionality also vice versa, i.e. that other information systems can use the UDK as input, a bidirectional interface based on SOAP technology (Simple Object Access Protocol) has been implemented in the WWW-UDK software.

The specifications of this interface are:

1. the exchange format for all exchanged data is XML
2. XML formatted retrieval requests to the WWW-UDK from any information system are possible
3. including additional catalogues in the search of the WWW-UDK
4. the HTTP-protocol is used for the data-transport

Realizing the XML interface with the SOAP technology offers many advantages, the most important of which is the already standardized SOAP as communication technique between single applications in the internet as well as the independence of operating systems and programming languages.

This technique promises of course a huge, new operational area. To prevent problematic interactions with other systems and their data structure so called udk-namespaces are used. The namespaces are declared via a URI uniquely, e.g. for the UDK xmlns:udk=http://www.umweltdatenkatalog.de/udk is defined. This declaration protects against name clashes, i.e. a name of an element can be used with different meanings in different DTDs (Data Type Definition) and attributes of one DTD can be used globally.

Conclusion and Perspectives

The UDK is a metainformation system and management tool for environmental metadata. Considering the great increase of interesting and important environmental data throughout the internet, a special interface based on the SOAP technology was implemented in the WWW-UDK software. This enables the WWW-UDK to communicate bidirectionally with other metadata information systems. Up to now the UDK communicates successfully with two other systems: Firstly, metadata of the Bavarian data catalogue UOK can be retrieved by the WWW-UDK and secondly GEIN (German Environmental Information Network, www.gein.de) has included the WWW-UDK so that the UDK database can be searched via the GEIN homepage also.

Due to the fact, that the management of geographical metadata will play a more important role in the future, the UDK data model will be adopted to be compatible to the international standard ISO 19115 for geographical metadata.