The UDK  
Present Status and Future Development  
An Overview

Fred Kruse¹, Oliver Karschnick¹, Anja Spöringer¹, Marco Eichler²,  
Sven Behrens³

Abstract
The Environmental Data Catalog (UDK) is a system for public access to environmental  
data for which various tools have been developed to collect Metadata and present them  
in the internet. Through various interfaces it is possible to do research in other metadata  
systems and merge the result in a common way. In the other direction the interfaces of  
the UDK can be used for research and data-retrieval by other information systems like  
the German Environmental Information Network (GEIN).

Introduction
The Environmental Data Catalog (German: Umweltdatenkatalog; UDK) is a meta-  
data information system. From the time of it’s conceptualization about 10 years ago,  
it has been refined and upgraded continuously. This article shows aspects of the  
organization, the data collection and the software. Furthermore it provides a view on  
future developments.

Organizational Evolution
The bases of the UDK were put in the years 1991 to 1995 in a research and develop-  
ment project at the Ministry of Environment of Lower Saxony promoted by the  
Federal Office for Environment Protection (Schütz/Lessing 1993, Lessing/Schütz  
1994).
Still at run time of the project a co-operation of the larger part of the German countries was formed, aiming for both the development and the dissemination of the UDK. In 1993 Austria entered a formal co-operation with the Federal Republic of Germany concerning the environmental data catalog. At the same time, an Environmental Information Law introduced the UDK as the official navigation tool for all environmental information on record in Austria.

Since 01.01.1996 an administrative agreement UDK exists between the Federal Republic of Germany and 13 of the 16 German states. In 2001 the states Rheinland-Pfalz and Bavarian also joined the agreement. As of today the state Berlin is the only missing member.

The coordination center UDK was installed in 1996 in order to handle the business of the administration agreement, to maintain the UDK software and to organize future developments.

**Evolution of data**

Since the major software of 1997, the number of data entered into the system has been constantly growing. From the end of 1997 to the end of 2001 in Germany the number of records has more than trebled to about 25,000 data records. In Austria about 12,000 data records have been collected. With the integration of the Bavarian UOK into the V-UDK, the number of records will be increased by 30,000 Objects.

During the last year the German focus of the cooperation changed from a more quantitative sight on the data to a more qualitative approach. Because of the federal character of the German environmental administration the states are responsible for the metadata collection themselves. The discussion was focused on a more homogeneous data gathering in the country as a whole.

A similar aspect is the degree of completion of the individual UDK fields and the incomplete data coverage over all areas concerning the environment. This is a central point of study during the year 2002.

**The UDK Data Model**

The UDK was developed for collection and retrieval of environmental metadata. The metadata should inform about 'who' holds 'where' 'what sort' of data in 'which format' (Swoboda/Kruse/Nyhuis/Rousselle 1998, Swoboda/Kruse/Nikolai/Kazakos/Nyhuis/Rousselle 1999). Both digital and analog data are of interest. The metadata records are stored as so-called UDK objects. The attributes of UDK objects are related to the identification of the object (object name, general description, key words, etc.), and to the technical, spatial, and time coverage of the data to be described. Corresponding to various types of data/tasks being described, the following UDK object classes are offered:
These classes reflect the most important types of environmental data represented as UDK objects as well as environmental data producing tasks and programs identified during the last couple of years.

To fulfill the needs of the various metadata providers an extensible data model was introduced. An optional administration module UDK-Z allows to define up to three additional classes and to add attributes to existing classes. This development reflects the use of UDK not only in state authorities concerned with environmental issues but also in further authorities such as geodetic authorities.

In addition to the management of UDK objects the UDK also provides a simple address management tool and a polyhierarchical thesaurus. Each UDK object can be linked to one or more addresses, which are administrated separately. This method avoids multiple input of addresses. The controlled vocabulary of the thesaurus as well as free terms can be used for indexing UDK objects and addresses.

**Software**

In the passed years UDK applications were developed for the collection, administration and retrieval of environmental metadata. The interaction of the single applications with each other can be represented as shown in Figure 1.

**Software for metadata collection**

A Windows based software was developed for metadata collection. This application offers the possibility of entering environmental objects and addresses. In addition an administration tool for the entire UDK data collection is applied also. The software was developed in Visual Basic 6. It is running on all Windows operating systems (Windows 9x, Windows NT 4.0, Windows 2000 and XP). The database management systems Oracle, MS-SQL-Server, Access, Informix, and Ingres are supported.
In most cases the collection of metadata takes place in decentralized databases (several databases per state). The Windows UDK offers a special update mechanism (updating cycle) to exchange the modifications between the decentralized databases. The mechanism can be automated through the so-called UDK-Agent using an eMail system.

The UDK supports data exchange with many external systems through a XML interface (Kruse/Eichler/Freitag/Sattler 2001), so metadata can be imported from many other systems.

The number of users of the UDK was growing continually in the last years. In many agencies a lot of users are working on one database. So in 2002 the Windows UDK was split into two Version to collect data:

1. A single user Version (UDK 4.3 E): Easy to install for local application.
2. A Client-Server-Version (UDK 4.3 CS): This software is based on a three-tier architecture. For communication between Client and Server the XML based SOAP-Protocol (Simple Object Access Protocol) is used (Kruse/Eichler 2001). The UDK 4.3 CS can not only be used in local area networks (LAN) and intranets but also in the internet.

The use of the Client-Server-Version provides the following benefits for installations of many clients:

- The whole software and database administration can be done centrally at the UDK-Server.
- The installation of the Client is very easy. There is no special knowledge of the user necessary.
- Updates of the clients can be done automatically.
The whole comfort of the windows UDK is provided to the user. Because of the use of XML and the SOAP-protocol it is easy to change the UDK-Client to a web-application in the future.

Figure 2: The architecture of the UDK 4.3 CS

Retrieval Software for the Internet

For retrieval of metadata via the internet the WWW-UDK was developed (Nikolai/Kazakos/Kramer/Behrens/Swoboda/Kruse 1999, Swoboda/Kruse/Legat/Nikolai/Behrens 2000). This application consists of two components:

1. The HTML-UDK is a Java servlet. It runs on a Web-server and communicates with a standard HTML-Browser. The application provides a quick search, an expert search (with various search filters, e.g., concerning the spatial, the temporal or the thematic relations of the UDK objects and addresses, respectively), and a thematic navigation along the thesaurus structure. Both quick and expert search can be used to retrieve for UDK objects and for addresses.

2. The Java-UDK is the enhanced application. It is oriented towards the retrieval functions of the Windows-UDK. It provides a UDK Navigator corresponding to the UDK Explorer of the Windows-UDK, an expert search with various search filters, a Thesaurus Navigator, and a geographical search called GeoView.
To compile a search in the catalogs of all states the so-called virtual UDK (V-UDK) was developed (Figure 3). This application provides a search in all UDK databases installed in the internet. The results are presented homogeneously in a single list.

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 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.

First candidate for using this interface is Bavarian Environmental Object catalog (UOK Bayern). A search in the V-UDK will search in the data of 14 partners. The results will be merged together and shown in a homogenous way. So the view of an user of the V-UDK is transparent.

The V-UDK provides also an other interface to the German Environmental Information Network (GEIN). Thereby, searches through GEIN are extended to search in the V-UDK.
Future Development

Certainly, the most important step during the next years will be the integration of GEIN and UDK into a common information system. In 2003 the administration offices for both systems will be merged at the current UDK offices at the Ministry of Environment of Lower Saxony in Hannover. The development of a common user interface for the retrieval of UDK and GEIN data is being discussed.

The administration of geographical metadata will play a very important role in future. In 2002 the international standard ISO 19115 will be published as a recommendation. The data model of the UDK will be improved with the aim to achieve compatibility to this standard. The UDK will get interfaces for both in- and export and for the connection to external systems like e.g. the German Broker for geographical data GeoMIS Bund.

The application for metadata collection, Windows-UDK, is available for Microsoft Windows OS only. The client server version of this software will be used for data collection via the internet. Many German states use Unix and Linux OS on their
web-servers. Thus, the development of a platform independent server application for the Windows-UDK is under discussion.

Furthermore, it is planned to extend the international cooperation. For the use of the UDK in other, non-German-speaking countries, it will be necessary to upgrade the software with a multilingual user interface.

Bibliography


