

Outlook-Based Concept for the Population and Updating of a Meta-Information System in Environmental Administration

P. Fischer-Stabel^a, R. Krieger^a and E. Rietzke^b

^a *Umwelt-Campus, University of Applied Sciences Trier, P.O.B. 1380, 55761 Birkenfeld, Germany
(fischer-stabel@umwelt-campus.de, krieger@umwelt-campus.de)*

^b *silvercreations software AG, Campus Allee, P.O.B. 319032, 55768 Hoppstädten-Weiersbach
(eric.rietzke@silvercreations.de)*

Abstract: Metadata is essential in retrieving information about the environment located at differently distributed organisational units of governmental administration. The population of the meta-information system may be done based either on a centralised concept (e.g. with support of external interviewers) or on a non-centralised concept requiring a software tool accessible at each desktop computer where the dataholder can enter the data directly. Often the implementation of such a tool is linked with a high effort in installation and training, especially if hundreds of clients have to be served. To overcome these disadvantages we developed a software tool based on MS Outlook supporting the non-centralised maintenance concept. The software tool is used in connection with the environmental data catalogue (Umweltdatenkatalog, abbr. UDK) developed by German and Austrian authorities.

Keywords: Metadata Information System, Environmental Data, Umweltdatenkatalog

1. INTRODUCTION

Information about the environment is complex and located at hundreds of distributed organisational units of governmental administration, in industry and non-governmental organisations. Therefore, there is a challenge for environmental meta-information systems (meta-IS) supporting experts but also interested citizens in the field of research and retrieval of environmental data and information.

Among the public authorities in Germany there exists a state-region co-operation which is responsible for the design and implementation of the environmental data catalogue (Umweltdatenkatalog, abbr. UDK) at the different organisational units in environment-related administration. This UDK was designed as a meta-IS for the navigation to environmental data collected by and located at multi-level public administration. It is providing its users with information about the following data: WHO is having WHICH information at WHAT place.

Essential in implementing and maintaining such a meta-IS is the data entering process and the data quality control. We have to face the fact that there exists a huge number of data providers located at different distributed organisational units. Another gap is periodical data collection and updating. Both points are outlined in the user requirements for the system: The tool should be made available without any further installation effort and it should be a load-and-go solution, easy to use.

As part of the work described in the paper we analysed the technical and organisational environment of the federal government of Rhineland-Palatinate to introduce and populate such an environmental IS in an efficient way. Special focus was given to the following requirements defined by the government:

- Online population of the database by the distributed data generating units.
- Expectation of about 2000 UDK-objects.
- Updating of objects should be done by the dataholders.

- No additional effort in installation and maintenance of pieces of software.
- No modifications at the existing UDK.

The paper starts with a brief description of the environmental data catalogue UDK. In section 3 we describe the actual procedures for entering and maintaining data. Moreover, we discuss the advantages and disadvantages of these procedures. In section 4 we propose a new procedure for maintaining the metadata. This procedure is based on MS Outlook and is completely integrated in the office workplace of the data holders. We conclude with a short summary and give some hints regarding future work.

2. UDK: AN ENVIRONMENTAL META-INFORMATION SYSTEM

The WinUDK (UDK 4.2 SP1) is an operational software system designed for the entering, management and retrieval of environmental metadata [Lessing et al. 1995, Swoboda et al. 1999]. The system is based on different commercial off-the-shelf database systems as MS Access, SQL-Server or Oracle. The link between database and WinUDK is made with ActiveX Data Objects (ADO) technology.

The metadata is represented as objects in the UDK-system. We have a taxonomy of different object classes as listed below:

- dataset / database
- service / application / information system
- document / report / literature
- geographical information / map
- organisational unit / task
- project / programs

From technical point of view, there are different system architectures available (see figure 1):

- Isolated local installations with MS Access.
- Extension of the local installation by generation of different UDK instances at different workplaces.
- Client/server based architecture. WinUDK's represent the clients who are accessing to a central database at the server side.

Communication between the different instances is conducted by an automated communication cycle described in Swoboda et al. [1999]. Within this

process the updated data-pools of every local catalogue will be harmonised using the central database.

Another option for the performing of the data exchange is given with the XML-interface of the WinUDK. The metadata from the UDK will be transferred and converted to XML-files which finally are used in updating the database.

Additionally, the Umweltdatenkatalog offers two other software components: The HTML-UDK and the Java-UDK. Both components are named by the term WWW-UDK.

3. METADATA MAINTENANCE: CURRENT SITUATION

In general, from an organisational and technical point of view we have to separate the two following situations: On the one hand, there is the initial data entering process for the very first population of the catalogue. On the other hand, we have the situation of updating and maintaining the database during its operational use. The initial data capture is mainly done by external interviewers.

In the domain of operational updating (insertion, modification and deletion of UDK objects), there are actually two different concepts implemented; namely, centralised and non-centralised maintenance:

3.1 Centralised Maintenance

Central entering of the data is performed by a manual transfer of the information to the central unit. The environmental information is made available by the data generating units which have stored the data as analogue or digital text-file. The data collection process is done periodically with a frequency of e.g. three months. The centre with its 1-3 UDK workplaces is responsible for quality control and the transfer of the information into the central catalogue.

An optimisation of this procedure can be reached with the use of Word-templates instead of analogue documents. The use of a well known system (e.g. MS Word) for data entering which requires no training and additional software installation is the advantage of this concept. The lack in end user support (e.g. tooltips) and in help functionality explaining the use of the form, but also in the necessary option for the generation of object hierarchies are main disadvantages. Therefore, in many cases there is a need of quality control to be done at the central unit.

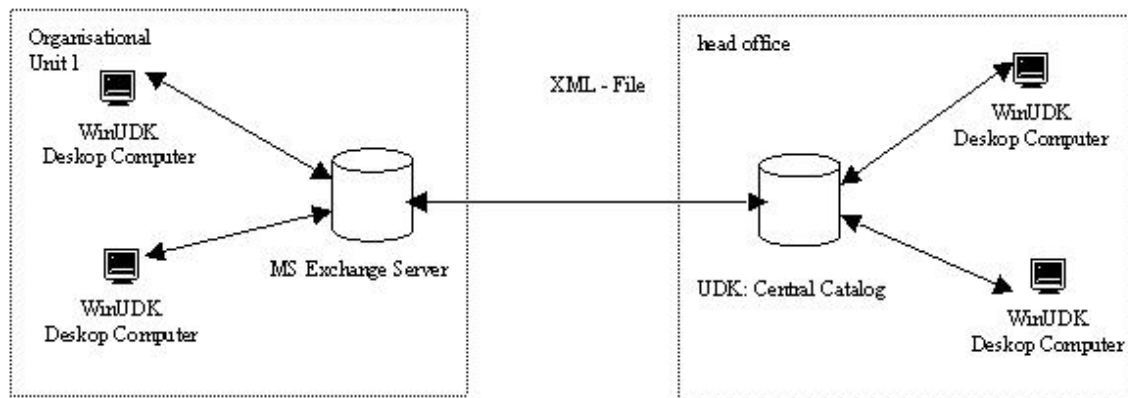


Figure 1: General Architecture of an UDK-installation.

3.2 Non-Centralised Maintenance

The data entering process is performed within the non-centralised concept with WinUDK clients installed at the data generating units. This concept is based on the client/server technology linked to the installation of a lot of UDK clients. The local expert enters the data, he/she or a responsible officer will do the quality check, and, with the next step, he/she will finally write the new or updated information into the central catalogue.

No breaks in the media used and, in addition, the location of the data quality control at the generating units are some of the advantages of that concept. Further benefits of distributed population and updating are given through the following points:

- sharing the effort for data entering by the units.
- responsibility in quality control is with the data generating units.
- near-real time update of the information.

There is much more environment-related competence within the data generating units compared to the centre. Therefore, the quality control is more efficient compared to the centralised solution.

Because of the high functionality of the WinUDK linked with the sporadic use of the client, there is a special effort required in training all the local experts at the different units. Other disadvantages of this concept are the high effort in installing and updating/maintaining the client's software.

4. METADATA MAINTENANCE BASED ON MS OUTLOOK

As discussed in the previous section the non-centralised data maintenance based on WinUDK has considerable disadvantages. For that reason we looked for an alternative software tool.

In doing so we analysed the technical infrastructure of governmental administration in Rhineland-Palatinate.

With respect to desktop computers the result can be summarised as follows:

- Operating systems: Windows 98 and Windows 2000.
- Software Standards: Microsoft Office, e.g. MS Outlook 98 and MS Exchange Server 5.5.
- Network: 64 kbit – 34 Mbit.

It can be assumed that this infrastructure is typical for most German and European public state authorities. MS Outlook is the standard e-mail and organiser program at each desktop computer. At each organisational location a MS Exchange Server 5.5 is installed for the management of the messaging infrastructure.

Microsoft offers a lot of possibilities to build and deploy custom collaboration applications that link MS Exchange Server and MS Outlook as client [Byrne 1999, Microsoft 2000]. Moreover, most of the users are familiar with the basic functionality of Outlook (e.g. cut, copy, paste information and find items). Based on these facts MS Outlook and Exchange Server provide a very attractive platform for the development of an application for entering and maintaining metadata if there is a large number of dataholders located at different distributed organisational units. The core of the developed application is a specialised form and is

implemented in VBScript. The user fills the meta data in this form. The entered metadata items are assigned to a Outlook-folder that is shared with all other people of the group responsible for the same thematic category. The permissions that allow others to add or change items are set by the administrator of the Exchange Server.

4.1 Technical Architecture

Fig. 2 depicts the technical architecture of the solution based on MS Outlook. At each desktop computer the HTML-UDK can be used for finding the environmental information. MS Outlook is installed as the standard e-mail and organiser program. As described before it is also used for maintaining the metadata about environmental information. All data including the form for

entering the data is stored in the MS Exchange Server that is installed at each location of the governmental administration. A program transforms the metadata items stored in Exchange Server into a XML-file according to the document type definition given in Koordinierungsstelle UDK [2001]. The XML-file is attached to an e-mail and is transmitted to the administrator of the central catalogue. The standard interface of the UDK allows the administrator to import XML-file containing the metadata records. The rights for changing the data will remain at the offices entering the data. Only the terms of the Thesaurus needed for the classification of environmental information and the address data records of the public authorities will be managed in the central catalogue. They will be sent to the Exchange Servers periodically.

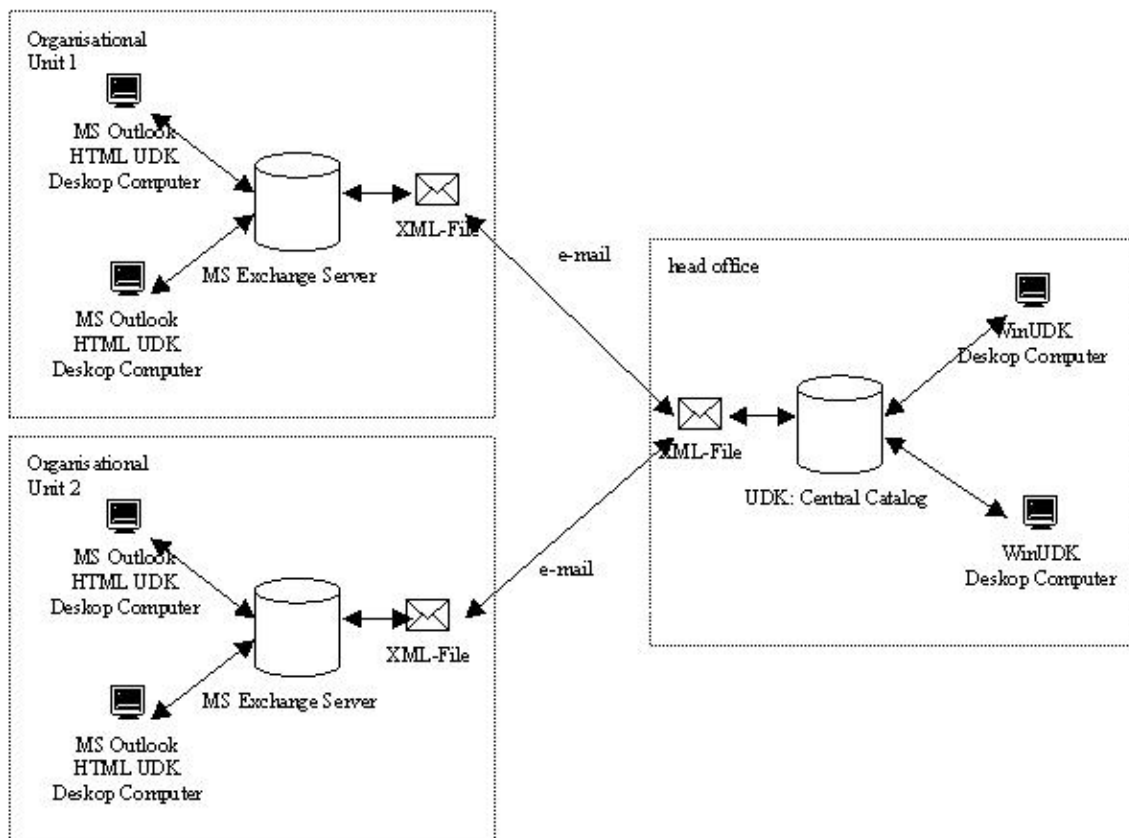


Figure 2: Outlook-based data maintenance: Technical Architecture.

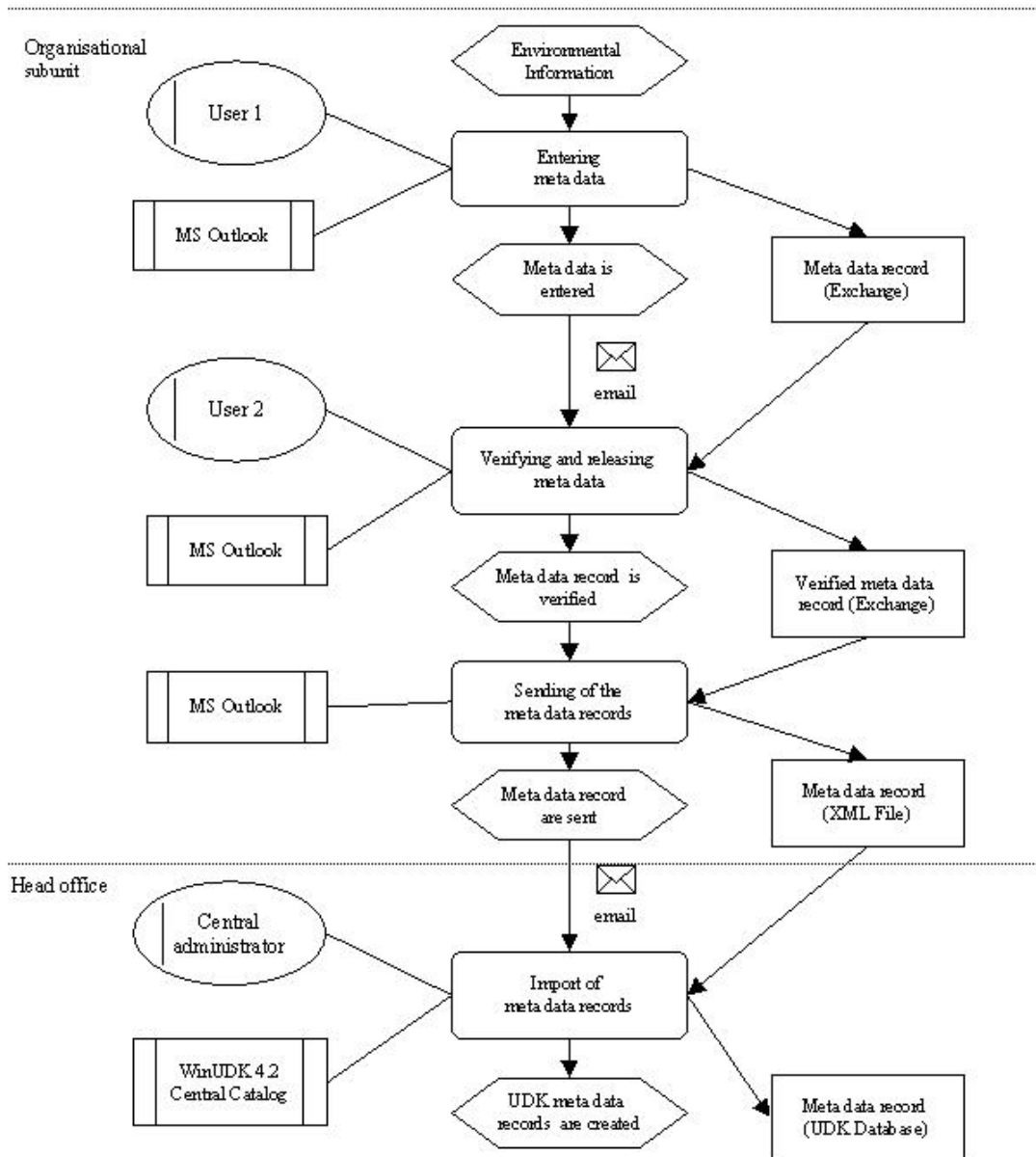


Figure 3: Maintaining metadata based on MS Outlook.

4.2 Process Description

Fig. 3 describes the process of entering and maintaining metadata based on MS Outlook. In a first step the user starts Outlook. On the Outlook bar or the folder list the user can select the folder called UDK. When clicking on this folder, the view pane displays existing metadata items of the folder. The items are grouped according to their status (*New*, *Ready for Quality Check*, *Released*). Similar to editing an e-mail the user can select an item or create a new item. The metadata is entered in the custom form. The data fields which have to be entered are thematically grouped into sub-forms

which are assigned to the main form. Explanatory messages are assigned to each data field. Required fields are clearly marked. Check routines prevent users from entering incorrect values. After completing the form the user receives a completion signal and the item receives the status *new*. When the OK button is pushed, the item receives the status *Ready for Quality Check* and an e-mail is sent to the user responsible for verifying the metadata items. This user has now the possibility to check and to correct the metadata items entered by the first user. After this quality check the user releases the items. A periodically planned job

transforms the items in a XML-file. The file is sent as an attachment of an e-mail to the administrator of the central catalogue. Based on the UDK-agent as described in Swoboda et al. [1999] the XML-file can be imported in the central catalogue and the corresponding UDK-objects are entered into the database.

4.3 Evaluation

An evaluation of our solution can be summarised as follows:

Technology and Administration :

- No additional software has to be installed at desktop computers. The installation of the form only affects MS Exchange server.
- Data security, back-ups and user profiles are managed as in Outlook and Exchange Server.
- No manual effort for transforming the metadata items entered in Outlook into UDK-objects.

Consequently, there is only a minimal additional effort for the system administrators.

Usability:

- Users are familiar with the basic functionality of MS Outlook.
- The Outlook-form offers a lot of explanatory messages and data checks in order to support the user.
- A simple workflow application automates and controls the data maintenance process.
- Functionality is limited to the requirements of casual users.
- Responsibility in quality control is with the data generating units.

Due to these facts we can assume that most of the users will accept the software tool.

The disadvantage of our solution is that the data will be stored locally in the MS Exchange Server and centrally in database of the central catalogue. This is a risk because we have to manage duplicate versions of the database. Due to the reduced functionality the application is not convenient to expert users or administrators of the UDK.

5. CONCLUSIONS

We developed a software tool for entering metadata into the environmental data catalogue UDK as implemented in Germany and Austria. The

software tool supports data maintenance in a non-centralised way. This means that data is entered and maintained by the dataholder directly. Most of the dataholders are casual users. Due to this fact we developed the tool based on MS Outlook and Exchange Server. Outlook is already installed at each workplace and most of the users are familiar with it. A simple workflow application controls the data maintenance process. Due to the high automation, the reduced functionality and the familiarity of the users with Outlook we expect that the learning curve for using the tool is minimal and that the degree of acceptance and the willingness to enter data will be very high. Because the tool is especially created for casual users there is still a need for a tool like WinUDK [Lessing et al. 1995] to cover the requirements of frequent and expert users.

Up to now we only implemented a prototype of the tool. Next this prototype will be installed at one public authority and the advantages and disadvantages of our solution will be analysed in more detail. We are sure that this analysis will confirm our strategy of unifying all information management tasks and office tasks in a single standard tool.

6. ACKNOWLEDGEMENTS

The authors wish to thank the Ministry of Environment and Forestry of Rhineland-Palatinate for funding the project.

7. REFERENCES

- Byrne, R. Anwendungsentwicklung mit Outlook 2000, Microsoft Press, 1999.
- Koordinierungsstelle UDK, Niedersächsischen Umweltministerium: Die XML-Schnittstellen des UDK 4.2, 2001
- Lessing, H., O. Günther, W. Swoboda, Ein objektorientiertes Klassenkonzept für den Umwelt-Datenkatalog (UDK), Tagungsband des 9. Internationalen Symposiums "Umweltinformatik 95" der Gesellschaft für Informatik, Berlin, 1995.
- Microsoft Corporation. Product information. <http://www.microsoft.com>, 2000.
- Swoboda, W., F. Kruse, R. Nikolai, W. Kazakos, D. Nyhuis and H. Rousselle, The UDK Approach: the 4th Generation of an Environmental Data Catalogue Introduced in Austria and Germany. Proceedings of the Third IEEE Meta-Data Conference, Bethesda, Maryland, USA (1999).