

Der Windows UDK 4.3 auf Basis der SOAP-Technologie

Fred Kruse¹, Marco Eichler²

Abstract

Der Umweltdatenkatalog (UDK) hat sich zum Quasi-Standard für Metadaten im Umweltbereich in Deutschland und Österreich entwickelt. Seit 1998 ist der UDK in der Version 4 im Produktionsbetrieb. Die Erfassungskomponente, der Windows-UDK, wird in der Version 4.3 zu einer Drei-Schichten-Architektur ausgebaut, wobei die Kommunikation zwischen dem UDK-Client und dem UDK-Server auf dem SOAP-Protokoll basiert.

Einleitung

Der Umweltdatenkatalog (UDK) ist ein Metainformationssystem zur Erfassung von Datenbeständen, die in den Umweltverwaltungen von Deutschland und Österreich vorgehalten werden. Von 1991 - 1995 wurde im Niedersächsischen Umweltministerium das Grundkonzept sowie eine einsatzfähige Software im Rahmen eines vom Bund geförderten Forschungs- und Entwicklungsvorhabens entwickelt.^{1,2} In Deutschland beruht die weitere Entwicklung und Einführung dieses Systems auf einer Verwaltungsvereinbarung, die seit 1.1.1996 in Kraft ist und vom Bund und mittlerweile 15 Bundesländern unterzeichnet wurde.

Seit 1998 ist der Windows-UDK^{3,4} mit seinen Versionen 4.0, 4.1 und 4.2 als Erfassungskomponente des UDK im Produktionsbetrieb. Obwohl Internet-Technologien im Umweltinformationsbereich eine immer größere Rolle spielen, ist der Windows-UDK durch die breite Palette seiner Funktionen die zentrale Anwendung im UDK-Umfeld.

Die ständig steigende Anzahl der Client-Installationen insbesondere im LAN-Bereich und der Druck aus dem Bereich der Internet-Entwicklungen machten es notwendig die Technologie des Windows-UDK weiterzuentwickeln.

Mit dem Windows-UDK 4.3 stellen wir hier eine Client-Server-Anwendung auf Grundlage des Drei-Schichten-Modells vor, das auf der Kommunikation mittels des

¹ Niedersächsisches Umweltministerium, Koordinierungsstelle UDK, Archivstraße 2, D-30169 Hannover, e-mail: udk@numis.niedersachsen.de

² BVG mbH, Menkestr. 7, D-26419 Schortens, e-mail: udk@bvg-fri.de

Simple Object Access Protocol (SOAP) aufbaut. Über diese Technologie hält sich der Win-UDK offen, zukünftig auf die .NET Plattform der Firma Microsoft aufzusetzen.

Der Windows-UDK

Der Windows-UDK ist zur Zeit die zentrale Anwendung im UDK-Umfeld. Er wurde in erster Linie als Erfassungs- und Administrationskomponente des UDK entwickelt, wird aber auch insbesondere von den Erfassern zur Recherche genutzt.

Auf der Ebene der Datenerfassung dient er dazu, sowohl Datenobjekte und Adressen als auch ganze Kataloge neu zu erfassen und zu pflegen.

Er verfügt weiterhin über umfangreiche Möglichkeiten zur Recherche (erweiterte Suche, Expertensuche, Thesaurussuche, Geo-Suche) und zur strukturierten Darstellung (UDK-Explorer).

Der Windows-UDK verfügt über eine Nutzerverwaltung, die über die Vergabe von Schreibrechten, den Schutz vor unautorisierten Änderungen der Dateninhalte sicherstellt, sowie über umfangreiche Möglichkeiten den Katalog zentral als Administrator zu pflegen.

Ein umfangreiches Angebot an Im- und Exportschnittstellen sorgt sowohl für die Kommunikation einzelner UDK-Anwendungen untereinander (Aktualisierungszyklus, Gesamtimport und -export, Teilbaumimport und -export) als auch mit Fremdsystemen (semantischer XML-Export, universelle Importschnittstelle). Alle Im- und Exportschnittstellen sind mit der Version 4.2 auf XML umgestellt worden.

Die DTD des UDK

Seit der Version 4.2 des UDK existiert eine DTD, die das gesamte fachliche Datenmodell des UDK abbildet⁵. Sie enthält nicht nur die Felder der UDK-Objekte sondern auch die der UDK-Adressen und UDK-Kataloge. Weiterhin werden alle Beziehungen zwischen den Objekten, zwischen den Adressen und zwischen Objekten und Adressen abgebildet.

In den Objekten werden die Felder in den vom UDK definierten Gruppen wie Fachbezug, Raumbezug usw. zusammengefasst. Der Fachbezug unterscheidet die einzelnen UDK-Klassen.

Die DTD wird im UDK 4.2 für die Im- und Export-Schnittstelle genutzt. Im UDK 4.3 ist sie die Beschreibung der Detaildarstellung der vom Server an den Client zurückgesandten Daten.

Von der Zwei-Schichten- zur Drei-Schichten-Architektur

Der Windows-UDK wird sowohl als lokale Standalone-Version als auch im Client-Server-Betrieb eingesetzt. Bis zur Version 4.2 beruht er auf der Zwei-Schichten-Architektur. Das heißt, es besteht eine Trennung zwischen Datenschicht und Anwendung. Durch diese Trennung ist es möglich, Datenbanksysteme unterschiedlicher Hersteller zu nutzen. Als mögliche Datenbankmanagementsysteme unterstützt der UDK Oracle, SQL-Server, Informix, Ingres und MS-Access.

Trotz eines ausgereiften Aktualisierungszyklus ist die Administration vieler Einzelinstallationen auf Grundlage lokaler MS-Access-Datenbanken sehr aufwendig. Häufig wird daher der Vorzug einer Installation einer zentralen Datenbank im LAN gegeben, auf die eine große Anzahl von UDK-Clients zugreift. Bedingt durch die Zwei-Schichten-Architektur ist für Versionen bis 4.2 aber dennoch der Aufwand sehr groß, die UDK-Anwendungen zu installieren, zu konfigurieren und das Versionsmanagement durchzuführen.

In der Version 4.3 ist der Windows-UDK daher in einer Drei-Schichten-Architektur realisiert. Die erste Schicht der alten Zwei-Schichten-Architektur wird so geteilt, dass eine Serverschicht, die Businesslogik des UDK abdeckt, und eine Clientschicht, das Userinterface bildet. Die Datenbankschicht bleibt als dritte Schicht erhalten. Durch diese Entwicklung werden zwei Ziele verfolgt:

- Auf den Clients wird nur eine schmale Installation ausgeführt und die gesamten administrativen Aufgaben konzentrieren sich auf den Server. Die Client-Installation kann zentral auf dem Server abgelegt und auch von Nutzern mit geringen administrativen Rechten selbst installiert werden. Das Versionsmanagement wird zentral von einer Serverkomponente verwaltet und Versionsupdates quasi automatisch auf den Clients ausgeführt. Durch diese Architektur wird daher der Administrationsaufwand wesentlich reduziert. Für die Nutzung im LAN bedeutet dies, dass die Installation des Windows-UDK auch an Arbeitsplätzen wirtschaftlich und organisatorisch leistbar wird, die den UDK nur zur Recherche benötigen. Eine solche Installation in den lokalen Netzen hat den Vorteil, dass die sehr gut ausgebauten Recherchemöglichkeiten und die ansprechend aufbereitete Benutzeroberfläche des Windows-UDK von nahezu allen Mitarbeitern eines Hauses genutzt werden können.
- Durch die Verwendung des von SOAP/HTTP als Kommunikationsprotokoll ist eine Umstellung und damit eine Verwendung des Win-UDK-Clients als Webkomponente mit geringem Aufwand möglich.

Architektur des UDK 4.3

Der UDK 4.3 setzt sich aus einem UDK-Server und einem UDK-Client zusammen. Der UDK-Server greift über eine ADO-Schnittstelle (ADO = ActiveX Data Objects) auf

die zentrale UDK-Datenbank zu. Unterstützt werden die Datenbanksysteme Oracle, SQL-Server, Informix, Ingres und MS-Access.

Auf dem Server liegt neben der UDK-Server-Anwendung auch die Client-Installationsroutine. Sie kann über das Intranet den Anwendern zugänglich gemacht werden. Dadurch wird es potentiellen Anwendern möglich, ohne größeren administrativen Aufwand den UDK-Client zu installieren.

Bei der Verbindung des UDK-Clients mit dem UDK-Server wird jeweils eine Versionskontrolle und gegebenenfalls eine Update des Clients durchgeführt.

Die Verbindung zwischen UDK-Client und UDK-Server wird über SOAP realisiert. Dieses Protokoll hat den Vorteil, dass es sowohl standardisierte Schnittstellen schafft als auch über das HTTP-Port mit anderen Rechnern kommuniziert. Dadurch sind Verbindungen auch über Firewalls hinaus möglich. Es wird daher möglich sein, auch über das Internet zwischen UDK-Server und UDK-Client zu kommunizieren.

Nachteil des SOAP-Protokolls ist, die durch den textbasierten Aufbau relativ schlechte Performance. Der Windows-UDK zeichnet sich durch große Benutzerfreundlichkeit aus, die sich in Form vieler Auswahllisten, einer kontextsensitiven Online-Erfassungsanleitung sowie einer umfangreichen Thesaurusfunktionalität äußert. Um diese Funktionalitäten zu gewährleisten, sind häufige performante Zugriffe auf Daten notwendig, die nur relativ selten aktualisiert werden. Diese Daten werden lokal in einer Systemdatenbank gehalten, auf die der UDK-Client direkt über ADO zugreift. Bei jedem Start des UDK-Client, wird die Aktualität der Systemdaten geprüft. Gegebenenfalls werden sie aus der zentralen Datenbank aktualisiert.

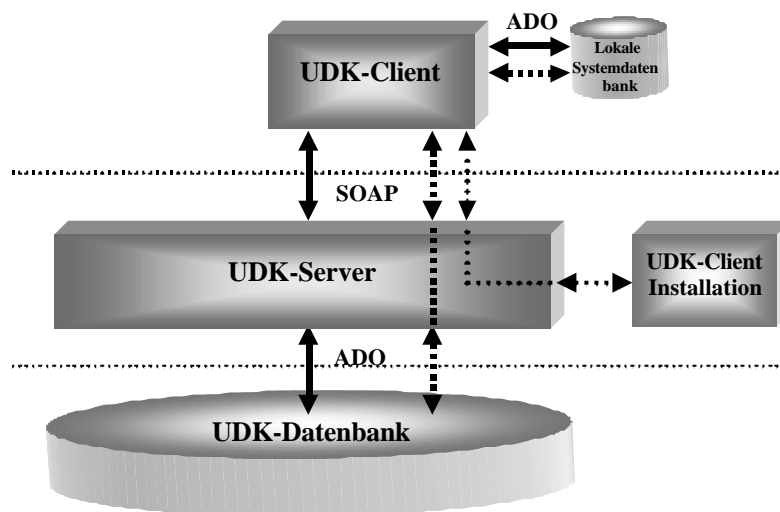


Abbildung 1: Die Architektur des UDK 4.3

Kommunikation über SOAP-Technologie

Bei der Weiterentwicklung des UDK wurde auf die neusten technischen Entwicklungen auf dem Gebiet der Anwendungskommunikation gesetzt. Die Kommunikation zwischen der User- und Business-Schicht basiert im UDK 4.3 auf dem Simple Object Access Protocol (SOAP). SOAP beruht auf XML (eXtended Markup Language). Es bietet einen allgemeinen RPC-Mechanismus (Remote Procedure Call), der den Internetstandard HTTP als Protokoll nutzt. Es stellt einen einfachen Mechanismus zum Austausch zwischen verteilten Anwendungen, und damit von dezentral gehaltenen Daten, bereit.

Für den UDK bieten sich durch die Nutzung von SOAP verschiedene Vorteile:

- Es werden definierte Schnittstellen geschaffen, über die andere Anwendungen aus dem Bereich der Umweltinformationssysteme direkt die Funktionalität des UDK nutzen können.
- Es ist denkbar das zukünftig UDK-Clients auch über das Internet auf einen UDK-Server (Dienst) zugreifen können. Dies ist heute schwierig zu realisieren, da die meisten Clients hinter Firewalls liegen, die nur für das Http-Protokoll geöffnet sind.
- SOAP unterstützt unter anderem die Kommunikation von Anwendungen, die in unterschiedlichen Programmiersprachen entwickelt wurden. Es ist daher ein Konzept denkbar, dass die mittelfristig Zusammenführung des Windows-UDK mit den in Java entwickelten Komponenten des WWW-UDK⁶ verfolgt.
- Verwendung von standardisierten Sicherheitsprotokollen wie SSL als Sicherheitsstruktur.
- Die .NET Strategie der Firma Microsoft baut auf SOAP auf. Eine Beurteilung dieser Entwicklung bezüglich der Relevanz für das UDK-Umfeld ist zur Zeit zwar verfrüht, aber es besteht die Möglichkeit, dass mittelfristig über diese Technologien der Windows-UDK mit geringem Aufwand zu einer Internetanwendung ausgebaut werden kann.

Über die Nutzung der SOAP-Technologie ist also zu erwarten, dass der Windows-UDK über eine reine Client-Server-Anwendung hinauswächst. Mit diesem zweiten Standbein neben den Java-Entwicklungen des WWW-UDK hält sich die UDK-Gemeinde offen, zügig auf zukünftige Entwicklungen im Bereich der Internet-Technologie zu reagieren.

Literatur:

- ¹ T. Schütz, H. Lessing, *Metainformation von Umwelt-Datenobjekten - Zum Datenmodell des Umwelt-Datenkataloges Niedersachsens*, in: "Informatik für den Umweltschutz", 7. Symposium (Ulm 1993), Springer-Verlag, Berlin (1993)

-
- ² H. Lessing, T. Schütz, *Der Umwelt-Datenkatalog als Instrument zur Steuerung von Informationsflüssen*, in: "Informatik für den Umweltschutz", 8. Symposium (Hamburg 1994), Metropolis-Verlag, Marburg (1994)
- ³ W. Swoboda, F. Kruse, D. Nyhuis, H. Rousselle, *Die Neukonzeption des Umweltdatenkatalogs*, in: "Umweltinformatik '98" (Bremen 1998), Metropolis-Verlag, Marburg (1998)
- ⁴ W. Swoboda, F. Kruse, R. Legat, R. Nikolai, S. Behrens, *Harmonisierter Zugang zu Umweltinformationen für Öffentlichkeit, Politik und Planung: Der Umweltdatenkatalog UDK im Einsatz*, in: "Umweltinformatik '00" (Bonn 2000), Metropolis-Verlag, Marburg (2000)
- ⁵ F. Kruse, M. Eichler, U. Freitag, T. Sattler, *Die XML-Schnittstelle des UDK 4.2 – Der Schlüssel zur Integration von Umweltsystemen*, in: "Neue Methoden für das Wissensmanagement im Umweltschutz", 4. Workshop des GI-Arbeitskreises Hypermedia im Umweltschutz und Workshop 3 der GI-Initiative Environmental Markup Language (Ulm 2001), Metropolis-Verlag, Marburg (2001)
- ⁶ R. Nikolai, W. Kazakos, R. Kramer, S. Behrens, W. Swoboda, F. Kruse, *WWW-UDK 4.0: Die neue Generation eines Web-Portals zu deutschen und österreichischen Umweltdaten*, in: "Umweltinformatik '99", Metropolis-Verlag, Marburg (1999)