

WWW-UDK 4.0: Die neue Generation eines Web-Portals zu deutschen und österreichischen Umweltdaten

Ralf Nikolai, Wassili Kazakos, Ralf Kramer¹, Sven Behrens²,
Walter Swoboda, Fred Kruse³

Abstract

Since its introduction in 1995 the environmental data catalogue UDK (Umweltdatenkatalog) has become a de-facto standard for environmental meta-information in German and Austrian governmental organisations. During the last year, the fourth generation of WWW-UDK, a tool for publishing UDK data on the Web, was developed based on a careful user requirement analysis. A complete redesign of the software makes possible that on the one hand advanced Internet technologies could be employed and on the other hand the various requirements of the different user groups could be fulfilled. To do so, three main components have been developed: (1) HTML-UDK which provides full retrieval functionality and requires minimal client equipment only. (2) J-UDK, an integrated Java applet of graphical, interactive tools for advanced users. (3) Virtual UDK which provides for the first time searches over distributed UDK data of different catalogues.

Furthermore the interaction of WWW-UDK and WinUDK is presented.

1. Einleitung

1.1 Motivation und Hintergrund

Seit 1990 haben die Mitgliedsstaaten der Europäischen Union die Anforderungen der EU-Direktive (90/313/EC) des freien Zugangs zu von Behörden erhobenen oder verwalteten Umweltinformationen zu erfüllen. Wesentliche Anforderung an einen solchen Zugang zu den heterogenen und verteilt vorliegenden Umweltinformationen ist

¹ Forschungszentrum Informatik (FZI), Haid-und-Neustr. 10-14, D-76131 Karlsruhe, Tel. ++49 721 9654 744 / 712, {nikolai, kazakos}@fzi.de

² DISY Informationssysteme GmbH, Stephaniestr. 30, D-76133 Karlsruhe, Tel ++49 721 9 290 290, behrens@disy.net

³ Niedersächsisches Umweltministerium, UDK-Koordinierungsstelle, Archivstr. 2, D-30169 Hannover, Tel. ++49 511 120 3697, udk@mu.niedersachsen.de

ein Werkzeug, das das gezielte und zuverlässige Wiederauffinden ermöglicht. In der ersten Phase des UDK-Projektes (UDK = Umweltdatenkatalog) von 1991 bis 1995 wurde daher ein Grundkonzept sowie eine erste einsatzfähige Software für ein Katalogsystem entwickelt. Dieses Katalogsystem erlaubt die Erfassung und die Verwaltung von Umwelt-Metadaten, also Daten die Umweltdaten beschreiben, in einem eigenständigen System.

Für die Erfassung der Metadaten wurde eine PC-basierte Software verwendet. Bereits Ende 1995 erkannte die Landesanstalt für Umweltschutz (LfU), Baden-Württemberg, das Potential der Web-Technologie, um die Metadaten des UDK für einem möglichst großen Kreis von Benutzern, die auf der Suche nach Umweltinformationen sind, recherchierbar zu machen. Im Auftrag der LfU entwickelte das FZI eine erste Version des WWW-UDK, der mit einer einfachen HTML-Oberfläche verschiedene Suchmöglichkeiten in den von einer relationalen Datenbank verwalteten UDK-Daten anbot. Diese Lösung übernahmen und erweiterten in den folgenden drei Jahren Österreich, Niedersachsen, Sachsen-Anhalt, Brandenburg und Hessen (Kramer et al. 1996, Kramer et al. 1997).

1.2 WinUDK 4.0 und WWW-UDK 4.0

Die Ergebnisse der ersten Phase, die Erfahrungen der letzten Jahre und eine ausführliche, systematische Benutzeranforderungsanalyse führten zu einem weitgehenden Re-Design des Datenmodells und der Windows- und Web-basierten Anwendungen, WinUDK und WWW-UDK.

Hauptziel des Re-Designs der WinUDK-Anwendung (Swoboda et al., 1998) war es, eine flexibleres und benutzerfreundlicheres Werkzeug als in den vorherigen Versionen bereitzustellen, um damit die Akzeptanz eines Umwelt-Katalogsystems bei den Metadaten-Lieferanten in den Behörden zu erhöhen. Diese Akzeptanz sollte wesentlicher Faktor sein, um sowohl die Qualität als auch die Quantität der UDK-Daten zu verbessern.

Der Komfort der Recherche in den UDK-Daten stand bei dem Re-Design des WWW-UDK, das in diesem Beitrag vorgestellt wird, im Mittelpunkt. Eine intuitive Benutzerschnittstelle sowohl für den UDK-Neuling als auch ausgefeilte graphische Werkzeuge für den regelmäßigen WWW-UDK-Benutzer sollten hierzu beitragen. Desweiteren soll der "Virtuelle UDK" erstmals eine gleichzeitige Suche in den von den Bundesländern und dem Bund dezentral verwalteten und bereitgestellten Katalogen ermöglichen. Die Entscheidung für eine weitgehende Neuimplementierung des WWW-UDK erlaubte auch, moderne Software-Entwicklungswerkzeuge (Java, Java-Servlets) einzusetzen, um so die Investitionen langfristig zu schützen.

2. Ergebnisse der Benutzeranforderungsanalyse

Der mehrjährige Einsatz sowohl des PC-basierten UDK 2.0 bzw. 3.x als auch des Web-basierten WWW-UDK 3.x bot exzellente Voraussetzungen, um Erfahrungen bei der Verwendung eines Kataloges für Umweltdaten unter realen Einsatzbedingungen zu sammeln. Die Benutzeranforderungen wurden auf mehreren Workshops und per E-Mail-Umfragen gesammelt. Wesentliche Ergebnisse in Bezug auf den WWW-UDK sind:

- Informationssuchende wollen schnellen, intuitiven und zielgerichteten Zugang zu den Umweltdaten. In den meisten Fällen ist die Bereitschaft nicht vorhanden, erst zu lernen, was ein Thesaurus ist, bevor mit diesem Hilfsmittel recherchiert werden kann.
- Die meisten Informationssuchenden innerhalb von Behörden, die Web-Zugang haben, benötigen (technisch) einfache HTML-Anwendungen, die möglichst keine Umkonfigurationen an Firewalls (z.B. das Öffnen von Ports) erfordern.
- Zur komfortablen Recherche mit Windows-ähnlicher graphischer Benutzungsoberfläche (z.B. bei einer karten-basierten Suche) kann es eine fortschrittliche Java-Anwendung ergänzend zur HTML-Version geben. Benutzer, die dies fordern, haben in der Regel Web-Browser der neuesten Generation und ausreichendes technisches Know-How.
- Der generelle Ansatz in einem Staat wie der Bundesrepublik mit föderaler Struktur ist es, Umwelt-Metadaten dezentral zu sammeln und zu pflegen. D.h. jedes der an der UDK-Kooperation teilnehmenden Bundesländer hat seinen eigenen Katalog. Da es in der Natur von Umweltfragen liegt, daß die zu ihrer Beantwortung nötigen Informationen häufig nicht nur aus einem Bundesland kommen (“Wie hat sich die Wasserqualität des Rheines in den letzten 10 Jahren verändert?”) wird eine katalogübergreifende - also verteilte - Suche benötigt.
- Der WWW-UDK soll in den einzelnen Ländern und für die verteilte Suche nicht einfach nur eine WWW-basierte Datenbankabfrage-Oberfläche sein. Statt dessen soll er ein “Portal” zu den Umweltinformationen des Bundes und der Länder sein. Der Begriff “Portal” stammt dabei aus dem Bereich der Web-Suchmaschinen und bezeichnet den Eingangsbereich zu verschiedenen, auch aggregierten, Inhalten und Diensten rund um ein Thema. So soll die Vernetzung der Umweltinformationen mit anderen Informationen, z.B. Rechtsinformationen und -vorschriften, ausgedrückt werden können. Es soll aktuelle Informationen rund um den UDK geben. Von den Metadaten soll möglichst einfach und schnell auf die eigentlichen Daten zugegriffen werden können. Kompetente Ansprechpartner sollen über Email weitergehende Auskünfte erteilen. Natürlich soll der WWW-UDK in das spezifische Bundes- bzw. Landes-WWW-Layout integriert werden können. Bisher ist der WWW-UDK Österreichs in dieser Hinsicht eines der wesentlichen Vorbilder (Umweltbundesamt Österreich, 1998).
- Insbesondere in den unteren Verwaltungsbehörden haben nur wenige Metadaten-Anbieter bereits Zugang zu Web-Diensten außerhalb ihrer Dienststelle. Sie

bevorzugen die off-line Erfassung von Metadaten, die durch einen Kollegen oder Vorgesetzten autorisiert werden, bevor sie der Öffentlichkeit frei zugänglich gemacht werden. Somit wird weiterhin eine Windows-basierte Erfassungssoftware benötigt. WWW-UDK wird von den meisten Metadaten-Anbietern im wesentlichen als Recherchemöglichkeit gesehen.

3. Architektur des WWW-UDK 4.0

Um die heterogenen Benutzeranforderungen zu erfüllen, wurden zwei Clients entwickelt: (1) HTML-UDK, eine Basisanwendung, die vollständige Retrieval-Funktionalität ähnlich den bekannten Web-Suchdiensten anbietet und minimale technische Anforderungen stellt, und (2) J-UDK, ein wesentlich interaktiveres, graphisches Werkzeug, das in Form eines integrierten Java-Applets bereitgestellt wird.

Die Gesamtarchitektur ist in Abbildung 3, links, dargestellt. Sie entspricht einer klassischen 3-Schichten-Architektur mit Client-Schicht (GUI etc.), Anwendungsschicht (UDK Object and Query Server) und Datenbankschicht. Die Anwendungsschicht bedient dabei die verschiedenen Clients, derzeit HTML-UDK und J-UDK. Die Hauptideen dieser Architektur sind die Wiederverwendung von Komponenten und die Verteilung der Komponenten und Dienste zu ermöglichen sowie unterschiedliche Client- und Server-Plattformen zu unterstützen. Diese Ziele wurden erreicht, indem die Architektur mit Java (als JavaServlet serverseitig, und für J-UDK als Applet clientseitig; Java RMI für die Verteilung, JDBC für den Datenbankzugriff) bzw. HTML-Benutzerschnittstelle (für HTML-UDK) umgesetzt wurde.

4. Die WWW-UDK-Komponenten

WWW-UDK 4.0 soll ab Juni 1999 verfügbar sein. Aktuelle Informationen und Versionen sind über <http://www.fzi.de/dbs/applAreas/udk4.html> zugänglich.

4.1 HTML-UDK: Recherchieren einfach und schnell

Die primäre Benutzergruppe von HTML-UDK sind Benutzer, die Umweltinformationen schnell und einfach finden wollen und bisher keine oder geringe Erfahrungen mit Metainformationen haben. Eine typische Fragestellung könnte z.B. sein, welche Informationen über Bodenverschmutzungen in einer bestimmten Stadt vorliegen.

Diese Benutzergruppe kennt in der Regel bereits klassische Web-Suchdienste, daher orientieren sich Funktionalität und Benutzerführung an diesen Systemen (die Grundidee der HTML-UDK-Benutzerschnittstelle stammt von www.libri.de). Berührungspunkte mit Katalogsystemen sollen somit vermieden werden.

Einfache Suchen ("Schnellsuchen") ermöglichen durch die Eingabe von einem oder mehreren Worten das Auffinden von Informationen (UDK-Objekten, d.h. den ei-

gentlichen Metainformationen, und Adressen). "Expertensuchen" erlauben eine detailliertere Einschränkung der Ergebnismenge durch Raum-, Zeit- und thematische Filter. Diese beiden Suchmodi entsprechen den "einfachen" und "erweiterten" Suchmodi, die von den meisten Web-Suchdiensten angeboten werden, und erlauben einen deskriptiven Zugang zu den Daten. Für Benutzer, die den UDK-Datenbestand explorativ erkunden möchten, wurde ein Zugang über Umweltthemen entworfen. Dieser Zugang entspricht in etwa der Idee des "Browseable Directories" von Yahoo (Yahoo, 1998). Die Kategorien entstammen dem Umweltthesaurus des Umweltbundesamtes, dessen Terme und deren hierarchische Struktur hier Verwendung finden. Da diese Terme bereits zur Indexierung der UDK-Objekte verwendet werden, kann der derart geordnete Zugang über Umweltthemen ohne jeweiligen manuellen Aufwand vollautomatisch bereitgestellt werden.

Alle Suchergebnisse werden zunächst in einer Übersicht aufgelistet, in der neben dem Namen des UDK-Objektes bzw. der UDK-Adresse auch der Typ des Ergebnisses angezeigt wird. Nach der Auswahl eines Ergebnisses werden die vollständigen Informationen angezeigt. Um gleichzeitig die Detailinformationen mehrerer Ergebnisse anzuzeigen, gibt es eine Sammelmappe, in der auch über verschiedene Suchanfragen hinweg Ergebnisse z.B. für eine Druckerausgabe gesammelt werden können.

Die Vernetzung der UDK-Daten zu anderen Datenbeständen wird hauptsächlich über URLs hergestellt. Das sind zum einen URLs, die direkt von den UDK-Objekten zu den beschriebenen Umweltinformationen – soweit im Web verfügbar – verweisen. Zum anderen können Regeln angegeben werden, wie aus einzelnen Feldern URLs generiert werden, die auf andere Systeme verweisen. In Österreich verweist z.B. das Feld "gesetzliche Grundlage" auf entsprechende Einträge im Rechtsinformationssystem und damit direkt auf die entsprechenden Gesetzestexte (Legat et al., 1999).

4.2 J-UDK: Web-Anwendung mit Windows Look-and-Feel

Die Gestaltungsmöglichkeiten von Benutzeroberflächen, die HTML als Markup-Sprache bietet, sind im Hinblick auf die vom Benutzer gewohnten fensterbasierten, wesentlich interaktiveren Programme oft nicht ausreichend. Für aufwendigere und komplexere Recherchen im UDK werden Werkzeuge benötigt, deren Funktionalität über HTML-UDK hinausgehen. Um solche Werkzeuge, wie z.B. eine karten-basierte Suche, sowohl im behördeninternen Intranet als auch im Internet anzubieten, wurde J-UDK entwickelt. Dabei wurden bereits für WWW-UDK 3.0 bzw. WebCDS (Kazakos et al., 1998; Nikolai et al., 1999) entwickelte Komponenten angepaßt, weiterentwickelt und in ein Gesamtsystem, J-UDK, integriert.

Die Retrieval-Funktionalität des J-UDK orientiert sich hierbei am WinUDK (Swo-boda et al., 1998) und erweitert das Paradigma auf die Suche im Intra- bzw. Internet: Neben den auch schon im HTML-UDK verfügbaren Suchmöglichkeiten, bietet der J-UDK einen kartenbasierten Zugang, einen Zugang über die Struktur der Objekte und Adressen sowie einen Zugang über Thesaurusbegriffe, die zur Verschlagwortung der

UDK-Objekte und -Adressen benutzt werden. Anfrageformulierung, -reformulierung und Ergebnisdarstellung werden vollständig vom J-UDK übernommen (s. Abbildung 1).

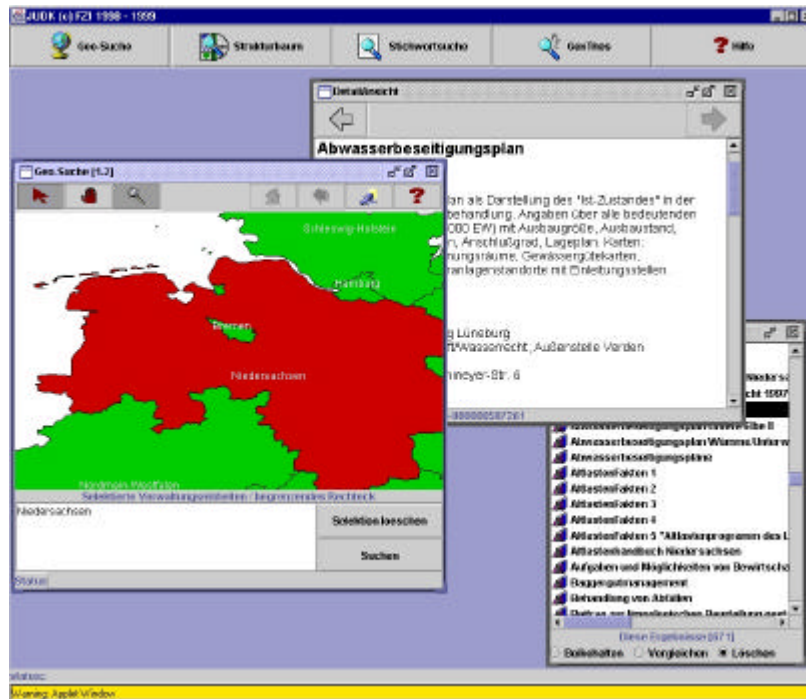


Abbildung 1: J-UDK Desktop mit Geo-Suche (GeoView), Ergebnisliste und Detaildarstellung des UDK-Objektes „Abwasserbeseitigungsplan“.

Über eine zentrale Navigationsleiste sind sämtliche im J-UDK integrierten Werkzeuge sowie die Hilfe erreichbar. Alle Suchwerkzeuge, sowie die Ergebnis- und Hilfeanzeige können im sogenannten J-UDK Desktop beliebig angeordnet und je nach Bedarf aktiviert oder deaktiviert werden. Im unteren Teil des Desktops wird nach jeder Benutzerinteraktion der aktuelle Status des J-UDK angezeigt.

Die Stichwortsuche ermöglicht die bereits aus dem HTML-UDK bekannte klassische Suche nach UDK-Objekten und -Adressen. Das Fenster wird durch Reiter unterteilt, über die der Benutzer entsprechend der HTML-Expertensuche die Suche weiter einschränken kann (Suchfilter).

Der Strukturbaum wiederum ordnet die Objekte und Adressen anhand ihrer Hierarchie im UDK. Der Benutzer kann hiermit durch die Ober- und Unterobjekte bzw. -Adressen durch den kompletten UDK navigieren, wodurch eine weitere explorative Sicht auf die UDK-Daten ermöglicht wird.

Über die kartenbasierte Suche kann der Benutzer den räumlichen Bezug der gesuchten Objekte einschränken. Das für den UDK angepaßte, ursprünglich für den WebCDS entwickelte GeoView-Modul enthält jeweils für die verschiedenen Verwal-

tungsebenen Karten Deutschlands bzw. Österreichs. Es ermöglicht einfache Karten-Viewer-Funktionalität (z.B. zoomen, verschieben) sowie die intuitive Markierung des Raumbezuges von Interesse durch Auswahl der Verwaltungseinheiten bzw. Aufziehen einer Bounding-Box.

GenThes, ein allgemeiner Thesaurus-Browser (Nikolai et al., 1999), wurde ebenfalls für den J-UDK angepaßt und integriert. Damit wird eine thesaurus-basierte Suche ermöglicht, die einerseits mit der Navigation in der hierarchischen Struktur des Thesaurus einen Zugang ähnlich den Umweltthemen des HTML-UDK (vgl. Abschnitt 5.1) ermöglicht. Allerdings ermöglichte Java eine optische bessere und interaktivere Darstellung. Desweiteren kann der Thesaurus weitergehend betrachtet und durchsucht werden (unter Berücksichtigung der Synonyme, Ober-, Unter- und verwandten Begriffe), so daß der Benutzer eine wichtige Unterstützung bei der Formulierung seiner Anfrage erhält.

Die Ergebnisse einer Suche werden unabhängig vom ausgewählten Werkzeug in einer Ergebnisliste dargestellt. Diese unterstützt zusätzlich die flexible Kombination von Suchergebnissen. So kann der Benutzer einstellen, ob er die alten Suchergebnisse bei einer erneuten Suche beibehalten (entspricht einer Oder-Verknüpfung), mit den neuen Ergebnissen konjunkt miteinander verknüpfen (Und-Verknüpfung) oder ersetzen möchte. Dies ermöglicht z.B. eine kombinierte Suche nach allen UDK-Objekten im Stadtkreis Karlsruhe (spezifiziert über GeoView), die mit Abwasser (ausgewählt in GenThes) verschlagwortet sind.

Interessieren den Benutzer eines oder mehrere der Ergebnisse, so kann er sich diese in der Detailansicht anzeigen lassen. Diese entspricht der auch im HTML-UDK eingesetzten Detailansicht und wird in einem in Java realisierten HTML-Ansichtsfenster angezeigt. Innerhalb dieses Fenster kann sich der Benutzer über HTML-Verknüpfungen z.B. Ober- und Unterobjekte anzeigen lassen.

Da das Drucken aus einem Java-Applet aus Sicherheitsgründen nur für zertifizierte Applets möglich ist, bedient sich der J-UDK hierfür des HTML-UDK, der dargestellte Ergebnisse mit Hilfe der Browser-Funktionalität ausdrucken kann.

4.3 Virtueller UDK: Verteilte Suche in Bund- und Länderkatalogen

Die in Abschnitt 2. aufgrund der eigenständigen Kataloge in den verschiedenen Bundesländern – es gibt keinen Zentralkatalog mit allen Daten sämtlicher Kataloge - geforderte verteilte Suche wurde in einer ersten Phase für den HTML-UDK realisiert. Der "Virtuelle UDK" ermöglicht es dem Informationssuchenden, in den Datenbeständen der verteilten Kataloge zu suchen, so als ob sie in einem zentralen Katalog verwaltet werden würden. Ziel der ersten Phase ist es, Erfahrungen mit einer verteilten Suche zu sammeln und die Benutzerakzeptanz zu testen.

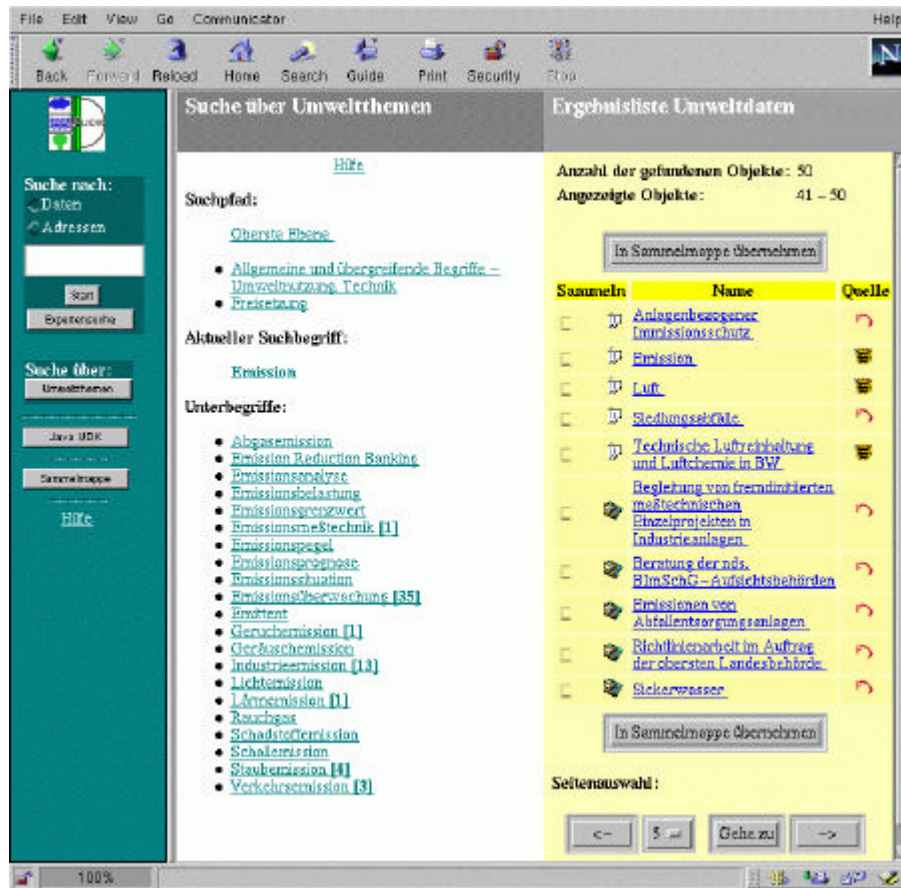


Abbildung 2: Suche über Umweltthemen im Virtuellen UDK (Zugriff auf Daten aus Niedersachsen und Baden-Württemberg)

In der ersten Phase des Virtuellen UDK hat der Benutzer keine Möglichkeit, einzelne Kataloge auszuwählen. Der Administrator des Virtuellen UDK-Servers stellt in einer Konfigurationsdatei die Adressen aller UDK-Server ein, die im Virtuellen UDK-Server zusammengefaßt werden sollen. Beim Start des Servers holt sich dieser sogenannte entfernte Referenzen (Remote References) auf die einzelnen UDK-Server, um anschließend durch Java Remote Method Invocation (RMI) mit ihnen zu kommunizieren.

Bei Suchanfragen wird vom Virtuellen UDK-Server ein parametrisiertes Anfrageobjekt an jeden dieser UDK-Server geschickt. Der UDK-Server stellt eine Infrastruktur für die Anfrageobjekte bereit. Ein Anfrageobjekt enthält eine vom Datenbankserver unabhängige Beschreibung der Datenbankabfragen, die erst auf dem jeweiligen UDK-Server durch einen Statement-Generator auf den lokal installierten Datenbankserver angepaßt wird. Dadurch ist es möglich, den Virtuellen UDK-Server mit ver-

schiedenen Datenbanksystemen zu verbinden. Da auch die Anfragelogik im Anfrageobjekt gekapselt ist, wirken sich Änderungen nicht auf die einzelnen UDK-Server aus, was den administrativen Aufwand bei zahlreichen an den Virtuellen UDK-Server angeschlossenen Systemen erheblich reduziert.

Wenn ein UDK-Server dem Virtuellen UDK-Server nicht antwortet oder ein Übertragungsfehler auftritt, wird nach mehreren Wiederholungsversuchen dieser Server ignoriert. Sobald die Teilergebnisse beim Virtuellen UDK-Server eintreffen, werden sie sortiert und in einer gemeinsamen Ergebnisliste aufgeführt (s. Abbildung 2). Die Ergebnisquellen werden durch Landeswappen symbolisiert.

Nach der Auswahl eines Ergebnisses für die Detailinformation wird das Anfrageobjekt nur an den UDK-Server geschickt, der tatsächlich weitere Informationen dazu liefern kann. Dieser Server ist eindeutig, da die Datenbestände in den verteilten Katalogen disjunkt sind. Auch alle weiteren Anfragen, z.B. nach unter- oder übergeordneten Objekten werden direkt an den richtigen UDK-Server geschickt.

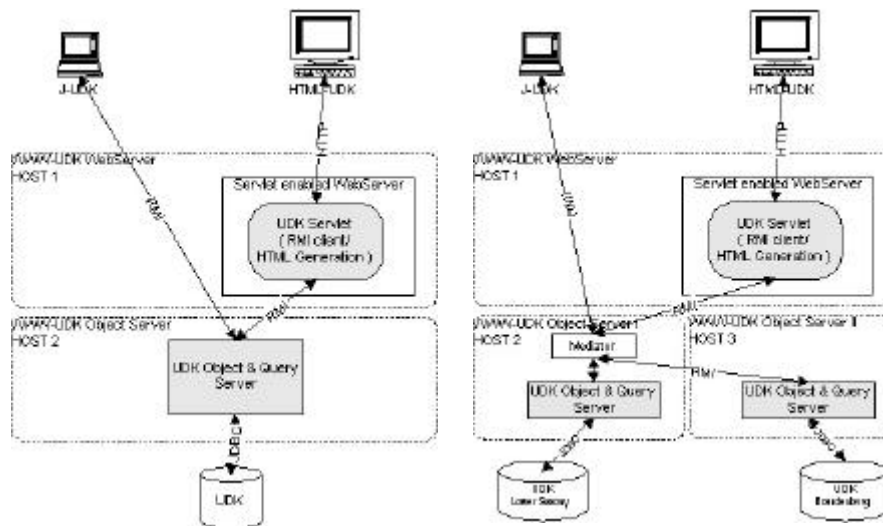


Abbildung 3: Architektur WWW-UDK (links) und Virtueller UDK (rechts)

Aufgrund der Architektur des WWW-UDK sowie der Homogenität der Kataloge (grundsätzlich identisches Datenmodell, das ausschließlich um maximal 3 eigene UDK-Klassen erweitert werden kann), sind komplexe Protokolle wie Z39.50 nicht erforderlich. Stattdessen ist es lediglich erforderlich, der Anwendungsschicht eine weitere Komponente, einen Mediator (siehe Abbildung 3, rechts), hinzuzufügen. Dieser verteilt die Anfragen an die verfügbaren Kataloge und sammelt und sortiert die Ergebnisse. Ein Roboter sammelt periodisch statistische Informationen über die Daten der verschiedenen Kataloge, z.B. welche Thesaurusterme in den einzelnen Katalogen

wie oft zur Indexierung verwandt wurden. Diese Informationen werden benötigt, damit dem Benutzer bei der Anfrageformulierung die gleichen Hilfestellungen gegeben werden können, wie beim Zugriff auf einen einzelnen Katalog. Die eigentlichen Daten werden nicht transferiert.

5. Technische Realisierung

Die gesamte Anwendung wurde in Java implementiert, um unabhängig von Betriebssystem und Computer-Architektur zu sein. Der UDK-Server implementiert die Datenbank- sowie die Anwendungsschicht, HTML-UDK und J-UDK zwei unterschiedliche Client-Schichten, die jeweils auf diesen Schichten aufsetzen.

5.1 UDK-Server

Die Verbindung zu einer Datenbank wird durch den UDK-Server über JDBC hergestellt. JDBC ist der Java-Standard für die Datenbankanbindungen und fast alle Anbieter einer Datenbank liefern mittlerweile den für JDBC benötigten Treiber mit. JDBC ist eine wenig abstrakte Schnittstelle, die immerhin den Verbindungsaufbau, das Abschicken einer SQL-Anweisung und das Auslesen eines Anfrageergebnisses soweit vereinheitlicht, daß bei einfachen SQL-Anfragen die zugrundeliegende Datenbank bei vorhandenem Treiber einfach ausgetauscht werden könnte, ohne auch nur eine Zeile im Java-Programmcode zu ändern.

Trotz aller Bemühungen den SQL-Standard voranzutreiben müssen komplexere Anfragen immer noch auf das verwendete Datenbanksystem angepaßt werden. Der UDK-Server bietet dazu Anfrageobjekten einen Statement-Generator. Dieser erzeugt aus der Datenbankserver-unabhängigen Beschreibung einer Datenbankanfrage eine SQL-Anweisung für das spezifische Datenbanksystem, z.B. Oracle oder Informix.

Das Anfrageobjekt kapselt die Anfragelogik für Anfragen nach einer bestimmten Entität, z.B. für ein UDK-Objekt. Anfrageobjekte können auf dem UDK-Server beliebig viele SQL-Anfragen stellen, Zwischenergebnisse auswerten und in Abhängigkeit davon weitere Anfragen stellen. Die Ergebnisse werden schließlich zu einer Entität zusammengefaßt und als Ergebnis der Anfrage zurückgeschickt.

Eine Entität besteht aus den zusammengefaßten Attributen einer oder mehrerer Datenbanktabellen und, ja nach Untertyp, verschiedenen Zugriffsmethoden auf die Attribute. Um die Anwendungslogik von der Darstellung zu entkoppeln, besitzt eine Entität keine Methode für die Visualisierung. Diese Trennung ermöglicht die Wiederverwendung der Anwendungslogik in verschiedenen Client-Anwendungen.

5.2 HTML-UDK

Eine dieser Client-Anwendungen stellt der HTML-UDK dar. Dieser ist eine Anwendung, die zwar auf einem Server läuft, dort aber dynamisch HTML-Seiten z.B. aus den Entitäten erzeugt und diese über einen HTTP-Server an HTML-Browser weiterreicht. Der HTML-UDK wurde als HTTP-Servlet realisiert. Servlets basieren auf der Java-Servlet API von Sun und stellen eine Alternative zum CGI dar. Vorteile von Servlets sind, neben den Java-eigenen Vorteilen, z.B. bessere Performance und einfacheres Session-Management, da die Servlets im Gegensatz zu CGI-Anwendungen nicht für jede Anfrage neu gestartet werden müssen. HTTP-Servlets werden inzwischen von fast allen HTTP-Servern direkt oder indirekt unterstützt.

Die vom HTML-UDK erzeugten HTML-Seiten sind template-basiert und können durch eigene Footer und Header sowie durch eigene Farbedefinitionen individuell angepaßt werden. Die Seitengenerierung ist unabhängig vom Servlet implementiert worden und kann dadurch für den J-UDK wiederverwendet werden, der die Detailinformationen ebenfalls als HTML-Seite darstellt.

Der HTML-UDK verzichtet auf den Einsatz von Javascript, um nur minimale Browseranforderungen stellen zu können. Javascript bietet Möglichkeiten für interaktivere HTML-Benutzeroberflächen, ist aber von den unterschiedlichen Browserherstellern unterschiedlich implementiert. Dadurch müßte häufig ein eigenes Javascript für jeden Browser erstellt werden. Zudem ist Javascript immer noch nicht sicher und wird deshalb häufig durch Browsereinstellungen nicht ausgeführt oder durch eine Firewall gar nicht erst durchgelassen.

Für die Seitengestaltung wird reines HTML 3.2 verwendet. Wesentliche Gestaltungsmittel sind Frames und Tabellen. Da HTML von den verschiedenen Browsern unterschiedlich dargestellt wird, lassen sich je nach Bildschirmauflösung und Schrift-Einstellungen einige Eigenwilligkeiten im Layout ohne die Verwendung von Cascading Style Sheets nicht verhindern. Das Layout wurde für eine Mindestauflösung von 800*600 Punkten bei einer Schriftgröße von ca. 10 Punkten optimiert.

5.3 J-UDK

J-UDK wurde als Java Applet realisiert, um der Heterogenität der Client-Plattformen im Internet und Intranet sowie den Sicherheitsanforderungen zu genügen. Für die Gestaltung der Benutzeroberfläche wurde die neue Swing-API, die einen wichtigen Teil der Java Foundations Classes (JFC) bildet (Sun Microsystems, 1999a), verwendet. Swing erlaubt im Gegensatz zum älteren AWT (Abstract Windowing Toolkit) eine komfortablere Gestaltung und leichtere Anpassung der Benutzeroberfläche.

Über eine zentrale Adapter-Klasse wurde die Koordination und Kooperation der einzelnen Werkzeuge sowie des UDK-Servers flexibilisiert, so daß in Zukunft Werkzeuge leicht hinzugefügt oder - bei Bedarf - entfernt werden können. Die Kommuni-

kation mit dem UDK-Server erfolgt über Java RMI. Da hier ggf. Ports für den Firewall-Durchgriff nach außen freigeschaltet werden müssen, mußte eine Workaround implementiert werden, der die freie RMI-Portvergabe auf feste Ports beschränkt.

Neben den vielen Vorteilen, die der Einsatz von Java und vor allem der Swing-API mit sich bringt, sind auch Probleme aufgetreten. Da die neue Java-Version Java Plattform 2 noch nicht ausreichend unterstützt wird, fand die Entwicklung mit JDK 1.1.7 statt. Dennoch sind die Browser-Hersteller mit ihren Entwicklungen noch nicht auf diesem Stand, so daß momentan J-UDK in einem Web-Browser nur mit dem separat zu beziehenden, kostenlosen Java Plug-In 1.1.2 von Sun (Sun Microsystems, 1999b) zu benutzen ist. Dieses Plug-In, das auf den unterschiedlichen Plattformen eine einheitliche virtuelle Maschine bereitstellt, funktioniert einwandfrei mit den unterschiedlichen Versionen des Microsoft Internet Explorers, jedoch erst ab Version 4.6 des Netscape Browsers. Für einen zukünftigen Einsatz als Java Application, also unabhängig von einem Web-Browser, gibt es keinerlei Beschränkungen.

6. Erfassung der Metadaten

Die Erfassung der UDK-Daten erfolgt dezentral über eine Windows-basierte Software (WinUDK 4.0). Jede an der Erfassung beteiligte Stelle wird UDK-Instanz genannt. Die Daten einer Instanz können sowohl in einer lokalen Access-Datenbank als auch über ODBC in einer beliebigen Datenbank gehalten werden. Über einen Aktualisierungszyklus werden die Daten an die Stelle weitergereicht, die den zentralen Datenbestand (Zentralkatalog z.B. eines Bundeslandes) führt. Von dieser zentralen Instanz werden alle durch die Aktualisierung veränderten Daten an alle nachgeordneten Instanzen verteilt (Swoboda et al., 1998).

Um das Volumen des Datentransfers zu minimieren, werden im Aktualisierungszyklus nur die Veränderungen des Datenbestandes weitergegeben. Es gibt zwei Möglichkeiten den Aktualisierungszyklus auszuführen:

1. Die Weitergabe der Daten erfolgt ‚per Hand‘. D.h. sie werden bei jeder Instanz über den WinUDK ausgelesen, per Diskette oder Email an den Zentralkatalog versandt und dort über den WinUDK in die Datenbank eingespielt. Das Verteilen der Daten vom Zentralkatalog an alle Instanzen erfolgt entsprechend.
2. Der Aktualisierungszyklus wird über den UDK-Agenten ausgeführt. Dieser liest die Veränderungen der Daten einer Instanz automatisch aus der Datenbank aus und verschickt sie per Email an den Zentralkatalog. Hier werden sie automatisch vom entsprechenden UDK-Agenten eingelesen. Die Verteilung der gesammelten Veränderungen vom Zentralkatalog an alle nachgeordneten Instanzen erfolgt automatisch in entgegengesetzter Richtung.

Der WWW-UDK wird in den Aktualisierungszyklus eingebunden, in dem er als eine weitere Instanz eingerichtet wird (s. Abbildung 4). Der WWW-UDK greift per JDBC auf die zugehörige Datenbank zu. Für den WinUDK wird ein ODBC-Zugriff realisiert.

Der WinUDK übernimmt die Integration der UDK-Datenbank in den Aktualisierungszyklus. Es ist dabei unerheblich, ob die Aktualisierung händisch oder über den UDK-Agenten realisiert wird.

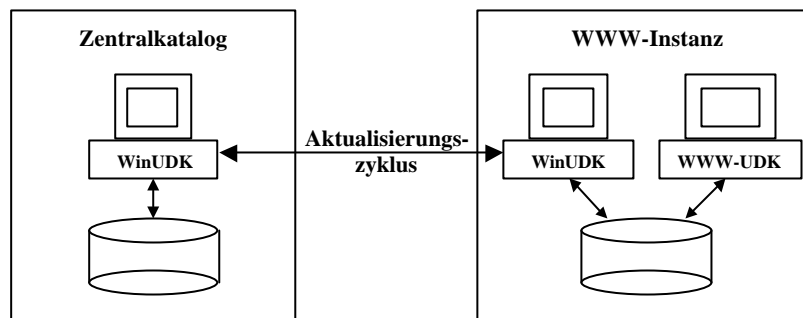


Abbildung 4: Einbindung des WWW-UDK in den Aktualisierungszyklus

Alternativ kann die Datenbank einer WinUDK-Instanz (z.B. Zentralkatalog) gespiegelt und außerhalb einer Firewall für den Zugriff des WWW-UDK installiert werden. Die Pflege der Datenbank kann dann über geeignete Replikationsmechanismen der Datenbankhersteller erfolgen.

7. Zusammenfassung und Ausblick

Der UDK ist inzwischen ein de-facto-Standard für Umweltmetadaten in Deutschland und Österreich geworden. Der Einsatz seit 1995 in verschiedenen Bundesländern bot einen exzellenten Ausgangspunkt, um die Benutzerzufriedenheit und weitergehenden Anforderungen zu analysieren. Die wesentlichen Anforderungen an eine Windows-basierte Erfassung und eine Web-basierte Retrieval-Software resultierten in einer Neukonzeption beider Anwendungen, die die Benutzerakzeptanz und somit die Qualität und die Quantität der Metadaten weiter erhöhen sollen.

Den vielfältigen Anforderungen der heterogenen Benutzergruppe des WWW-UDK wurde durch zwei Anwendungen Rechnung getragen, HTML-UDK und J-UDK. Durch Synergieeffekte innerhalb des Projektes (eine Anwendungsschicht bedient beide Anwendungen) und projektübergreifend (Wiederverwendung bereits existierender Komponenten wie einer kartenbasierter Suche und einem allgemeinen Thesaurusbrowser), konnten die zusätzlichen Kosten in engen Grenzen gehalten werden.

Mit dem "Virtuellen UDK" existiert erstmals eine Möglichkeit in den verteilten Bund- und Länder-Katalogen des UDK gemeinsam zu recherchieren. Somit wird eine wesentliche Basis geboten, um Zugang zu den heterogenen und verteilten Umweltinformationen in Deutschland und in Österreich zu erhalten. Der "Virtuelle UDK" bietet zudem hervorragende Voraussetzungen, der "zentrale Ansprechpartner" für weitergehende Umweltsuchdienste wie z.B. GEIN2000 auf deutscher und GELOS auf internationaler Ebene zu sein.

Die Bereitstellung eines “Portals zu Umweltinformationen” ist allerdings nicht allein durch die Bereitstellung der Software getan. Sie bietet lediglich die Voraussetzungen für eine ständige Erfassung wichtiger Ressourcen, der Qualitätssicherung und Aktualisierung vorhandener Metadatenbestände und der Einrichtung von Querweisen zu anderen relevanten Ressourcen. Im Laufe der nächsten Jahre müssen sich sowohl die Werkzeuge hierfür bewähren, als auch durch eine Steigerung der Anfragen an den UDK der zunehmende Wert der gesammelten Metadaten unter Beweis gestellt werden.

8. Literaturverzeichnis

- Haasis, H.-D. und Ranze, K.C. (Hrsg.) (1998): Umweltinformatik '98. Vernetzte Strukturen in Informatik, Umwelt und Wirtschaft. Proceedings des 12. Internationalen Symposium “Informatik für den Umweltschutz”, Metropolis, 1998.
- Kazakos, W., Kramer, R., Nikolai, R., and Rolker, C. WebCDS - A Java-based Catalogue System for European Environment Data. In Proceedings des International Workshop Issues and Applications of Database Technology (IADT'98), S. 482 – 490, URL: <http://www.fzi.de/dbs/publications/kazakos/IADT98.ps>, 1998.
- Kramer, R., Nikolai, R., Keitel, A., Legat, R. und Zirm, K. (1996): Enhancing the Environmental Data Catalogue UDK for the World Wide Web, in: Umweltinformatik '96. Proc. 10. Internat. Symposium “Informatik für den Umweltschutz”, Metropolis, 1996.
- Kramer, R., Nikolai, R., Koschel, A., Rolker, C., Lockemann, P., Keitel, A., Legat, R. und Zirm, K. (1997): Web-based Environmental Metainformation System, in: ACM SIGMOD Record, Vol. 26, No 1, S. 16-25, New York, 1997.
- Legat, R., Batschi, W.-D., Hashemi-Kepp, H., Kruse, F., Nikolai, R., Nyhuis D, Pultz, F., Stallbaumer, H., Swoboda W. (1999): Der Umweltdatenkatalog in Österreich – 5 Jahre Erfahrungen, erscheint in: Umweltdatenbanken im Web, Proc. Workshop Umweltdatenbanken, Hrsg.: Kramer, R., Hosenfeld, F., Ortleb, H. Metropolis, 1999.
- Nikolai, R., Kramer R., Steinhaus M., Felluga B., Plini, P.. Genthies (1999): A general thesaurus browser for web-based catalogue systems. In Proc. IEEE Meta-Data'99, Bethesda, Maryland, USA, April 1999. URL: <http://computer.org/conferen/proceed/meta/1999/papers/49/rnikolai.html>.
- Sun Microsystems (1999a): Java Foundation Classes: Now and the Future. URL: http://java.sun.com/marketing/collateral/foundation_classes.html
- Sun Microsystems (1999b): Java Plug-In Software – Overview. URL: <http://java.sun.com/products/plugin/1.1.2/index-1.1.2.html>.
- Swoboda, W., Kruse, F., Nyhuis, D. und Rousselle, H. (1998): Die Neukonzeption des Umweltdatenkatalogs, in: Haasis, H.-D. et al., 1998.
- Umweltbundesamt Österreich (1998): Umweltdatenkatalog Österreich im WWW. URL <http://udk.bmu.gv.at/>.
- Yahoo! (1998) Browsing Yahoo! URL: <http://howto.yahoo.com/chapters/8/1.html>, 1998.