

Die XML-Schnittstelle des UDK 4.2 - Der Schlüssel zur Integration von Umweltanwendungen

Fred Kruse¹, Marco Eichler², Ulrike Freitag³ und Thomas Sattler⁴

Abstract

The Umweltdatenkatalog (UDK = Environmental data catalogue) is a metainformation system for locating environmental information available in public authorities. The UDK is introduced in Austria and Germany. This article describes the Data Type Definition (DTD) for an XML-Interface.

1. Einleitung

Der Umweltdatenkatalog (UDK) ist ein System zum Nachweis von Umweltdaten des Bundes und der Länder in Deutschland und Österreich. Die Software des UDK besteht aus einer Windows-Anwendung zur Erfassung und Administration der UDK-Datenbestände und einer Recherche-Komponente die auf Java-Technologien beruht.

Das Datenmodell des UDK hat sich in den letzten Jahren zu einem de-facto Standard für Umweltdaten in Deutschland und Österreich entwickelt. Im Umweltbereich aber auch zunehmend im Bereich der Vermessung und der Bodenforschung wird der UDK als Metainformationssystem eingesetzt. Insbesondere durch den Einsatz als Metakomponente von Umweltinformationssystemen beim Bund und in den Ländern wird die Entwicklung einer Schnittstelle zu anderen Systemen unumgänglich.

Weitere Anforderungen für eine möglichst allgemeingültige Schnittstelle entstehen durch die Kopplung des UDK an andere Metainformationssysteme in gemeinsamen

¹ Niedersächsisches Umweltministerium, Koordinierungsstelle UDK, Archivstr. 2, D-30169 Hannover, Tel. +49 511 120 3436, udk@numis.niedersachsen.de, <http://www.umweltdatenkatalog.de>

² BVG mbH, Menkestr. 7, D-26419 Schortens, udk@bvg-fri.de, <http://www.bvg-fri.de>

³ Condat AG, Alt Moabit 91d, D-10559 Berlin, Tel. +49 30 390 94 222, uf@condat.de, <http://www.condat.de>

⁴ DECON-network, Griesbachstraße 10, D-76185 Karlsruhe, Tel. +49 721 830 11 34, udk@lfuka.lfu.bwl.de.

Rechercheanwendungen. So wird zur Zeit der Zugriff des Geographischen Metainformationssystems des InGeoForums entwickelt und ist ein Zugriff des virtuellen UDKs auf den bayrischen Umweltobjektkatalog (UOK) geplant.

Für alle Formen der Kommunikation mit anderen Systemen eignet sich besonders XML als Austauschformat. Zur Zeit ist eine Export- und Importschnittstelle in der Entwicklung. Diese Schnittstelle beruht auf einer bereits entwickelten Data Type Definition (DTD) für XML, die im Folgenden beschrieben werden soll.

2. Komponenten des UDK Datenmodels

Das Datenmodell des UDK unterscheidet Objekte und Adressen. Objekte und Adressen sind jeweils einem Katalog zugeordnet. Es ist möglich in einer Datenbank mehrere Kataloge zu führen.

Die Objekte werden untereinander hierarchisch, aber auch mit Querverweisen in Relation gesetzt. Da die Relationen prinzipiell bidirektional angelegt sind, werden diese getrennt von den Objekten beschrieben. Ähnliches gilt für die Relationen zwischen Adressen, die Adresshierarchien abbilden (z.B. zwischen Institution und der in ihr tätigen Person). Jedem Objekt können beliebig viele Adressen über Relationen zugeordnet werden. Eine Adresse, die auskunftsgibende Stelle, ist als Pflichtfeld im UDK vorgeschrieben.

Der UDK besitzt ein Klassenmodell, das unterschiedliche Fachbezüge für verschiedene Arten von Daten definiert. Eine besondere Eigenschaft sind dabei drei Klassen deren Fachbezüge frei definierbar sind. Für die Definition der drei freien Klassen steht ein eigenes Element in der DTD (`udk:data-model-extension`) zur Verfügung. Entsprechend enthält die Definition folgende Hauptelemente:

```
<!ELEMENT udk (udk:catalogue*, udk:data-source*, udk:address*,
udk:data-source-relation*, udk:data-source-address-relation*,
udk:address-relation*, udk:data-model-extension?)*>
```

3. Komponenten eines UDK-Objekts

Ein UDK-Objekt setzt sich aus Elementen zusammen, die die einzelnen Kategorien des UDK beschreiben, die sich im Wesentlichen auch in den einzelnen Karteikarten des Windows-UDK widerspiegeln. Für ein Objekt werden die Elemente Allgemein, Fachbezug, Raumbezug, Zeitbezug, Zusatzinformation und Verschlagwortung unterschieden.

```
<!ELEMENT udk:data-source (udk:general, udk:technical-domain?,
udk:additional-information?, udk:spatial-domain?, udk:temporal-domain?,
udk:subject-terms?, udk:available-linkage*)>
```

Das Element Allgemein enthält unter anderem die Pflichtfelder des UDK (Name des Objektes, Klasse des Objektes, Beschreibung des Objektes, Objekt-Identifikation, Identifikation und Datum der Ersterfassung, Identifikation und Datum der letzten Änderung). Einige dieser Felder dürfen beim Import aus anderen Systemen leer sein (siehe unten) und werden dann vom UDK beim Import erzeugt.

```
<!ELEMENT udk:general (udk:object-identifier, udk:catalogue-identifier,
udk:creator-identifier, udk:modifier-identifier, udk:object-class, udk:title,
udk:abstract?, udk:language-of-resource?, udk:language-of-record?,
udk:date-of-last-modification?, udk:date-of-creation?,
udk:original-control-identifier?, udk:no-of-parents?,
udk:general-additional-values?)>
```

Der Raumbezug unterscheidet insbesondere zwischen administrativen Einheiten, wobei sich diese zur Zeit auf die deutschen und österreichischen beschränken, und naturräumlichen Einheiten. Zusätzlich können insbesondere bei den naturräumlichen Einheiten die Geographischen Koordinaten einer umschließenden Boundig-Box übergeben werden.

```
<!ELEMENT udk:spatial-domain (udk:description-of-spatial-domain?,
udk:height?, udk:geo-location*)>
```

```
<!ELEMENT udk:geo-location (udk:bounding-coordinates?, (udk:place |
udk:region)?)>
```

Der Zeitbezug enthält natürlich die Felder von und zu, um einen Zeitraum zu beschreiben. Es können zusätzliche Informationen wie Periodizität und Intervall übergeben werden.

```
<!ELEMENT udk:temporal-domain (udk:description-of-temporal-domain?,
udk:beginning-date?, udk:ending-date?, udk:time-step?, udk:time-scale?,
udk:time-period?, udk:time-status?, udk:time-type?)>
```

Die Zusatzinformationen setzen sich in erster Linie aus den gesetzlichen Grundlagen und einer Reihe von Feldern zu Nutzungsanmerkungen (wie Datenformat oder Medium) zusammen.

```
<!ELEMENT udk:additional-information (udk:export-to*, udk:legislation*,
udk:internal-identification?, udk:legislation-note?, udk:access-constraints?,
udk:use-constraints?, udk:available-medium*, udk:available-data-format*,
udk:technical-documentation?, udk:publication-condition?)>
```

Für die Verschlagwortung wird der Umweltthesaurus und die Umweltklassifikation des Umweltbundesamtes verwendet. Es ist ebenso möglich, freie Suchbegriffe zu übergeben.

```
<!ELEMENT udk:subject-terms (udk:controlled-term* , udk:classification* ,
udk:uncontrolled-term*)>
```

Der Fachbezug eines Objektes wird je nach Klasse des jeweiligen Objektes gebildet. Der UDK sieht sechs Klassen vor: Datensammlung / Datenbank, Dokument / Bericht / Literatur, Geoinformation / Karte, Anwendung / Programm / Informationssystem, Vorhaben / Projekt / Programm und Organisationseinheit / Fachaufgabe. Die einzelnen Fachbezüge haben den beschriebenen Informationen entsprechend spezifische Felder (z.B. Geoinformation / Karte: Blattname, Blattnummer, Erstellungsmaßstab, ... oder Dokument / Bericht / Literatur: Autor, Herausgeber, ...). Zusätzlich können die Felder des Fachbezuges der frei definierbaren Klassen übergeben werden.

```
<!ELEMENT udk:technical-domain (udk:description-of-tech-domain? ,
(udk:dataset | udk:service | udk:document | udk:map | udk:project |
udk:additional-class-attribute-values)?)>
```

4. Import aus anderen Systemen

Die derzeit für den Windows-UDK entwickelte Schnittstelle ist in der Lage Objekte zu importieren. Weiterhin können Relationen zwischen Objekten und zwischen Objekten und Adressen importiert werden. Zur Zeit ist es nicht möglich, auch Adressen zu importieren, was aber für die Zukunft geplant ist. Die Relationen können sich, soweit sie bekannt sind, auf die UDK-Identifikation beziehen oder aber auf die Identifikation des Ursprungsystems. Letztere wird in dem Feld Original-Identifizierung gespeichert und bei späteren Importen zur Identifikation des Objektes genutzt. Die XML-Elemente zu einigen UDK-Pflichtfeldern können leer sein. Wird z.B. keine Objekt-Identifikation beim Import übergeben, wird diese vom Programm erzeugt. Sind die Pflichtfelder „erstellt von“ und „zuletzt geändert durch“ leer, wird hier der importierende Nutzer eingesetzt. Ist für ein Objekt die Pflichtrelation zur auskunftsgewährenden Adresse nicht erfüllt, wird diese per Dialog während des Importvorganges abgefragt. Für ein Objekt, das keine Relation zu einem Elternobjekt übergibt, wird diese ebenfalls dialoggesteuert abgefragt.

5. Zukünftige Entwicklungen

Weitere Entwicklungen werden sich insbesondere im Recherchebereich ergeben. Die bisher entwickelte DTD eignet sich zur Übergabe von Detailinformationen zu Objekten und Adressen. Was aussteht ist ein Profil, das Suchanfragen und die Rückgabe von Suchergebnissen definiert. Entsprechende Profile werden im Laufe dieses Jahres entwickelt werden.