

Texte

**07**  
**08**

ISSN  
1862-4804

## Umweltinformationssysteme - Anwendungsbeispiele für den Themenschwerpunkt Wasser

Umwelt  
Bundes  
Amt 

Für Mensch und Umwelt





## Umweltinformationssysteme - Anwendungsbeispiele für den Themenschwerpunkt Wasser

Workshop des Arbeitskreises  
„Umweltdatenbanken/Umweltinformationssysteme“  
der Fachgruppe „Informatik im Umweltschutz“,  
veranstaltet in Zusammenarbeit mit dem Umweltbundesamt  
am 21. und 22. Mai 2007 in Hamburg

Im Auftrag des Umweltbundesamtes

Diese Publikation ist ausschließlich als Download unter  
<http://www.umweltbundesamt.de>  
verfügbar.

**Herausgeber:** Umweltbundesamt  
Postfach 14 06  
06813 Dessau-Roßlau  
Tel.: 0340/2103-0  
Telefax: 0340/2103 2285  
Internet: <http://www.umweltbundesamt.de>

**Redaktion:** Fachgebiet IV 2.1  
Gerlinde Knetsch  
Uta Zacharias

Dessau-Roßlau, Februar 2008

## Vorwort

Angesichts neuer Herausforderungen in der Umsetzung von Europäischen Richtlinien ist der Einsatz von Informations- und Kommunikationstechnologien eine entscheidende Triebkraft für die Verbesserung der Bereitstellung von Umweltinformationen auf verschiedenen Ebenen. Dies kann unterstützt werden durch den Aufbau und die Nutzung einer Daten- und Informationsinfrastruktur zum Beispiel für Überwachungsaufgaben der Länder, für Berichtspflichten des Bundes und der Länder im Zusammenhang mit der Umsetzung der Richtlinien auf nationaler Ebene aber ebenso für die Entwicklung von Modellen zur prognostischen Darstellung der Entwicklung des Zustandes unserer Umwelt.

Aufbauend auf diesem Gedanken setzte sich der Arbeitskreis *"Umweltdatenbanken/Umweltinformationssysteme"* des Fachausschusses *"Informatik im Umweltschutz"* am 21. und 22. Mai 2007 in Hamburg den Themenschwerpunkt *Wasser*. Die vom Bundesamt für Seeschifffahrt und Hydrographie (BSH) organisierte Veranstaltung bezog den Blickwinkel der Umsetzung einer Europäischen Richtlinie auf nationaler Ebene ein. Mit der europäischen Wasserrahmenrichtlinie (WRRL) 2000/60/EG vom 07. September 2000 schuf das Europäische Parlament einen neuen Ordnungsrahmen für Maßnahmen der Mitgliedstaaten im Bereich der Wasserpolitik. Die explizite Aufnahme der Aspekte *Wechselwirkungen von Grundwasser, Oberflächenwasser, Auen und Küstengewässer* sowie der *Erstellung von Gewässerbewirtschaftungsplänen* in die WRRL fördert die gezielte Datenaufbereitung und Nutzung für konkrete Handlungsfelder in der Gewässerpolitik.

Der Workshop bot den Teilnehmenden Gelegenheit zur Präsentation und Diskussion von innovativen Anwendungsentwicklungen unter Einsatz von modernen Technologien, zur Vorstellung konzeptioneller Lösungsansätze für Modellierungen und Ausbreitungsrechnungen in marinen Ökosystemen aber auch zur Erörterung der Herausforderungen des Datenmanagements für das Langzeitmonitoring. Die offene und konstruktive Arbeitsatmosphäre mit Fachexperten aus Behörden, Instituten und Forschungseinrichtungen Deutschlands und Österreichs erlaubte an der einen oder anderen Stelle eine Vision des zukünftigen Datenmanagements zu diskutieren und den Blickwinkel auf gesundheitliche Aspekte der Umweltinformation auszuweiten.

Vier Vorträge waren eng mit dem Themenschwerpunkt „Wasser“ verbunden. Dies betraf einen interessanten konzeptionellen Ansatz für die Implementierung von Fachdatenbanken, Geo- und Metadateninformationssystemen in eine lokale Dateninfrastruktur des Nationalparks Schleswig-Holsteinisches Wattenmeer. Die Zielsetzung dieser Konzeption richtet sich vorrangig an den Anforderungen der WRRL aus, die ein umfangreiches Monitoring auf Basis von Kartierungen und Messreihen für Küstengewässern und Ästuaren vorschreibt, ein national koordiniertes Berichtswesen verlangt und die Information der Öffentlichkeit unterstützt. Ein im Hessischen Reed in Vorbereitung befindliches Projekt zum Grundwassermonitoring dokumentierte die automatisierte und qualitätsgesicherte Datenerfassung unter Einsatz von Datenloggern. Eine in das Projekt integrierte Kosten-Nutzen-Analyse beleuchtet den wirtschaftlichen Aspekt dieser Technologie. Die Vorgehensweise bei der Integration verschiedenster Daten von Betreibern von Abwasseranlagen in Hessen zu einem Jahresbericht verdeutlichte den hohen informationstechnischen Anspruch an die Umsetzung in eine nutzergerechte Technologie. Auch bei der Lösung zum Data Warehouse Wasser (DWW), Teil des vom Geschäftsbereich des Bayerischen Staatsministeriums für Umwelt, Gesundheit und Verbraucherschutz entwickelten und betriebenen Informationssystems Wasserwirtschaft (INFO-Was) überzeugte der hohe Grad der Integration von wasserwirtschaftlichen Daten, die für fachübergreifende Fragestellungen nutzerspezifisch aufbereitet werden. Innovative Lösungen des von Bund und Ländern gemeinsam betriebenen Umweltportal Deutschland PortalU<sup>®</sup>, der webbasierten Aufnahme von Daten aus dem bundesweiten Moos-Monitoring, zu Harmonisierungsbestrebungen von Bodendaten und des Einsatzes des Software-Pakets InGrid im Umweltinformationssystem Baden-Württemberg sowie eines Überblicks zu chemischen Stoffinformationssystemen rundeten das Themenspektrum des Workshops ab.



Gerlinde Knetsch

Umweltbundesamt  
Fachgebiet Informationssysteme Chemikaliensicherheit  
Februar 2008

# Inhaltsverzeichnis

## Vorwort

<b>Aufbau einer Dateninfrastruktur für Monitoring und Berichtswesen am Nationalparkamt</b>	<b>1</b>
J. Kohlus, B. Stahl, K. Eskildsen Landesamt für den Nationalpark Schleswig-Holsteinisches Wattenmeer	
<b>'MossMett' – Ein WebGIS auf Basis von Open Source - Software für das bundesweite Moos-Monitoring 2005</b>	<b>11</b>
L. Kleppin, C. Aden, R. Pesch, G. Schmidt, W. Schröder HS Vechta	
<b>Automatische Überführung und Harmonisierung von Bodendaten mit Hilfe digitaler Übersetzungsschlüssel</b>	<b>19</b>
F. Hosenfeld, D. Kuhn	
<b>Datenlogger-gestütztes Grundwassermonitoring</b>	<b>29</b>
Th. Gutzke, G. Seewald envi-systems GmbH	
<b>Data Warehouse Wasser – ein integratives Recherche- und Informationssystem innerhalb des bayerischen Informationssystems Wasserwirtschaft</b>	<b>37</b>
A. Reineke Bayerisches Landesamt für Umwelt C. Hofmann disy Informationssysteme GmbH	
<b>Dezentrale EKVO-Datenerfassung per Rich-Client-Anwendung - Praktische Erfahrungen mit Offline-Datenbanken</b>	<b>43</b>
W. Müller; HLUG H. Pfaff; RP-U Gießen J. van Nouhuys, T. Fischer, I. Mohr; Condat AG	

<b>A Modeling Methodology for Scientific Processes</b>	<b>55</b>
St. Jablonski, B. Volz, M. Abdul Rhemann Lehrstuhl für Angewandte Informatik IV Universität Bayreuth	
<b>Gefahr in Verzug – Stoffinformationssysteme für die Bewertung von Chemikalien</b>	<b>61</b>
G. Knetsch Umweltbundesamt Dessau	
<b>PortalU-Komponenten und –Schnittstellen zur Datenrecherche und -erfassung</b>	<b>69</b>
M. Klenke, F. Kruse, Th. Vögele Koordinierungsstelle PortalU im Niedersächsischen Umweltministerium	
<b>Ein Konzept zum Fachdokumenten-Management und zum Einsatz von InGrid im Umweltinformationssystem Baden-Württemberg (UIS BW)</b>	<b>85</b>
R. Weidemann, Th. Schlachter Forschungszentrum Karlsruhe, IAI	
R. Ebel Landesanstalt für Umwelt, Messungen und Naturschutz BW	

# Aufbau einer Dateninfrastruktur für Monitoring und Berichtswesen am Nationalparkamt

Jörn Kohlus, Britta Stahl, Kai Eskildsen

[joern.kohlus@nationalparkamt.de](mailto:joern.kohlus@nationalparkamt.de)

[britta.stahl@nationalparkamt.de](mailto:britta.stahl@nationalparkamt.de)

[kai.eskildsen@nationalparkamt.de](mailto:kai.eskildsen@nationalparkamt.de)

## Abstract / Einleitung

Das Landesamt für den Nationalpark Schleswig-Holsteinisches Wattenmeer ist heute eine der kleinsten Oberen Landesbehörden in der Bundesrepublik. Trotzdem gehen bis heute Impulse und Konzepte zur Datenverarbeitung von dort auf vielfach größere Institutionen und Organisationseinheiten über. „Not macht erfinderisch“ beschreibt treffend die Notwendigkeit einer kleinen Organisationseinheit ihre Datenhaltung auch mit Hilfe Dritter zu optimieren.

Der Beitrag beschreibt die Konzeption der lokalen Implementierung von Datenbanken, Geoinformationssystem und Metadaten für die lokale Dateninfrastruktur.

Besonders durch die Umsetzung der europäischen Umweltrichtlinien sind in den vergangenen sieben Jahren zusätzliche Aufgaben für die Datenhaltung des Monitoring im Wattenmeer hinzugekommen. Mit technischen Komponenten für die erweiterten Berichtspflichten soll auf sie reagiert werden.

## 1 Aufgaben und Instrumente

Die Aufgaben des Landesamtes für den Nationalpark Schleswig-Holsteinisches Wattenmeer, kurz Nationalparkamt, leiten sich aus dem 1999 novellierten Nationalparkgesetz her. Sie umfassen die Information der Öffentlichkeit, Bildungsarbeit, Besucherlenkung, die ökologische Umweltbeobachtungen (Monitoring) und die Erstellung wissenschaftlicher Planungsgrundlagen sowie die Betreuung und Verwaltung des Nationalparks. Zudem ist der Nationalpark im Netzwerk Natura 2000 als Schutzgebiet sowohl nach Vogelschutz- (EWG 1979) als auch nach Flora-Fauna-Habitat-Richtlinie (EWG 1992) ausgewiesen.

Grundlagenwissen für diese Aufgaben wurden vor allem im Rahmen der Ökosystemforschung Wattenmeer (Stock et al. 1996) von 1989 bis 1996 geschaffen. In deren Rahmen wurde seit 1990 auch das Geoinformationssystem (GIS) des Nationalparkamtes eingerichtet (Kohlus 1992).

Aus dem Vorhaben gingen auch erste Vorschläge für eine integrierte Umweltüberwachung des Wattenmeeres hervor. Zusammen mit den dänischen und niederländischen Partnern wurde daraus das trilaterale Monitoring- und Bewertungsprogramm (TMAP) für das gesamte Wattenmeer entwickelt, trilateral

beschlossen und implementiert (CWSS & TMAG 1997). In Tönning am Nationalparkamt wurde einer der vier regionalen Datenköpfe für das TMAP eingerichtet (CWSS & TMAG 2004).

Die europäischen Umweltrichtlinien – dabei besonders die Wasserrahmenrichtlinie (WRRL) (EG 2000) - stellen zum Teil sehr konkrete Anforderungen, denen nur mit Hilfe einer Dateninfrastruktur effektiv nachgekommen werden kann. Hierzu gehören ein umfangreiches Monitoring auf Basis von Kartierungen und Messreihen, ein national koordiniertes Berichtswesen und die Information der Öffentlichkeit. Aus den zum Nationalpark gehörenden Küstengewässern und Ästuaren sind entsprechende Daten aus einem für die Erfordernisse der WRRL angepassten Monitoring verfügbar zu machen und aufzubereiten.

Als Reaktion auf die Richtlinie über den Zugang der Öffentlichkeit zu Umweltinformationen (EG 2003) fordert das Umweltinformationsgesetz (UIG) in seiner 2005 überarbeiteten Fassung die aktive Information der Öffentlichkeit über Umweltdaten. Diese Fachinformationen sind daher prinzipiell zugänglich zu machen. Dem kann explizit mittels im Internet bereitgestellter Metadaten und Fachinformationen nachgekommen werden.

Neben Anpassungen des GIS und der Datenbank auf diese neuen Anforderungen wurde von Seiten des Nationalparkamtes mit der Beteiligung am NOKIS-Projekt (Lehfeldt et al. 2006) und aktuell mit der laufenden Einrichtung des Berichtswerkzeuges Cadenza reagiert.

## **1.1 Basisinfrastruktur**

Das Nationalparkamt ist seit 2006 in das Landesnetz Schleswig-Holstein eingebunden. Alle angeschlossenen Einrichtungen haben dem IKOTECH III-Standard zu entsprechen. Damit ist die Festlegung auf eine Windows-Umgebung und eine Palette von Softwareprodukten, für die auch Support angeboten wird, gegeben (FM 2007). Die Anbindung des Nationalparkamtes an das Landesnetz erfolgt über einen Übergabeserver. Hierdurch können die Arbeitsplätze den Arbeitsanforderungen entsprechend leicht abweichend von IKOTECH III ausgestattet und betrieben werden. Der zentrale Dienstleister übernimmt den Support für das Landesnetz Schleswig-Holstein. Allerdings sind Wartung und Pflege der Bürosysteme und Spezialsoftware sowie die Systemsicherung in Eigenleistung durchzuführen. Den Sicherheitsanforderungen des Landesnetzes ist dabei zu entsprechen. Folglich sind nur einige Ports für die Kommunikation geöffnet und ein aktiver Zugriff auf Rechner im LAN des Nationalparkamtes ist ausgeschlossen.

## **1.2 Aufgaben und Struktur der TMAP-Datenbank**

Das TMAP wurde von den Wattenmeer-Anrainerstaaten mit dem Ziel eingeführt, die Umweltuntersuchung im Wattenmeer zu koordinieren. Nur so kann der Zustand des Wattenmeeres in seiner Ganzheit adäquat bewertet werden. Regelmäßig wird koordiniert durch das gemeinsame Wattenmeersekretariat (CWSS) ein Bericht über den Zustand dieses sensiblen Lebensraums erstellt und es werden Handlungsempfehlungen für den Schutz formuliert. Diese Berichte werden von Fachleuten erstellt, die Zugang zu den Daten des Gesamtgebietes benötigen. Wie die Evaluierung bei der Erstellung des letzten Wattenmeerzustandsberichts (Quality Status Report – Essink et al. 2005) gezeigt hat, wird durch die abgestimmten

Erfassungsmethoden und die einheitliche Ablage der Daten (CWSS & TMAG 2000) in den Länder diese Auswertung erheblich erleichtert (TMAG 2006). Mit Passwort geschützt, wird im Internet der Datenzugriff und -austausch über eine einheitliche Benutzeroberfläche ermöglicht.

Die Daten für die schleswig-holsteinische TMAP-Datenbank werden von vielen verschiedenen Personen und Institutionen erfasst. Für eine effektive und einfache Aktualisierung der TMAP-DB ist im Auftrag des Nationalparkamtes in den letzten Jahren in vielen Gesprächen mit den Datenoriginatoren eine formalisierte Datenlieferung (i.d.R. Excel-Tabellen) erarbeitet worden. Importwerkzeuge wurden entwickelt, die bei der Datenübernahme die Eingabewerte nach logischen Kriterien auf Vollständigkeit und Plausibilität prüfen und entsprechend kenntlich machen. Erst nach dieser Qualitätssicherung werden die Daten zunächst in die Datenbank des Nationalparkamtes eingelesen.

Der Füllungsgrad der Datenbank für das TMAP konnte hierdurch in den letzten beiden Jahren erheblich erhöht werden. So enthält die schleswig-holsteinische TMAP-DB heute lange Datenreihen biologischer, chemischer und geographischer Parameter sowie allgemeine Daten zu Wetterbedingungen und Hydrographie auf einem jährlich aktualisierten Stand.

### **1.3 Nutzung und Einbindung des GIS**

Das GIS auf der Basis von ESRI-Produkten wird für die Aufgaben von Verwaltung, Planung und Monitoring genutzt. Räumliche Daten des Monitoring wie Robbenzählungen (Kohlus & Abt 2006), Seegraskartierungen oder detaillierte Vegetationskarten des Vorlandes (Stock et al. 2005) werden neben Geodaten über rechtliche Grenzen oder historische Zustände erfasst, aufbereitet und für die Nutzung vorgehalten. Seit 1990 wurden mehrere tausend Geodatenätze erstellt. Neben der grafischen Umsetzung in Karten sind Analysen und statistische Auswertungen zum Berichtswesen und Datenlieferungen zu externen Anfragen bei Eingriffsvorhaben im Nationalpark zu erstellen. Für Aufgaben der Analyse, Statistik aber auch der Datenhomogenisierung werden topologische Eigenschaften genutzt und benötigt.

Die Geo-Datenhaltung ist nicht nur aus Kostengründen Datei-basiert. Seit 2007 können neben eigenen Daten die Geobasisdaten der Landesvermessung einschließlich Rasterprodukten – gescannten und georeferenzierten Karten und Luftbilder – genutzt werden. Zusammen mit einem ab Frühjahr 2007 verfügbaren Raster-Höhenmodell einer Laservermessung sind insgesamt rund drei Terrabyte Geodaten zu verwalten und im Zugriff zu halten (Abb. 1).

Ein stetiger Zuwachs an Daten ist absehbar. Zwar werden die Daten vorbearbeitet von der Geodatenkopfstelle im Rahmen des Natur- und Umweltinformationssystems Schleswig-Holstein (NUIS-SH, Zölitz 1991) geliefert und am Nationalparkamt müssen diese Datenmengen nicht gesichert und dauerhaft archiviert werden. Anders als im terrestrischen Bereich ist jedoch ein Rückgriff auf Altdaten regelmäßig bei der Datennutzung und Interpretation aufgrund der extremen Morphodynamik des Wattgebietes vorzunehmen. Als räumliches Referenzsystem wird die Gauß-Krüger Projektion verwendet, für die Berichte im Rahmen von TMAP werden die Daten aber trilateral vereinheitlicht in die UTM-Projektion zur Basis des europäischen Ellipsoiden überführt. Eine Umstellung auf ETRS-89 (GRS80-Ellipsoid) ist geplant. Als Standard im TMAP ist die Weitergabe der Geodaten im offenen ESRI-shape Format – mit dem Verlust topologischer Information – vereinbart.



## 1.4 Metainformation

Metainformationen als Daten über Daten verbinden im heutigen Kontext den Gedanken, Dokumentation Informationen bereitzustellen, mit denen Daten nicht nur beschrieben, sondern auch auffindbar gemacht werden. Dies schließt Angaben zum technischen und rechtlichen Zugang ein. Als erstes relevantes Metadatenvorhaben wurde seit 1991 der Umweltdatenkatalog (UDK) entwickelt. Eine Reaktion auf den starken Zuwachs an verfügbaren, digitalen Daten. In übergeordneten Einrichtungen – besonders in den sich auf neu erkannte Herausforderungen einstellende Umweltministerien – entstand der Wunsch, mittels des UDK Überblicksinformationen zu den Daten der nachgeordneten Einrichtungen zu erhalten. Allerdings konnten für diese neue Aufgabe der Bereitstellung von Metadaten, neben der Fortführung der Dokumentation, keine adequate Mittel bereitgestellt werden.

Auf diese Problematik reagierte eine Initiative aus dem Bereich GIS und des Nationalparkamtes, die zielgebend für NOKIS wurde: die konzeptionelle Verbindung von Dokumentation und der Aufgabe Metadaten bereit zu stellen. Konzeption und Softwaretools von NOKIS lassen Mehrfachaufwand vermeiden und unterstützen eine effektive Erhebung der Metainformation. NOKIS hat sich von Beginn an auf den internationalen ISO 19115 Standard bezogen und Erweiterungen des Metadatenprofiles vereinbart, die auch die konsequente Beschreibung von Geodaten ohne publizierte Objektkataloge ermöglicht. Mit der Möglichkeit, lokale Profile zu erweitern, kann NOKIS als System für die lokale Dokumentation genutzt werden (Kohlus & Heidmann 2004). Die konsequent auf den Küstenbedarf ausgerichteten Parameter erlauben aber konzeptadequat auch größeren Einrichtungen NOKIS-Metadaten aus lokalen Metainformationssystemen zu generieren. Mit den NOKIS-Erweiterungen für Zeitreihendaten sollen auch die sekundär räumlichen Daten für das TMAP oder die WRRRL-Aufgaben mit dem System beschrieben werden. Erste Tests 2006 verliefen weitgehend erfolgreich.

Mit der unüblichen Förderung der Datenerfassung für NOKIS hat das BMBF es ermöglicht, einen Großteil der Datenerfassung nachträglich mit Metadaten zu beschreiben und eine Basis als Handlungsgrundlage geschaffen. Ohne die Verbindung von Dokumentation und Metadaten könnte das Nationalparkamt aber nicht adequat Metadaten zu den Datenbeständen liefern. Entsprechend der Pläne zur Verwaltungsreform in Schleswig-Holstein soll das Nationalparkamt mit der Küstenschutzverwaltung, dem Amt für ländliche Räume (ALR), zusammen gehen. Auch dort wird NOKIS als Metadateninformationssystem genutzt. NOKIS erlaubt ein kaskadiertes Konzept der Metadatenverwaltung. So werden weder bei der Verbindung mit einer anderen Einrichtung noch bei einer möglichen Neuordnung umfangreiche Zusatzarbeiten notwendig.

## 1.5 Unterstützung des Berichtswesens

Das Nationalparkamt hat sich 2006 entschieden, die Berichtssoftware Cadenza einzuführen. Erwartet wird, dass typische Anfragen an den Datenbestand vorgegeben werden und jeweils aktualisierte Ergebnisse entsprechend der Datenlage geliefert werden können. Von besonderem Interesse ist die Möglichkeit, die Abfragen im Intranet wie im Internet für den jeweiligen Nutzerkreis verfügbar zu machen. Ein zweiter Grund für das Interesse an der Einführung der Software liegt in der engen Verknüpfung von Datenabfragen mit einer räumlichen Darstellung.

Aus der Praxis des TMAP und sich regelmäßig wiederholenden Anfragen ist ein Teil des Bedarfes bekannt, der z. T. auch in stilisierter Form durch prioritäre Lösungen bedient wird. Der Hauptaufwand wird bei der Definition und Verabredung von Datenabfragen für die verbleibenden Parameter des Monitoring erwartet. Nach erfolgter Installation und erfolgreicher Implementierung der ersten Abfragen werden derzeit weitere Parameter in das System einbezogen.

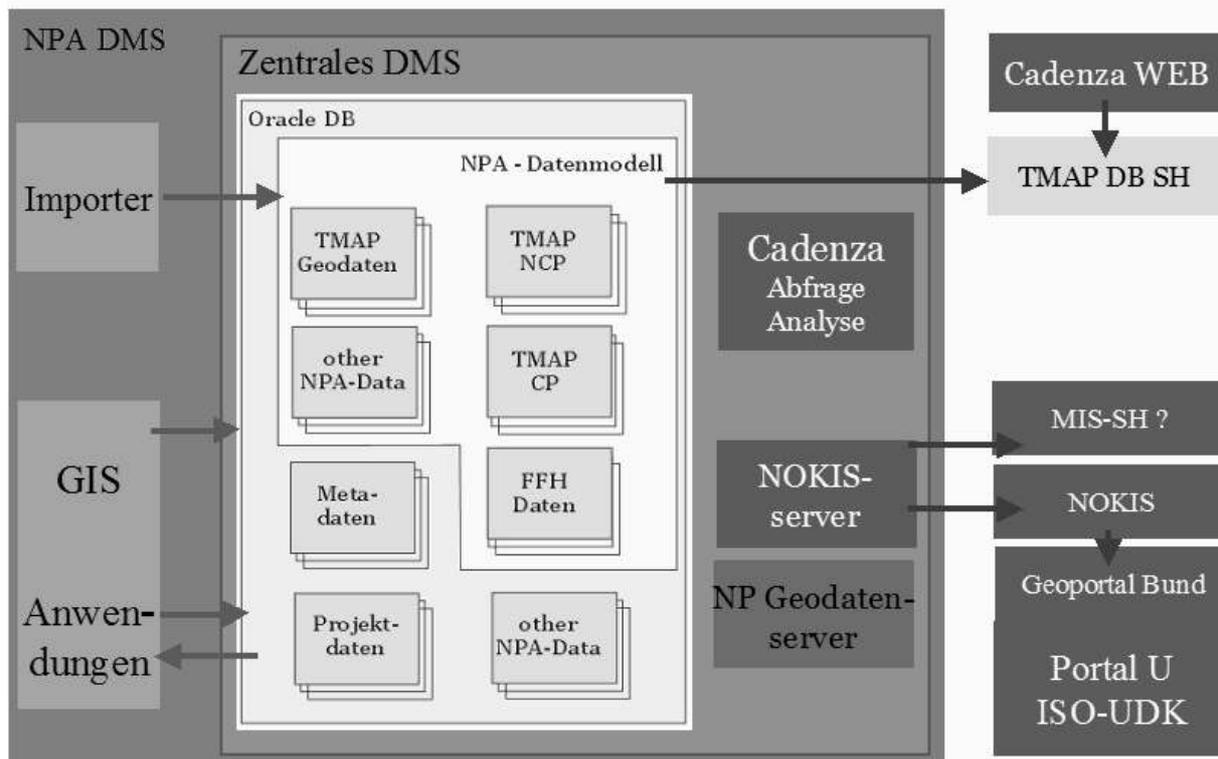


Abb. 2: Komponenten des Datenmanagementsystems des Nationalparkamtes sowie die Weitergabe von Metainformation und die geplante Bereitstellung von Daten des Monitorings aus dem TMAP.

## 2 Weitere Integration der Komponenten

### 2.1 Nationalpark DB und TMAP-Datenbank

Nicht nur für das TMAP liegen Daten im Nationalparkamt vor. Weitere Daten aus Einzeluntersuchungen, anderen Aufgabenbereichen des Amtes sowie Aufbereitungen historischer Datenbestände werden auf lokalen Arbeitsplatzrechnern vorgehalten. Ziel ist, die gesamten tabellarischen Daten des Nationalparkamtes in einer Datenbank zu halten (Abb. 2).

Aus dieser Nationalparkdatenbank heraus werden schon jetzt mit feststehenden Routinen Tabellen nach trilateraler Vorgabe generiert, die dann auf eine Datenbank jenseits der Firewall gespiegelt und extern verfügbar gemacht werden. Mit ähnlichen Verfahren wollen wir dem Bedarf aus den Anforderungen von WRRL und NATURA2000 (EWG 1979, 1992) nachkommen.

## **2.2 Bedarf für ein teilautomatisiertes Berichtswesen**

Mit dem wachsenden Umfang an Berichtsverpflichtungen muss eine Formalisierung der Berichtsform und der Berichtserzeugung einhergehen. Zu Teilen kann eine solche Aufgabe mittels Cadenza erfolgen, allerdings werden hier nur einzelne Grafiken und Diagramme produziert. Berichte, im Sinne der Richtlinien, sind jedoch komplexere Dokumente.

Sie bestehen aus Teilen mit nahezu statischer Information wie Erläuterungen zu Methoden, Zuständigkeiten, Gebietskulissen und Layout. Andere Teile dagegen sind dynamisch. Entsprechend des Standes der Daten verändern sich Karten oder Grafiken und damit zugehörige, die Daten und Darstellungen erläuternde Texte. Die Passagen werden üblicherweise von unterschiedlichen Bearbeitern erstellt, die häufig auch in verschiedenen Einrichtungen wirken. Letztlich werden als Ergebnisdokumente PDF-Dateien für den Druck und HTML-Seiten für eine Präsentation im Internet benötigt.

## **2.3 Verbindung von Metadaten und Datensichten**

Aus zweifacher Hinsicht wünschen wir uns eine Optimierung der Verbindung von Sichten auf Daten und der sie beschreibenden Metadaten. Bei der Suche und Auswahl von Metadaten ist zumindest eine voreingestellte Sicht auf den Dateninhalt sehr hilfreich. Bei NOKIS unterstützen Dienste (Lehfeldt & Reimers 2004) im Web – Webservices und Web-Anwendungen – nicht nur ein Datenvorschaubild, sondern erste einfache Manipulationen der Betrachtung zulassen und damit den Nutzer.

Umgekehrt besteht der Bedarf bei der Nutzung von Daten mittels Auswertungssoftware, deren Eignung, Verfügbarkeit und Ursprung nachvollziehen zu können. Ein zumindest lesender Zugriff auf die Metadaten aus einer Berichtssoftware heraus würde diesem Bedarf entgegenkommen.

Aus Sicht der Datennutzung wird eine serviceorientierte Architektur der Softwarekomponenten gewünscht, und zwar so, dass einzelne Komponenten – eine Diagrammdarstellung oder eine Darstellung der Metadaten – aus verschiedenen Anwendungen heraus verwendet werden können.

## **3 Literaturverzeichnis**

CWSS [Common Wadden Sea Secretariat]; TMAG [Trilateral Monitoring and Assessment Group] (Hrsg.): Implementation of the Trilateral Monitoring and Assessment Program (TMAP) - Final Report, Wilhelmshaven, 1997.

CWSS; TMAG (Hrsg.): CHAPTER III, TMAP Data Unit - User Manual, 2004. In: TMAP MANUAL - The Trilateral Monitoring and Assessment Program (TMAP), Wilhelmshaven, 2000 - 2007. [<http://www.waddensea-secretariat.org/TMAP/guidelines/Manual.html>]

CWSS; TMAG (Hrsg.): CHAPTER II, Guidelines, 2000; mit Annex 1: TMAP Hypotheses and Parameter Groups und Annex 2: Ecological Targets and Targets on Cultural and Historical Aspects. In: TMAP MANUAL - The Trilateral Monitoring and Assessment Program (TMAP), Wilhelmshaven, 2000 - 2007. [<http://www.waddensea-secretariat.org/TMAP/guidelines/Manual.html>]

- EG (Hrsg.): Richtlinie 2000/60/EG des Europäischen Parlaments und des Rates vom 23. Oktober 2000 zur Schaffung eines Ordnungsrahmens für Maßnahmen der Gemeinschaft im Bereich der Wasserpolitik. [http://europa.eu.int/smartapi/cgi/sga\\_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=DE&numdoc=32000L0060&model=guichett](http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=DE&numdoc=32000L0060&model=guichett), 2000.
- EG [Europäische Gemeinschaft] (Hrsg.): Richtlinie 2003/4/EG des europäischen Parlaments und des Rates vom 28. Januar 2003 über den Zugang der Öffentlichkeit zu Umweltinformationen und zur Aufhebung der Richtlinie 90/313/EWG des Rates. [http://eur-lex.europa.eu/LexUriServ/site/de/oj/2003/l\\_041/l\\_04120030214de00260032.pdf](http://eur-lex.europa.eu/LexUriServ/site/de/oj/2003/l_041/l_04120030214de00260032.pdf), 2003.
- Essink, Karel; Dettmann, Carsten; Farke, Hubert; Laursen, Karsten; Luerßen, Gerold; Marencic, Harald; Wiersinga, Wim (Eds.): Wadden Sea Quality Status Report 2004, Wadden Sea Ecosystem, No. 19, 2005.
- EWG [Europäische Wirtschaftsgemeinschaft] (Hrsg.): Richtlinie 79/409/EWG des Rates vom 2. April 1979 über die Erhaltung der wildlebenden Vogelarten. [http://www.bmu.de/gesetze\\_verordnungen/eg-vo\\_eg-richtlinien/doc/35475.php](http://www.bmu.de/gesetze_verordnungen/eg-vo_eg-richtlinien/doc/35475.php), 1979.
- EWG (Hrsg.): Richtlinie 92/43/EWG des Rates vom 21. Mai 1992 zur Erhaltung der natürlichen Lebensräume sowie der wildlebenden Tiere und Pflanzen. [http://europa.eu.int/eur-lex/de/consleg/pdf/1992/de\\_1992L0043\\_do\\_001.pdf](http://europa.eu.int/eur-lex/de/consleg/pdf/1992/de_1992L0043_do_001.pdf), 1992.
- FM [Finanzministerium des Landes Schleswig-Holstein] (Hrsg.): <http://www.ikotech.de/>. Kiel, 2007.
- Hofmann, Claus (2003): Wasserdichte Planung - Das übergreifende Fachinformationssystem in Niedersachsen. In: GeoBIT H. 5/2003, S. 16- 17.
- Kohlus, Jörn: Der GIS-Einsatz am Nationalparkamt für das Schleswig-Holsteinische Wattenmeer - ein Subzentrum der WATIS. In: Günther, Oliver; Schulz, Klaus-Peter; Seggelke Jürgen (Hrsg.): Umweltsanwendungen geographischer Informationssysteme.: S. 156-168; Karlsruhe, 1992.
- Kohlus, Jörn; Abt, Kai: Die Umsetzung von Beobachtungen der Robbenliegeplätze bei Befliegungen in Liegeplatz-Information. In: Traub, Karl-Peter; Kohlus, Jörn (Hrsg.): GIS im Küstenzonenmanagement. S. 227 – 236, Heidelberg, 2006.
- Kohlus, Jörn; Heidmann, Carsten: NOKIS Nord- und Ostsee Informationssystem. In: Schernewski, Gerald; Dolch, Tobias (Hrsg.): Geographie der Meere und Küsten. Coastline Report 1(2004), ISSN 0928-2734, S. 239-248.
- Lehfeldt, Rainer; Heidmann, Carsten; Reimers, Hans-Christian; Kohlus, Jörn; von Weber, Malte: NOKIS - Nord- und Ostsee KüstenInformationSystem - Netzwerk der Metadaten In: Traub, Karl-Peter; Kohlus, Jörn (Hrsg.): GIS im Küstenzonenmanagement. S. 150 – 160, Heidelberg, 2006.
- Lehfeldt, Rainer; Reimers, Hans -Christian: The WFD Reporting Process - A German Approach to Information Management in the Coastal Zone. In D. García de Jalón, P. Vizcaino (eds) 5th International Symposium on Ecohydraulics. Aquatic Habitats: Analysis & Restoration, September 12-17. IAHR, Madrid, Spain, pp175-181, 2004.

- Stock, Martin; Schrey, Eckart; Kellermann, Adolf; Gätje, Christiane; Eskildsen, Kai; Feige, Matthias; Fischer, Gundula; Hartmann, Frank; Knoke, Vera; Möller, Andrea; Ruth, Maarten; Thiessen, Anja, Vorberg, Ralf: Ökosystemforschung Wattenmeer – Synthesebericht Grundlagen für einen Nationalparkplan, Schriftenreihe des Nationalparks Schleswig-Holsteinisches Wattenmeer, Heft 8, 784 S., Heide, 1996.
- Stock, Martin; Gettner, Sabine; Hagge, Heinke; Heinzel, Kirsten; Kohlus, Jörn; Hartmut Stumpe: Salzwiesen an der Westküste von Schleswig-Holsteins 1988 – 2001. In: Schriftenreihe des Nationalparks Schleswig-Holsteinisches Wattenmeer, H. 15, Tönning, 2005.
- TMAG [Trilateral Monitoring and Assessment Group]: QSR data handling evaluation, unveröffentlichter Bericht zum TMAG-Treffen, TMAG-06-2-05.2-2, 6 S., 2006.
- Zölitz-Möller, Reinhard (1991): Vorstudie zum Aufbau eines Natur- und Umweltinformationssystems Schleswig-Holstein (NUIS-SH) - Hrsg. vom Minister für Natur, Umwelt und Landesentwicklung des Landes Schleswig-Holstein, Kiel. 96 S., 24 S. Anhang.



# 'MossMet' – Ein WebGIS auf Basis von Open Source-Software für das bundesweite Moos-Monitoring 2005

Kleppin, L., Aden, C., Pesch, R., Schmidt, G., Schröder, W.

[kleppin@iuw.uni-vechta.de](mailto:kleppin@iuw.uni-vechta.de)

[caden@iuw.uni-vechta.de](mailto:caden@iuw.uni-vechta.de)

[rpesch@iuw.uni-vechta.de](mailto:rpesch@iuw.uni-vechta.de)

[gschmidt@iuw.uni-vechta.de](mailto:gschmidt@iuw.uni-vechta.de)

[wschroeder@iuw.uni-vechta.de](mailto:wschroeder@iuw.uni-vechta.de)

## Einleitung

Seit 1990 wird in einem 5-Jahres Rhythmus unter Beteiligung von bis zu 32 Nationen das europäische Umweltbeobachtungsprogramm „Heavy Metals in Mosses Survey“ als Bestandteil des UNECE-Kooperativprogramms ICP Vegetation (Working Group on Effects) durchgeführt. Mit dem Moos-Monitoring sollen die zeitlichen und räumlichen Trends der Schwermetallanreicherung erfasst werden. In Deutschland wurden 1990, 1995, 2000 und 2005 an 592, 1026, 1028 bzw. 726 Standorten Moose entnommen (Schröder & Pesch 2005, Pesch et al. 2007b, Pesch & Schröder 2007). Zur Optimierung des Informationsaustausches des bundesweit durchgeführten UNECE Moos-Monitorings 2005 wurden Web-gestützte Eingabe- und Recherche-Applikationen auf Basis von Open Source-Software entwickelt (Pesch et al. 2007a). Deren Aufgabe ist es, die deutschlandweit im Moos-Monitoring erhobenen Daten zentral zu verwalten und für statistische Analysen und Berichtspflichten abfragbar zu machen. Das WebGIS ermöglicht die räumliche Darstellung und Analyse der Moos-Monitoring-Daten sowie weiterer Umweltinformationen.

## 1 Technik

Die eingesetzten Software-Komponenten bestehen ausnahmslos aus Open Source-Produkten, die inzwischen eine anerkannte Alternative zu kommerziellen Produkten darstellen. Sie arbeiten nach Standards, die durch das Open Geospatial Consortium (OGC) auf Grundlage von ISO-, CEN-Standards und weiteren Normierungen erarbeitet werden. Standards wie der „Web Map Service“ (WMS) oder der „Web Feature Service“ (WFS) bilden eine Grundlage für die Vernetzung und Verarbeitung von dezentral gehaltenen Geodaten zu Geodaten-Infrastrukturen (GDI) (PENG 2003, MITCHELL 2005). Sie werden inzwischen von den meisten Herstellern von GIS-Software in die Produkte implementiert. Dabei ist mindestens der Einsatz eines WMS notwendig, um an einer GDI teilhaben zu können.

Die für das WebGIS verwendete Software besteht aus dem Linux-Betriebssystem Debian, dem darauf installierten http-Server der Apache Foundation, einer PHP-Umgebung und dem Datenbank-Managementsystem PostgreSQL mit der „spatial ex-

tension“ PostGIS für die Vorhaltung von vektorbasierten Geodaten. Bei dem eingesetzten Mapserver handelt es sich um den UMN Mapserver der University of Minnesota, der als WMS/WFS eingerichtet wurde. Als Benutzeroberfläche für die Ansicht der Geodaten und die Einrichtung der Webservices wird die WebGIS-Clientsuite Mappender der Firma WhereGroup<sup>1</sup> verwendet. Als GIS-Backend wird zur Zeit noch die “spatial-Extension“ PostGIS für das DBMS PostgreSQL eingesetzt, da die bisher benötigten GIS-Funktionen über die enthaltene Funktions-Bibliothek in die Anwendung eingebunden werden konnten. Die Implementation von GRASS-GIS als zukünftiges GIS-Backend (NETELER & MITASOVA 2004) ist Bestandteil aktueller Arbeiten.

## 2 Systemarchitektur

Die Abhängigkeit der einzelnen Systemkomponenten ist in der Abbildung 1 dargestellt und wird im Folgenden beschrieben.

Die Kommunikation zwischen Client und Server erfolgt über HTTP-Get/Post-Variablen (DREESMANN 2004). Ein Client (Web-Browser) versendet damit Parameter an einen Server (request), die je nach Nutzeranforderung ausgeprägt sind und die Antwort (response) des Servers beeinflussen. Die Parameter können bspw. eine nach Nutzerbedürfniss veränderte Größe für die Einstellung einer Kartenausdehnung enthalten. Der Webserver (Apache http-Server) nimmt die Parameter entgegen und gibt diese, je nach Art der Anfrage, an die Metadaten-Applikation oder den Mappender weiter. Richtet sich die Anfrage an die Metadaten-Applikation, kann die PostgreSQL-Datenbank über PHP-Funktionen mit SQL-Anweisungen z.B. abgerufen oder der Datenbankinhalt modifiziert werden. Die Ergebnisse der Anfrage werden an den Webserver (Apache) zurückgegeben, in HTML interpretiert und per HTTP an den Client gesendet. Derartige Anfragen können sich z.B. auf das Ausfüllen eines Probenentnahmeformulars beziehen. Ist die Anfrage an den Mappender gerichtet, um z.B. Standorte der aktuellen Messkampagne 2005/06 kartografisch darzustellen, leitet dieser die Anfrage an den UMN-Mapserver weiter. Der UMN-Mapserver seinerseits kommuniziert über Map-Dateien mit der PostgreSQL-Datenbank, in der die zur Verarbeitung oder Abbildung notwendigen Geodaten abgelegt sind. Der UMN-Mapserver liefert die Ergebnisse in Form von georeferenzierten Rasterbildern (z.B. als GIF) an den Webserver (Apache), dort werden sie in HTML interpretiert und per HTTP an den Client gesendet.

---

<sup>1</sup> <http://www.ccgis.de>

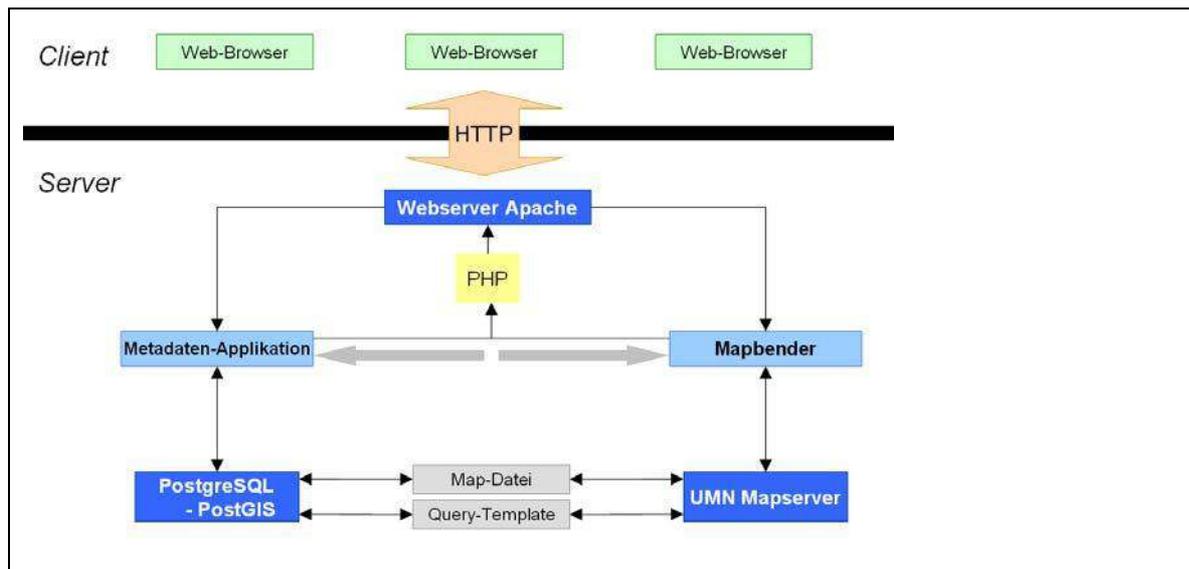


Abb. 1: Systemarchitektur und Kommunikation der 'MossMet'-Applikation

### 3 Die Web-Applikation 'MossMet'

Die passwortgeschützte 'MossMet'-Anwendung besteht aus einer Metadaten- und einer WebGIS-Applikation und dient der Web-basierten Eingabe und Ausgabe von Metadaten und Messwerten des bundesweiten Moos-Monitorings sowie der kartografischen Darstellung u.a. daraus mit geostatistischen Verfahren berechneter Flächenkarten der in den Moosen gemessenen Schwermetalle.

Bevor die digitale Erfassung der jeweiligen Probenentnahmeprotokolle erfolgen kann, erscheint eine Aufforderung zur Eingabe der Gauß-Krüger-Koordinaten des beprobten Standortes. Durch Abgleichen der Standortkoordinate mit allen in der Datenbank vorhandenen Koordinaten wird automatisch ermittelt, ob dieser Standort bereits bei einer früheren Kampagne beprobt wurde. Gemäß der UNECE-Probenentnahmerichtlinie (UNECE 2005) gilt eine Probenentnahmestelle als wiederbeprobt, wenn sie im 2-km-Umkreis zu einem Standort aus vergangenen Monitoring-Kampagnen liegt. Ist dies nicht der Fall, erhält dieser automatisch entsprechend der fortlaufenden Nummerierung eine neue ID. Liegt ein bereits existierender Standort im 2-km-Umkreis des neuen Standorts, nimmt der neue Standort automatisch die ID des alten Standortes an. Ist der Standortabgleich in der Datenbank abgeschlossen, wird der Nutzer zum digitalen Probenentnahmeformular weitergeleitet.

**Allgemeine Angaben:**

<b>Wetter*:</b> <input type="radio"/> regnerisch <input type="radio"/> sonnig <input type="radio"/> bewölkt <input type="radio"/> nebelig <input type="radio"/> Sonstiges	<b>Moosart*:</b> <input type="radio"/> <i>Pleurozium schreberi</i> <input type="radio"/> <i>Scleropodium purum</i> <input type="radio"/> <i>Hypnum cupressiforme</i>	<b>Bewuchs*:</b> <input type="radio"/> spärlich <input type="radio"/> polsterartig <input type="radio"/> rasenbildend	<b>Verbreitung*:</b> <input type="radio"/> selten <input type="radio"/> häufig	<b>Kalkpartikel*:</b> <input type="radio"/> keine <input type="radio"/> vereinzelt <input type="radio"/> zahlreich
--	---	--	--	---

<b>Rechtswert:</b> 3560274	<b>Höhe ü. NN* (m):</b> <input type="text"/>	<b>Nr. der TK 25:</b> <input type="text"/>	<b>Moosprobenvolumen*:</b> <input type="radio"/> < 1 l <input type="radio"/> 1 - 2 l <input type="radio"/> > 2 l	<b>Sprosslänge*:</b> <input type="radio"/> 5 - 10 cm <input type="radio"/> 10 - 15 cm <input type="radio"/> > 15 cm
<b>Hochwert:</b> 5734100	<b>Neigung (°):</b> <input type="text"/>	<b>Exposition*:</b> wählen ▾		

**Angaben zur Probenentnahmestelle**

**Lage der Probenentnahmestelle\*:**

Lichtung Nadelwald <input type="radio"/>	Lichtung Laubwald <input type="radio"/>	Lichtung Mischbestand <input type="radio"/>	Lichtung Schonung <input type="radio"/>	Heidefläche <input type="radio"/>	Grasland <input type="radio"/>	Sonstiges <input type="radio"/>
--	---	---	---	--------------------------------------	-----------------------------------	------------------------------------

50 m → Hauptwindrichtung

Abb. 2: Auszug aus dem digitalen Probenentnahmeformular des 'MossMet'

Das digitale Probenentnahmeformular enthält alle gemäß der Probenentnahmerichtlinie relevanten, Standort beschreibenden Kriterien, die aufgenommen werden sollten, um eine statistische Evaluierung der Messergebnisse zu ermöglichen (UNECE 2005). Der Probenentnehmer kann die entsprechenden Metadaten über mehrere Drop-Down-Menüs sowie Radiobuttons, die die Eingabemöglichkeiten für einige Standort beschreibende Merkmale einschränken, auswählen. Die Eingabemöglichkeit bei Freitextfeldern ist insofern eingeschränkt, als dass z.B. in den Eingabefeldern Neigung oder Höhe über NN nur Zahlen akzeptiert werden, die einen bestimmten Wert nicht überschreiten (z.B. Höhe über NN < 2964). Unvollständige oder nicht korrekt ausgefüllte Probenentnahmeformulare werden im Zuge einer automatischen Plausibilitätsprüfung nicht angenommen, sondern dem Bearbeiter zur Korrektur noch einmal vorgelegt. Als Korrekturhilfe erscheint dabei am Anfang des Formulars eine Fehlerbeschreibung. Ergänzend zu diesen Standort beschreibenden Metadaten verwaltet das 'MossMet' auch die Messdaten über die Metallakkumulation in den Moosen. Mess- und Metadaten können aus der Datenbank nach Vorgabe individueller Kriterien abgerufen und in ein Excel-fähiges Format exportiert werden. Auf Bundesebene wurde in der aktuellen Monitoring-Kampagne 2005 ausschließlich die 'MossMet'-Anwendung zur Erfassung Standort beschreibender Angaben sowie zur Verwaltung von Messdaten eingesetzt.

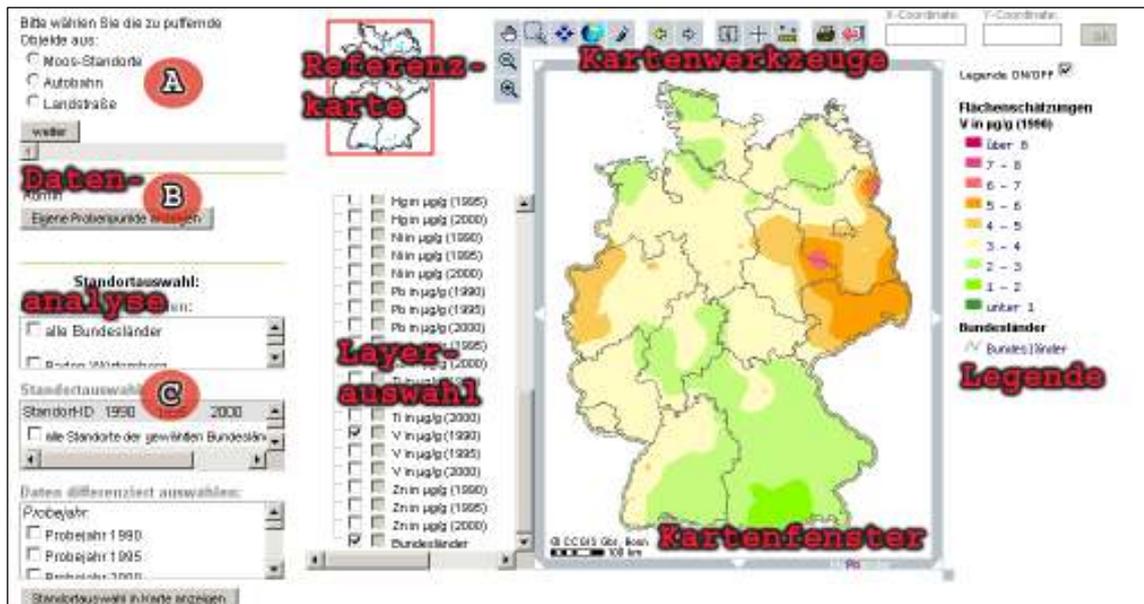


Abb. 3: WebGIS-Applikation von 'MossMet' (A=Layer, B=Kennung, C=Standortwahl)

Die kartografische Abbildung der räumlichen Informationen erfolgt durch die WebGIS-Applikation, dessen Nutzeroberfläche mit der Open Source-Software „Mapbender“ entwickelt wurde (Abb. 3). Das **Kartenfenster** dient als Ausgabefenster für alle aktivierten Layer. Das zentrale Element bildet hierbei eine Karte, eine Maßstabsleiste sowie Navigationspfeile, mit denen sich die Karte in alle Richtungen verschieben lässt. Die **Referenzkarte** ist ein verkleinertes Abbild der Gesamtkarte und dient der Orientierung. Die **Werkzeugleiste** ermöglicht grundlegende GIS-Operationen, wie z.B. das interaktive Festlegen eines Kartenausschnitts, die Abfrage ausgewählter Layer oder die Durchführung von Distanzmessungen. Der Aufbau der **Layerauswahl** ähnelt einer Windows-Explorer-Leiste und stellt Daten zum Moos-Monitoring sowie zusätzliche relevante Geodaten in Form von administrativen Grenzen, Klima-Messstationen, Verkehrswegen und Informationen zu landschaftsökologischen Raumeinheiten sowie zur Landbedeckung bereit, die in der PostgreSQL-Datenbank vorgehalten werden. Mit der linken Checkbox lässt sich der Layer im Kartenfenster ein- bzw. ausblenden. Entsprechend den aktivierten Karten wird automatisch eine **Legende** generiert. Die Aktivierung der zweiten Checkbox ermöglicht eine interaktive Abfrage des jeweiligen Layers im Kartenfenster mit dem Info-Button. Im Teilbereich **Datenanalyse A** (Abb. 3) ist die GIS-Funktion Puffer integriert, welche auf Probenentnahmestellen, Bundesstraßen und Autobahnen angewendet werden kann. Die Ergebnisse lassen sich anschließend nach Auswahl der entsprechenden Layer im Kartenfenster anzeigen. Die Anfertigung einer Anfahrtsskizze wird so ermöglicht und dadurch das Auffinden und die anschließende Wiederbeprobung eines Standortes erleichtert. Der Teilbereich **B** der **Datenanalyse** ermöglicht die kartografische Darstellung der vom jeweiligen Bearbeiter bzw. Probenentnehmer dokumentierten Beprobungsstandorte. Der Teilbereich **C** dient der Suche nach Beprobungsstandorten anhand der Metadaten aus dem Probenentnahmeprotokoll.

Bisher wurden im 'MossMet' die Vorteile der OGC-Standards nur in geringem Umfang genutzt (Datenhaltung, GIS-Funktionen). Diese bilden jedoch eine Grundlage für die europaweite Erweiterung und Anwendung von MossMet. Neben den Daten aus dem deutschen Moos-Monitoring sollen in Zukunft auch Geo- und

Metadaten von Messstandorten weiterer Teilnehmerländer des europäischen Moos-Monitoring einbezogen werden.

Die Beschreibung einzelner Geodatenätze durch Metadaten erfolgt über eine eigens für die WebGIS-Applikationen des Lehrstuhls für Landschaftsökologie entwickelte Eingabemaske. Hiermit können z.B. Angaben zur Qualität, zu dem Bearbeiter oder Schlagworte erfasst werden. Dies erfolgt in Anlehnung an die Standards der ISO 19115 sowie des Umweltdatenkataloges (UDK). Die Recherche nach Geodaten wird damit ermöglicht.

## 4 Diskussion

Webbasierte Informationssysteme werden auch außerhalb Europas in unterschiedlichem Kontext eingesetzt. AFIS (Advanced Fire Information System) ist ein Feuer-Informationssystem, das in Südafrika Anwendung findet. Ursprünglich als lokale Applikation genutzt, wurde für das AFIS eine aus Open Source-Produkten bestehende Client-Server-Architektur aufgebaut, in die das AFIS integriert wurde. Das Ziel der Applikation ist es, Waldbrand gefährdete Gebiete mit einer möglichst hohen Wahrscheinlichkeit und hohen räumlichen Genauigkeit ausfindig zu machen. Hierzu werden z.B. aktuelle Luftbilder und Temperaturkarten mit archivierten Brandflächenkarten verschnitten (McFERRIN et al. 2007).

Eine Alternative zur Mapbender-Software ist das Open Source-Programm Chameleon (DM Solutions Group<sup>2</sup>), das ebenfalls für die Entwicklung von Web-Mapping Anwendungen dient. Mit Blick auf die integrierten Standard GIS-Funktionen und der Layerverwaltung würde diese GUI für die 'MossMet'-Anwendung ebenfalls in Frage kommen. Allerdings ist die Nutzerfreundlichkeit des Mapbender-Programms etwas größer. Im Gegensatz zum Mapbender benötigt Chameleon zur Aktualisierung des Kartenfensters einen Update-Button.<sup>3</sup> Das liegt daran, dass Mapbender mit Hilfe von Javascript und Chameleon mit Hilfe von PHP und Mapscript programmiert wurde.

## 5 Ausblick

Durch die Verwendung lizenzfreier Open Source-Software kann die Web-Anwendung 'MossMet' zeitlich unbegrenzt genutzt und weiterentwickelt werden, ohne dass zusätzliche Kosten etwa für Software-Lizenzen entstehen. Das 'MossMet' kann so für weitere Moos-Monitoring-Kampagnen in seiner Funktionalität ausgeweitet werden. Zur Auswertung der Daten des Moos-Monitorings wäre es z.B. wünschenswert, Verschneidungs-Operationen (intersect) mit unterschiedlichen Layern durchzuführen. Darüber hinaus soll das 'MossMet' um Upload- und Download-Funktionen erweitert werden, damit der Nutzer zusätzliche Geodaten in das WebGIS einbinden (upload) oder eigens berechnete Karten z.B. im ESRI shape-Format lokal abspeichern kann (download). Mit Blick auf weitere Moos-Monitoring-Kampagnen sollten bereits bei der Standortbeprobung Metadaten unter Verwendung von mobilen Computern wie z.B. Personal Digital Assistants (PDA) digitalisiert und drahtlos in die zentrale Moos-

---

<sup>2</sup> <http://chameleon.maptools.org/>

<sup>3</sup> <http://www.mapsherpa.com/hawaii2/>

Datenbank übertragen werden. Das 'MossMet' wurde für das deutsche Moos-Monitoring-Programm entwickelt. Sein Ausbau für das europaweite Moos-Monitoring wurde auf dem 20. UNECE ICP Vegetation Treffen in Dubna (Russland, März 2007) empfohlen. Hierbei hat sich eine Kooperation zwischen einigen europäischen Ländern und Deutschland entwickelt, deren Ziel die Erarbeitung einer europäischen 'MossMet'-Anwendung ist. Hierbei sollen die Standards des OGC in größerem Umfang zum Tragen kommen.

## 6 Literaturverzeichnis

- DREESMANN, M. (2004): Geoservices nach dem Open Geospatial Consortium. Dokumentation auf dem GIB-Portal, <http://gdi.berlin-brandenburg.de>.
- MCFERREN, G., ROOS, S., TERHORST, A. (2007): Fire Alerts on the Geospatial Semantic Web. In: Scharl A, Tochtermann K (eds), *The Geospatial Web. How geo-browsers, social software and the Web 2.0 are shaping the network society*. Springer, London.
- MITCHELL, T. (2005): *Web Mapping illustrated*. O'Reilly, Sebastopol, Cambridge.
- NETELER, M., MITASOVA, H. (2004): *Open Source GIS: A GRASS GIS Approach*. 2nd edition. Kluwer Academic Publishers, Boston.
- PENG, Z.-R., TSOU, M.-H. (2003): *Internet GIS, distributed geographic information services for the internet and wireless networks*. John Wiley and Sons, Hoboken NJ.
- PESCH, R., SCHMIDT, G., SCHRÖDER, W., ADEN, C., KLEPPIN, L., HOLY, M., (2007a): Development, Implementation and Application of the WebGIS 'MossMet'. In: Scharl, A., Tochtermann, K. (eds), *The Geospatial Web. How geo-browsers, social software and the Web 2.0 are shaping the network society*. Springer. London, S. 191 - 200
- PESCH, R., SCHRÖDER, W., DIEFFENBACH-FRIES, H., GENßLER, L. (2007b): Optimierung des Moosmonitoring-Messnetzes in Deutschland. UWSF – Z Umweltchem Ökotox (Online First 14.03.2007) 1–12 <DOI: <http://dx.doi.org/10.1065/uwsf2007.03.166>>
- PESCH, R., SCHRÖDER, W. (2007): Entwicklung der Metall-Bioakkumulation in Deutschland zwischen 1990 und 2000 auf zwei naturräumlichen Differenzierungsstufen. In: Strobl, J., Blaschke, Th., Griesebner, G. (eds.), *Angewandte Geoinformatik 2007. Beiträge zum 19. Agit-Symposium*, S. 527 - 536
- SCHRÖDER, W., PESCH, R. (2005): Correlation and time series of metals in mosses with selected sampling site specific and ecoregional characteristics in Germany. *Environ Sci Pollut Res* 12 (3), S. 159 - 167.
- UNECE (United Nations Economic Commission for Europe Convention on Long Range Transboundary Air Pollution) (2005). *Monitoring of Atmospheric Heavy Metal Deposition in Europe Using Bryophytes. Experimental Protocol 2005 / 2006 Survey*. Bangor, UK: UNECE ICP Vegetation.



# Automatische Überführung und Harmonisierung von Bodendaten mit Hilfe digitaler Übersetzungsschlüssel

Friedhelm Hosenfeld<sup>4</sup> und Dieter Kuhn<sup>5</sup>

[hosenfeld@digsyland.de](mailto:hosenfeld@digsyland.de)

[kuhn@agenda-regio.de](mailto:kuhn@agenda-regio.de)

## Abstract

Die Bundesanstalt für Geowissenschaften und Rohstoffe (BGR)<sup>6</sup> initiierte die Entwicklung eines digitalen Übersetzungsschlüssels zwischen unterschiedlichen Bodenkartieranleitungen. Seit 2005 wurde der zuerst geschaffene Übersetzungsschlüssel 1.0 (ÜS 1.0) in mehreren Zweigen erweitert, die die Transformation zwischen verschiedenen Ausgaben der deutschen Bodenkundlichen Kartieranleitung (KA) und der internationalen Klassifikation World Reference Base for Soil Resources (WRB) und auch unterschiedliche Software-Plattformen unterstützen: Ergänzend zu der Microsoft-Access-basierten Standalone-Lösung wurde eine Internet-Applikation entwickelt, um allen Interessierten den Zugang zu dem Übersetzungsschlüssel anzubieten, auch ohne eine spezielle Software zu installieren. Für die Nutzung des Internet-Übersetzungsschlüssels werden nur ein Standard-Web-Browser und eine Internet-Verbindung benötigt.

Die Software ist modular aufgebaut: Die Übersetzungsregeln für jeden Parameter werden in speziellen Übersetzungstabellen innerhalb der Datenbank verwaltet. Dieses Konzept ermöglicht die Anpassung der Übersetzungsregeln durch die Bodenexperten ohne Änderungen des Programmcodes.

## 1 Einführung

In Deutschland stellt die Bodenkundliche Kartieranleitung (KA) eine wesentliche Basis zur Beschreibung von Böden dar. Zusätzlich zu Anweisungen zur Durchführung von Bodenkartierungen definiert die KA die wesentlichen Eigenschaften und Klassifikation von Böden wie Bodentypen, Bodenart, Ausgangsgestein und ähnliches. Die Bodenkundliche Kartieranleitung erschien 1965 in der ersten Auflage; aktuell gültig ist die 5. Auflage (KA5: Ad-hoc-AG-Boden 2005). Neben dieser nationalen Nomenklatur existieren einige internationale Standards, von denen die Klassifikation nach FAO/WRB (FAO 2006, FAO, ISRIC, IUSS 2006) von besonderer Bedeutung ist, da sie einen weltweiten Vergleichsmaßstab von Böden liefern soll.

---

<sup>4</sup> DigSyLand - Institut für Digitale Systemanalyse & Landschaftsdiagnose, Zum Dorfteich 6, D-24975 Husby, EMail: [hosenfeld@digsyland.de](mailto:hosenfeld@digsyland.de), Internet: <http://www.digsyland.de/>

<sup>5</sup> AgendaRegio - Institut für regionale Analyse und Entwicklung, Marienthaler Straße 17, D-24340 Eckernförde, EMail: [kuhn@agenda-regio.de](mailto:kuhn@agenda-regio.de), Internet: <http://www.agenda-regio.de/>

<sup>6</sup> Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), Stilleweg 2, D-30655 Hannover, EMail: [poststelle@bgr.de](mailto:poststelle@bgr.de), Internet: <http://www.bgr.bund.de/>

Da einerseits Bodendaten bereits seit langer Zeit mit unterschiedlichen Methoden und Standards erhoben werden und andererseits die Auswertung und Nutzung dieser Daten standardisierte und homogene Informationen erfordern, war der Bedarf für eine automatisierte Übersetzung oder zumindest für eine automatisierte Unterstützung des Übersetzungsprozesses zwischen verschiedenen Kartierstandards in der Bodenkunde offensichtlich. Diese Anforderung gewinnt zunehmend an Relevanz, wenn die Interoperabilität von Bodendaten nicht nur auf der nationalen sondern auch auf internationaler Ebene betrachtet wird.

Die Nachfrage nach standardisierten Bodendaten bleibt aufgrund der vielfältigen Einsatzzwecke weiterhin steigend, so können Bodendaten zur Beurteilung von Erosion, Wasserhaltevermögen, Pufferkapazität, Schwermetallbindungsvermögen, CO<sub>2</sub>-Bindung, sowie für Aussagen zur Eignung von Böden als Kultur- und Archivgut, Produktionsstandort oder Bauland herangezogen werden.

Zu dem Zeitpunkt, zu dem die Bundesanstalt für Geowissenschaften und Rohstoffe (BGR) die Entwicklung digitaler Übersetzungsschlüssel veranlasste, existierten keine allgemeinen Werkzeuge die eingesetzt werden konnten, um Bodendaten, die gemäß der 3. Auflage der Bodenkundlichen Kartieranleitung (KA3: AG Bodenkunde 1982) erhoben worden waren, in Formate gemäß der 4. Auflage der Kartieranleitung (KA4: AG Bodenkunde 1995) oder in Standards gemäß FAO oder WRB (FAO 2006, FAO, ISRIC, IUSS 2006) zu überführen. Unterstützt durch die Ad-hoc Arbeitsgruppe Boden der Staatlichen Geologischen Dienste (SGD) der Länder und der BGR wurde von der BGR zunächst ein Forschungsauftrag mit dem Titel „Vergleich von Daten der KA3 – KA4 – FAO im Hinblick auf deren Nutzung für die Erstellung von nationalen und internationalen Bodenkarten“ vergeben (Kneib, Bongard & Nachbar 1998). Darauf aufbauend wurde die Entwicklung des ersten digitalen Übersetzungsschlüssels ÜS 1.0 (Adler et al. 2004) begonnen, der die Übersetzung der Klassifikation gemäß KA3 in die KA4 durchführen sollte. In der Konzeption des Übersetzungsschlüssels wurde gefordert, dass der Übersetzungsschlüssel einerseits dem komplexen Übersetzungsprozess gerecht wird; andererseits sollte der Übersetzungsschlüssel aber flexibel erweiterbar und gut handhabbar bleiben.

Im Anschluss an die Freigabe der ersten Version der Software beauftragte die BGR die Institute AgendaRegio and DigSyLand, unterstützt durch eine WRB-Expertengruppe und die BGR selbst, mit der Evaluation und Erweiterung des Übersetzungsschlüssels, um sowohl die Bedienbarkeit als auch die Qualität der Ergebnisse zu optimieren.

## **2 Ziele und Produkte**

Drei Ziele sollten durch die Überarbeitung des Übersetzungsschlüssels ÜS 1.0 erreicht werden:

Verbesserung der Übersetzungsqualität,

Steigerung des Bedienungskomforts und des Datenmanagements und

die Übersetzung weiterer Standards und Klassifikationen.

Um diese Ziele umzusetzen, wurden verschiedene Software-Produkte entwickelt, die auf ähnlichen Übersetzungstechniken und Programmcodes basieren, aber jeweils am besten für einen speziellen Anwendungszweck geeignet sind:

- **ÜS 1.2 (KA3/KA4)**  
Der Übersetzungsschlüssel ÜS 1.2 übersetzt Bodendaten, die gemäß KA3 aufgenommen wurden, in das KA4-Format. Dieses Produkt entspricht in den Grundfunktionen dem ersten Übersetzungsschlüssel ÜS 1.0, bietet aber optimierte Benutzungsschnittstellen mit besserer Bedienbarkeit und zusätzlichen Funktionen. Zudem wurden die Übersetzungsregeln überarbeitet, so dass deutlich bessere Übersetzungsergebnisse erzielt werden. ÜS 1.2 ist ausschließlich als Standalone-Anwendung auf der Basis von Microsoft Access erhältlich.
- **ÜS 2.0 KA4/KA5 Access**  
Diese Version bietet die gleichen Benutzungs- und Programmschnittstellen wie ÜS 1.2, enthält aber die Übersetzungsregeln zur Erzeugung von Bodendaten gemäß KA5 aus den Eingabedaten gemäß KA4.
- **ÜS 2.0 KA5/WRB Access**  
Die Microsoft Access-basierte Version ÜS 2.0 KA5/WRB wird eingesetzt, um Daten, die im Format der deutschen KA5 (Ad-hoc-AG-Boden 2005) vorliegen, in ein Format gemäß der internationalen FAO- und WRB-Klassifikationen (FAO 2006, FAO, ISRIC, IUSS 2006) zu transformieren. Die dafür benötigten Übersetzungsregeln befinden sich derzeit noch in der Entwicklung.
- **ÜS 2.0 Internet**  
Neben den Access-basierten Versionen des ÜS 2.0 wurde eine Internet-Anwendung geschaffen, die insbesondere die Nutzenden adressiert, die keine spezielle Übersetzungsschlüssel-Software installieren möchten. Bisher ist ausschließlich der KA4→KA5-Übersetzungsprozess im Internet verfügbar. Die KA5→WRB-Übersetzung wird jedoch folgen.

Die Entscheidung, verschiedene Produkte für unterschiedliche Übersetzungsaufgaben zu entwickeln, basierte auf mehreren Gründen: Einerseits benötigen viele Anwendende aufgrund ihrer Datensituation hauptsächlich eine Übersetzungsfunktion. In diesem Fall ist es empfehlenswert, sie mit dem jeweils speziell für den gewünschten Anwendungszweck geeigneten Programm zu versorgen, das in Gestaltung und Funktionalität ganz auf diesen Zweck ausgelegt ist. Derartige Software ist zudem unkomplizierter zu warten und konfigurieren. Andererseits befinden sich alle angebotenen Übersetzungsfunktionen noch in der Entwicklung, weitere Optimierungen sind geplant. Der Entwicklungs- und Evaluationsprozess lässt sich besser handhaben, wenn in der Software nur jeweils ein Übersetzungsprozess betrachtet werden muss.

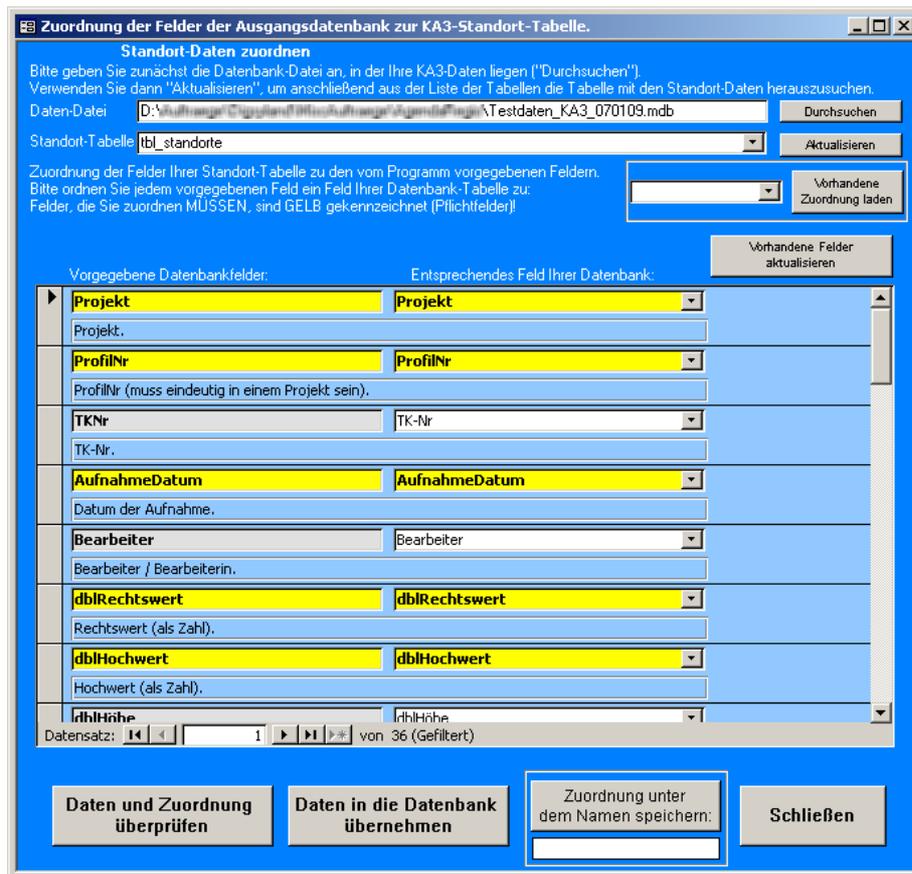


Abbildung 1: Übersetzungsschlüssel ÜS 1.2

Zuordnung der Eingangsdaten zu den Datenbank-Attributen (Pflichtfelder sind gelb gekennzeichnet).

### 3 Standalone-Übersetzungsschlüssel auf der Basis von Microsoft Access

Die Standalone-Versionen des digitalen Übersetzungsschlüssels basieren auf dem Datenbank-Produkt Microsoft Access (ab Version 2000). Die Anwendung verteilt sich auf eine Frontend-Datenbank, die die Benutzungsschnittstellen und die Programmlogik umsetzt, und eine Backend-Datenbank, die die Tabellen mit den Übersetzungsregeln bereitstellt.

Die meisten Übersetzungsregeln werden mit Hilfe von Übersetzungstabellen implementiert, die bestimmen, welche Eingabeparameterwerte auf welche Zielwerte abgebildet werden. Dieses Konzept der Übersetzungstabellen hat sich als sehr geeignet erwiesen, da die Übersetzungsregeln auf diese Weise direkt von den Bodenexperten verwaltet und optimiert werden können.

In einigen Ausnahmefällen ist der Einsatz komplexerer Übersetzungsmechanismen erforderlich, die durch spezielle Programmfunktionen umgesetzt werden.

Anpassungen dieser speziellen Übersetzungsregeln können Überarbeitungen des Programmcodes erforderlich machen.

Der gesamte Übersetzungsprozess wird durch eine Masterdatenbank-Tabelle gesteuert, die alle für den Übersetzungsprozess erforderlichen Angaben zu jedem Parameter verwaltet, wie etwa Art der Übersetzung und die zu verwendende Übersetzungs-Regeltabelle.

### 3.1 Übersetzungsbeispiel

Zur Verdeutlichung des Übersetzungsprinzips wird ein einzelner Übersetzungsschritt anhand eines Beispiels vereinfacht skizziert. Exemplarisch soll der KA3-Wert „Hnc,Hnp“ für „Torfart“ in die Nomenklatur der KA4 übersetzt werden:

- Zunächst wird die Master-Übersetzungstabelle abgefragt, um folgende Übersetzungsangaben für den Parameter „Torfart“ zu bestimmen:
  - *Übersetzungsart,*
  - *Übersetzungs-Regeltabelle* und
  - *Übersetzungs-Vergleichsbedingung.*
- In diesem Fall wird die Übersetzungsart „Regeltabelle mit Einzelwert-Zerlegung“ ermittelt, so dass der Originalwert in zwei Einzelwerte „Hnc“ und „Hnp“ zerlegt wird, die einzeln nach der geeigneten Methode übersetzt werden.
- Beide Einzelwerte werden in der Übersetzungs-Regeltabelle „tbl\_43d“ nachgeschlagen und gemäß der Vergleichsbedingung in den entsprechenden KA4-Wert übersetzt.
- Dies führt zu den Werten „Hnr“ und „Hnp“,
- die dann gemäß der Übersetzungsart zu dem Endergebnis „Hnr, Hnp“ zusammengesetzt werden.

In gleicher Weise werden alle Werte der Eingangsparameter übersetzt. Zur Zeit werden vier verschiedene Übersetzungsarten unterstützt:

- 1:1-Übernahme des Originalwertes,
- Verwendung einer Regeltabelle,
- Verwendung einer Regeltabelle mit vorheriger Einzelwert-Zerlegung,
- sowie die Übersetzungsart „separat“, bei der spezielle Übersetzungsfunktionen zum Einsatz kommen.

### 3.2 Übersetzungsschlüssel ÜS 1.2

Der Übersetzungsschlüssel ÜS 1.2 liest KA3-Eingabedaten ein und erzeugt daraus übersetzte Daten gemäß der KA4. In bestimmten Fällen können manuelle Entscheidungen zur Übersetzung bestimmter Bodenparameter erforderlich sein. Da vorhandene Bodendaten nicht immer den Standards entsprechend erhoben wurden, können die Übersetzungsregeln an regionale oder andere spezielle Bedürfnisse angepasst werden. Das Ziel des Übersetzungsschlüssels ist es aber, nur die Anforderungen für die Eingabedaten den vorhandenen heterogenen Beschaffenheiten anzupassen. Um eine zukünftige standard-gemäße Harmonisierung der Bodendaten – insbesondere auch bei heterogenen

Eingabedaten – zu erreichen, dürfen die Ausgabeparameter nicht verändert werden. Damit wird sichergestellt, dass die Ausgabedaten der einheitlichen KA4 entsprechen.

Die Bedienungsoberfläche des Übersetzungsschlüssels unterstützt die Nutzenden dabei, ihre eigenen Daten in die vom System vorausgesetzte Struktur abzubilden. Die Eingabedaten müssen zur Zeit im Access-Format vorliegen. Ein XML-Import ist als alternativer Eingabeweg geplant. Momentan werden geeignete XML-Spezifikationen für jeden der unterstützten Klassifikationsstandards von der BGR erarbeitet. Derzeit bietet der ÜS 1.2 bereits einen prototypischen KA4-XML-Export an.

Die Nutzenden haben die Wahl, jeweils einzelne Parameter in den KA4-Standard zu übersetzen und die Übersetzungsergebnisse einzeln zu überprüfen oder alle Parameter insgesamt in die KA4 zu überführen. Die Übersetzung wird mit Hilfe der Access-Abfragetechnik durchgeführt, die zulässt, bereits die Ergebnisse jeweils einer Zahl von Datensätzen zu betrachten und zu evaluieren, bevor der Übersetzungsprozess für alle Datensätze abgeschlossen ist. Die Speicherung des Übersetzungsergebnisses erfordert selbstverständlich die Übersetzung aller Datensätze, so dass hier je nach Umfang der Daten eine längere Laufzeit einkalkuliert werden muss.

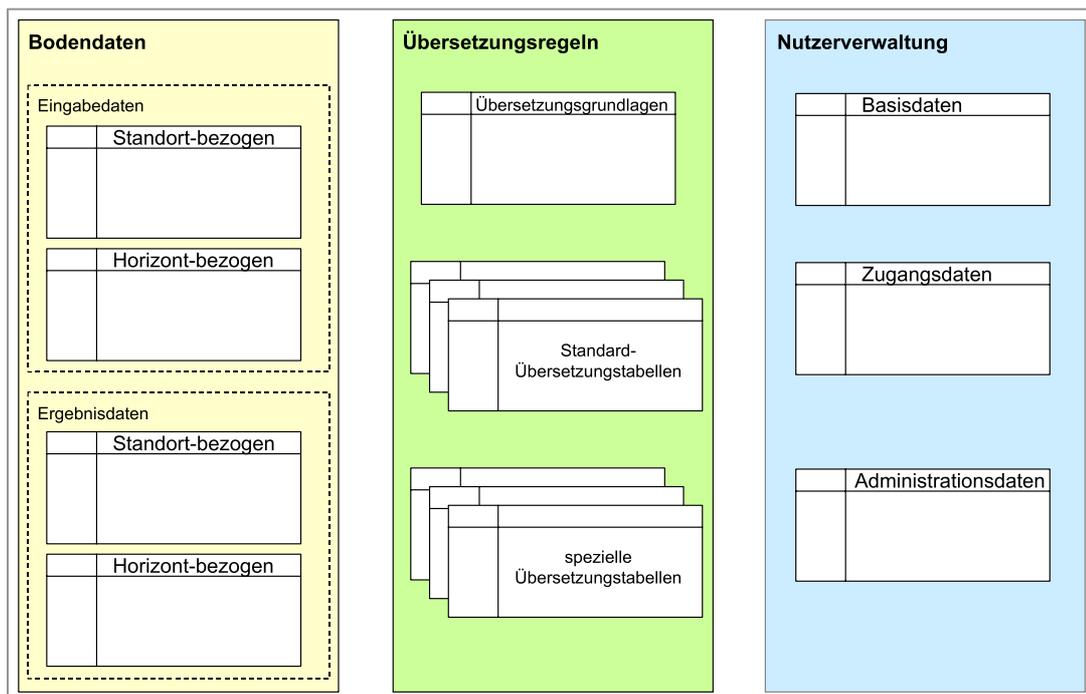


Abbildung 2: Internet-Übersetzungsschlüssel: vereinfachtes Datenmodell

### 3.3 Übersetzungsschlüssel ÜS 2.0

Der neue Access-basierte ÜS 2.0 besteht aus zwei Instanzen: Die eine Version überträgt KA4-Daten in die KA5, die andere übersetzt KA5-Daten in die internationalen Klassifikationen FAO/WRB. Der ÜS 2.0 KA4/KA5 ist äquivalent zum ÜS 1.2, nutzt aber andere Übersetzungsregeln und enthält weitere Optimierungen des Programmcodes.

Der ÜS 2.0 WRB befindet sich derzeit noch in der Entwicklung. In einem ersten Schritt werden die gleichen Übersetzungstechniken wie beim ÜS 1.2 und ÜS 2.0 KA4/KA5 implementiert. Da jedoch die Abbildung der deutschen Standards auf die internationalen Klassifikationen deutlich komplexer ist, müssen zusätzliche Arten von Übersetzungsregeln erarbeitet werden. Neue Übersetzungstabellen wurden eingerichtet, die Wahrscheinlichkeiten für den Einsatz bestimmter Übersetzungsregeln enthalten. In einigen Fällen kann so dem Umstand Rechnung getragen werden, dass mehr als ein Ausgabewert als korrekte Übersetzung in Frage kommt.



Abbildung 3: Internet-Anwendung: Startseite.

#### 4 Übersetzungsschlüssel ÜS 2.0 – Internet-Anwendung

Die Internet-Anwendung wurde realisiert mit PHP und verwendet Microsoft SQL Server als Datenbanksystem. Wie bei den Access-Anwendungen wird der Übersetzungsprozess durch Datenbank-Tabellen gesteuert, die die Übersetzungsregeln liefern. Bei beiden Architekturen, Access und Internet, werden die gleichen Übersetzungstabellen eingesetzt. Daher wurden die Anwendungen mit Administrationsfunktionen ausgestattet, die den Export der Übersetzungsregeln aus der Access-Version und den Import in die Internet-Anwendung ermöglichen. Damit wird der Aufwand zur Pflege der Übersetzungsregeln in der Internet-Version auf ein Minimum beschränkt. Ausschließlich die komplexeren Übersetzungsfunktionen müssen in äquivalenten PHP-Code umgesetzt werden.

Zur Nutzung der Internet-Anwendung müssen sich Interessierte zunächst beim System registrieren. Nach dem Erhalt der Zugangskennung können die Nutzenden eigene KA4-konforme Bodendaten im leicht mit den meisten Standard-Anwendungen

erzeugbaren CSV-Format<sup>7</sup> in das System hochladen. Für die nahe Zukunft ist eine zusätzliche Möglichkeit zum Daten-Import im XML-Format geplant. Nach dem Hochladen müssen die eigenen Datenstrukturen auf die Spezifikation des Systems abgebildet werden. Während einige Parameter optional angegeben werden können, wird ein Teil der Parameter als „Pflichtfelder“ angesehen, die angegeben werden müssen, um den Übersetzungsprozess zu starten. Derzeit können mit der Internet-Applikation Bodendaten in das KA5-Format überführt werden. Die Übersetzung in den FAO/WRB-Standard wird folgen, wenn der ÜS 2.0 WRB fertig gestellt ist (s. Kap. 3.3).

Die Übersetzungsergebnisse einschließlich der Fehlermeldungen und Warnungen werden in der Datenbank verwaltet. Nach 7 Tagen werden diese Daten auf dem Datenbankserver automatisch gelöscht. Während dieses Zeitraums können die Anwendenden ihre Daten im CSV-Format exportieren, was bequem in Anwendungen wie Microsoft Excel oder Access eingelesen und weiter verarbeitet werden kann. Eine XML-Ausgabe ist ebenfalls geplant.



Abbildung 4: Internet-Anwendung: Hauptauswahl-Seite.

## 5 Zusammenfassung und Ausblick

Die Hauptziele der beschriebenen Software-Lösungen bestehen in der Unterstützung der Harmonisierung von Bodendaten, die derzeit noch in verschiedenen Formaten der Bodenkundlichen Kartieranleitung (KA) vorliegen, und in der Verbesserung des internationalen Datenaustauschs durch die Erzeugung von

<sup>7</sup> CSV: Comma-separated values, ASCII-Dateiformat, bei dem die Werte jeden Datensatzes durch Komma oder andere Trennzeichen (meistens Semikolon) getrennt werden.

Bodenprofilbeschreibungen gemäß der neuen WRB-Klassifikation. Der zweite Schritt zur Entwicklung der digitalen Übersetzungsschlüssel ist noch nicht vollständig abgeschlossen, zeigt aber dennoch bereits wichtige Ergebnisse, indem hilfreiche Werkzeuge zur Harmonisierung von Bodendaten bereitgestellt werden.

Die Verkettung der einzelnen Übersetzungsprozesse ist möglich, so dass KA3-Daten auch in das KA5- oder in das WRB-Format überführt werden können, wenn die verschiedenen Versionen des Übersetzungsschlüssels nacheinander ausgeführt werden („Baukasten-Prinzip“). Es hat sich gezeigt, dass der Übersetzungsschlüssel nicht nur zum Transfer zwischen verschiedenen Klassifikationen genutzt werden kann, sondern auch zur Aufdeckung von Fehlern in den Eingabedaten hilfreich sein kann. Auf diese Weise trägt der Übersetzungsschlüssel zusätzlich zur Qualitätsverbesserung von Bodeninformationen bei. Während die Access-Versionen eher für Häufig-Nutzende geeignet sind, die auch eigene Anpassungen vornehmen möchten, bietet die Internet-Anwendung einen schnellen und leicht zu nutzenden Weg zur Übersetzung von Bodendaten.

Momentan befinden sich die aktuellen Software-Produkte noch in der Evaluationsphase und stehen daher öffentlich noch nicht zur Verfügung. Die Übersetzungsqualität wird mit unterschiedlichen Datensätzen geprüft, um die Auswirkungen der aktuellen Optimierungen zu verifizieren. Während die Entwicklung des Übersetzungsprozesses zur Transformation verschiedener deutscher Kartieranleitungen KA3/KA4/KA5 bereits zum größten Teil abgeschlossen ist, erfordert die automatisierte Erzeugung von Daten gemäß FAO/WRB noch Forschungsarbeit.

## 6 Literaturverzeichnis

- Adler, G. et al. (eds., 2004): Übersetzungsschlüssel zum Transfer von Bodendaten aus den deutschen Klassifikationen KA3 nach KA4 und von KA4 in die Internationale Klassifikation WRB und FAO, Arbeitsanleitung, Version 1.0, Arbeitshefte - Boden, 2004 Heft 1.
- AG Bodenkunde (1982): Bodenkundliche Kartieranleitung, 3. Auflage. Bundesanstalt für Geowissenschaften und Rohstoffe und Geologische Landesämter der Bundesrepublik Deutschland (eds.), Hannover, 331pp.
- AG Bodenkunde (1995). Bodenkundliche Kartieranleitung, 4. Auflage. Bundesanstalt für Geowissenschaften und Rohstoffe und Geologische Landesämter der Bundesrepublik Deutschland (eds.), Hannover.
- Ad-hoc-AG-Boden (2005): Bodenkundliche Kartieranleitung, 5. Auflage. Bundesanstalt für Geowissenschaften und Rohstoffe und Geologische Landesämter der Bundesrepublik Deutschland (eds.), Hannover, 438 pp.
- FAO (2006): Guidelines for soil description. 4th edition. Rome.
- FAO, ISRIC, IUSS (2006): World reference base for soil resources 2006. A framework for international classification correlation and communication. World soil resources reports (103), 128pp.
- Kneib, W. D., Bongard, B. & Nachbar, M. (1998): Vergleich von Daten der KA3 – KA4 – FAO im Hinblick auf deren Nutzung für die Erstellung von nationalen und internationalen Bodenkarten. – büro für bodenbewertung, Kiel

(Forschungsbericht i.A. der Bundesanst. f. Geowiss. u. Rohst, Hannover;  
unveröff.).

# Datenlogger-gestütztes Grundwassermonitoring

Dr.-Ing. Thomas Gutzke

Dr.-Ing. Gerrit Seewald

[gutzke@envi-systems.com](mailto:gutzke@envi-systems.com)

[seewald@envi-systems.com](mailto:seewald@envi-systems.com)

## Einleitung

Deutschland ist aus Sicht der Wasserversorgung ein begünstigtes Land. Über die Ressource Grundwasser wird ca. 70 % des Trinkwasserbedarfs gedeckt. Für die Gewinnung, Aufbereitung und Weiterleitung sind die ortsansässigen Wasserversorgungsunternehmen zuständig. Alle getätigten Förderaktivitäten sowie die Grundwassersituation müssen dabei dokumentiert und ausgewertet werden. Datenlogger übernehmen in diesem Prozess sukzessiv die manuelle und kostenintensive Grundwasserstandserfassung. Der sinnvolle Einsatz von Datenloggern im Bereich Grundwassermonitoring muss jedoch unter verschiedenen Gesichtspunkten analysiert und bewertet werden. Neben finanziellen Aspekten spielt hierbei insbesondere die Problematik der automatisierten Datenübernahme, -prüfung und -filterung eine große Rolle.

Nach der Datenübernahme erfolgt die Weiterverarbeitung, Aus- und Bewertung in sogenannten Monitoringberichten, die in regelmäßigen Intervallen an die zuständigen Behörden übermittelt werden müssen. Anhand dieser Berichte, eigener Landesmessstellen sowie der erteilten Wasserrechte erfolgt die Prüfung der Behörden hinsichtlich der Einhaltung maximaler Fördermengen Grenzgrundwasserständen etc., um ggf. steuernd eingreifen können.

## 1 Grundwassermonitoring

Aufwändig gestaltet sich für alle Beteiligten (Wasserversorger, Behörden etc.) die Überwachung der Grundwassersituation hinsichtlich des Wasserstandes. Hierfür werden die räumlich verteilten Grundwassermessstellen in einem regelmäßigen Turnus angefahren, der Abstich mit Hilfe eines Lichtlot ermittelt und diese Daten in das vom jeweiligen Unternehmen verwendete Bewirtschaftungssystem übernommen. Auf dieser Grundlage lassen sich Auswertungen in Form von Ganglinien (zeitlicher Verlauf an einem Ort) und 2 ½-dimensionale Pläne (räumliche Situation zu einem Zeitpunkt) erzeugen.



Abb. 1a: Grundwassermessstelle Abb. 1b: Ablesung mit Lichtlot

Die Erfassung der Grundwasserstände beinhaltet neben des personellen und zeitlichen Aufwands auch eine Reihe von Problemen und Fehlerquellen:

- Das Auffinden von Messstellen wird speziell durch starken Bewuchs erschwert
- In vielen Regionen (Wälder, Moorlandschaften etc.) gestaltet sich der Zugang häufig sehr schwierig und erfordert zusätzliche Hilfsmittel
- Nicht selten kommt es zu Verwechslungen von Messstellen
- Auch Meterfehler und Zahlendreher werden häufig erst bei der Verarbeitung der Rohdaten im Rahmen des Grundwassermonitorings entdeckt.

Eine nachträgliche Prüfung und Redigierung von Abstichswerten ist mit erneutem Personal- und Zeitaufwand verbunden und ist bei zu großem Zeitverzug häufig nicht mehr reproduzierbar.

## 1.1 Datenlogger

Eine Alternative zur beschriebenen herkömmlichen Erfassung des Grundwasserstandes bilden die so genannten Datenlogger. Datenlogger sind Speichereinheiten, die die von einem Sensor ermittelten Informationen in einem Ringspeicher ablegen.

Datenlogger werden seit Jahren im Bereich der Fließgewässer intensiv genutzt. In den letzten Jahren ist auch im Bereich des Grundwassermonitorings vermehrt der Einsatz von Datenloggern zu beobachten.

Gründe für den Einsatz von Datenloggern:

- Schwer auffindbare und schlecht erreichbare Messstellen
- Steigende Personalkosten
- Erhöhte Anforderungen (kürzere Messintervalle, erhöhte Datenverfügbarkeit)
- Technische Verbesserungen seitens der Datenlogger (längere Laufzeiten, weniger Drifts und Messlücken)
- Bessere und günstigere Möglichkeiten der Datenfernübertragung (DFÜ) mittels SMS etc.

Auf der anderen Seite sprechen eine Reihe von Punkten gegen einen Einsatz von Datenloggern:

- Hohe Anschaffungskosten
- Laufende Betriebskosten  
(Mitarbeiterschulung, Datenauslesung, Kontrollmessungen, Wartung, Reparatur)
- Mögliche Geräteausfälle (Messfehler, Messlücken, Drifts, Beschädigungen)
- Hohe Datendichte (und die damit verbundene problematische Weiterverarbeitung)
- Heterogene Datenformate der Hersteller (meist Ascii-Dateien) erfordern den Einsatz proprietärer Import- und Auswertungssoftware
- Leistungsschwache Auswertungssoftware

Insbesondere die herstellereigenen Auswertungswerkzeuge unterstützen den Anwender nur unzureichend bei der Fehleridentifizierung/-korrektur, Datenreduzierung und im Datenaustausch mit marktüblicher Bewirtschaftungssoftware.

Die Sensoren zur Erfassung der Grundwasserstände unterscheiden sich in ihrer prinzipiellen Funktionsweise (Winkelkodierer, Einperlschnur etc.) sowie deren technischen Ausführungen, wie Speichergöße, Ausleseschnittstellen, Messgenauigkeiten etc. Die gebräuchlichste Sensorart stellt der Drucksensor dar. Hier wird der Druck der Wassersäule abzüglich des Luftdrucks in frei einstellbaren Intervallen (meist zwischen 1 Sekunde und 1 Tag) gespeichert. Durch Kenntnis der geodätisch eingemessenen Messpunkt- bzw. Geländehöhe können NN-Höhe und Flurabstand ermittelt werden (Abb. 2).

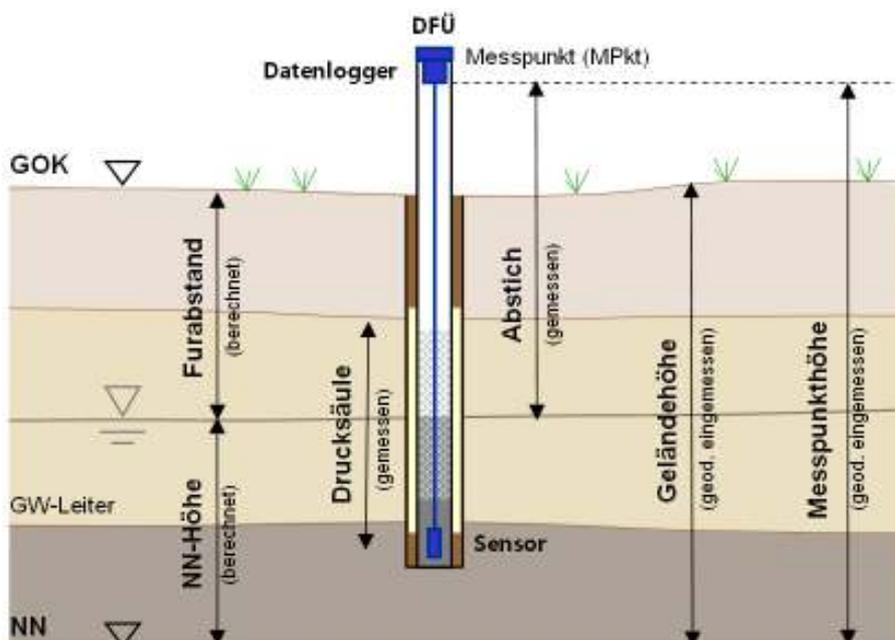


Abb. 2: Schematischer Aufbau der Grundwasserstandsmessung

Die Möglichkeiten des Auslesens der Loggerdaten sind mittlerweile vielfältig. Grundsätzlich muss hier zwischen Online- und Offline-Varianten unterschieden werden. Während bei einer Online-Verbindung eine Sendeeinheit am Datenlogger installiert sein muss (z.B. eine GSM-Karte), existieren eine Reihe von Schnittstellen zur manuellen Auslesung vor Ort, die die Datenübertragung auf ein Notebook oder einen PDA erlauben (Abb. 3).

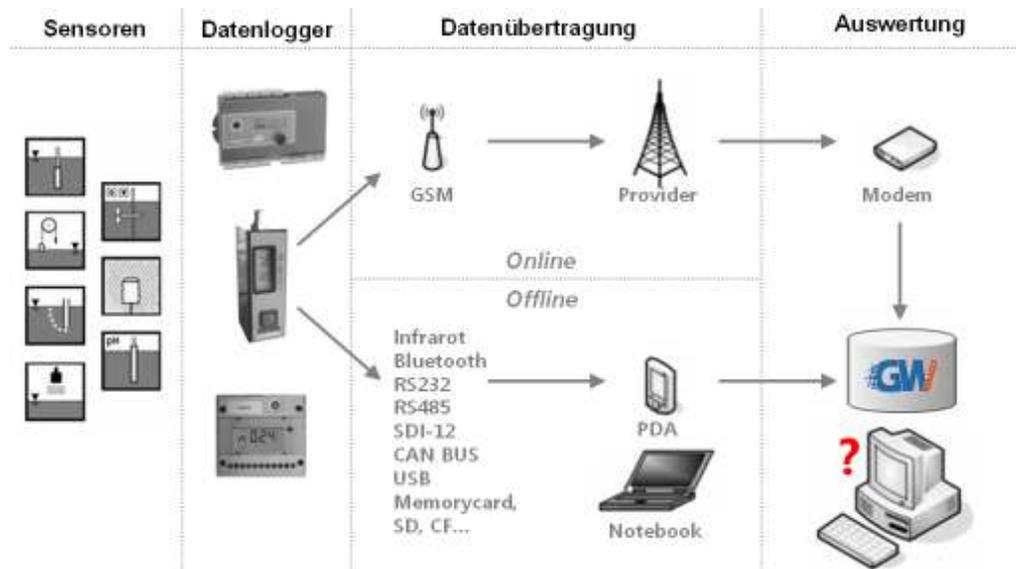


Abb. 3: Auslesen und Weiterverarbeiten von Loggerdaten

## 1.2 Konzept und Umsetzung

Auf Grund der problematischen und derzeit mehrstufig manuell vorzunehmenden Integration der Loggerdaten, wurde ein Konzept entwickelt, um eine automatisierte Datenübernahme zu realisieren (Abb. 4).

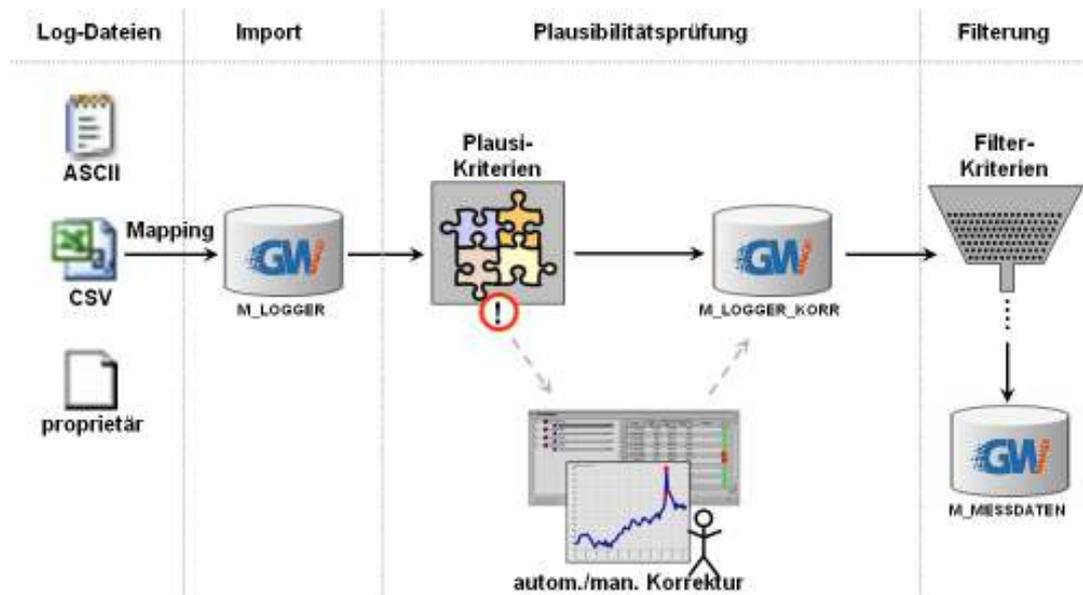


Abb. 4: Konzept zur automatischen Datenübernahme

Das Konzept umfasst dabei drei Prozesse.

### 1. Mapping und Import

In einem ersten Schritt werden die unterschiedlichen Herstellerformate der eingesetzten Datenlogger auf das gewünschte Zielformat gemappt (LIT-Import). Für das Mapping kommt das seit Jahren im Praxisbetrieb befindliche Modul ImportGWM des Softwaresystems GW-Manager zum Einsatz. Ein einmal erstelltes Mapping-Schema wird in einer XML-Datei gespeichert und kann so jederzeit wieder verwendet werden.

### 2. Plausibilitätsprüfung

Die in der Datenbank vorliegenden Daten werden in einem zweiten Schritt anhand von speziellen Kriterien hinsichtlich ihrer Plausibilität geprüft. Hierzu zählen:

- Über-/Unterschreitung von Grenzwerten
- Überschreitung von max. Abständen zu Vorwerten
- Korrelationsverhalten zu benachbarten Messstellen
- Vorhandensein von Messlücken

Darüber hinaus können mathematische Methoden eingesetzt werden, die es erlauben, auf Basis von Messreihenanalysen weiterführende Fehlerquellen aufzudecken. Entdeckt das System einen möglichen Fehler innerhalb einer Datenreihe, so erfolgt eine Benachrichtigung des Anwenders, der manuell die erforderlichen Korrekturen durchführen muss. Für bestimmte Fehlerarten (Messlücken, Ausreißer etc.) wird geprüft, ob eine automatische Korrektur ohne Benutzerinteraktion möglich ist. Alle Einstellungen zur Plausibilitätsprüfung sowie zur Korrektur werden in einer Projektdatei gespeichert.

### 3. Datenfilterung

In einem dritten Schritt werden die Daten gefiltert, d.h. sinnvoll reduziert. Während für die Identifikation von Messfehlern kurze Messintervalle hilfreich sein können, ist für die weitere Auswertung der Daten häufig eine reduzierte Datenmenge anzustreben (Abb. 5). Die Einstellungen für die Datenreduzierung erfolgt auf Objektbasis, so dass für jede Grundwassermessstelle individuelle Filterkriterien definiert werden können.

Filterschema Datei: C:\Filterschema\Filterschema\_Fuer\_Nord.xml  
 Eingaben Für Messtyp: Gueltig\_fuer\_Alle  
 Ein Filterschemadefinition für alle Messtypen  
 Für jede Messtyp unterschiedliche Filteroption definieren

Ziel-Datum in GW-Manager

Stündlich  
 Jede: [ ] Stunde

Täglich  
 Jeder: [ 1 ] Tag

Wöchentlich  
 Jeden: [ ]  
 alle 2 Wochen

Monatsdatum  
 1. Montag im Monat [DVWK-Stichtag]  
 Jeden [ ] [ ] im Monat  
 Datum wählen [ ]

benutze dafür in der (Logger-) Datei

Optionen

Zeitlich nächsten Wert zu [ 12:00 ] Uhr  
 Mittelwert von [ ] Einzelmessungen [ ] : [ ] Uhr  
 Statistischen Wert des betreffenden Tages : [ ]  
 Statistischen Wert der letzten : [ ] Tage : [ ]  
 1. Wert des Tages  
 Letzten Wert des Tages

Zusatzoptionen

"0" - Werte ausschließen  
 Tage mit unvollständigen  
 Logger-Messreihen ausschließen

Abb. 5: Filterkriterien

### 1.3 Zusammenfassung

Das beschriebene Konzept ermöglicht eine herstellerunabhängige Integration von Datenloggern in die Prozesse zur Überwachung und Steuerung der Grundwassersituation. Über den Schritt des Datenmappings können alle Datenloggerformate automatisiert in ein Zielsystem importiert werden. Dort findet eine unmittelbare Prüfung der Datenqualität statt. Nach erfolgreicher Prüfung bzw. Korrektur erfolgt eine anforderungsspezifische Reduzierung der Datenmenge auf Basis entsprechender Filterkriterien.

Aktuell wird im Hessischen Ried ein „Datenlogger-Projekt“ vorbereitet, bei dem die entwickelten Methoden im Praxisbetrieb angewandt und erprobt werden können. Ein wichtiger Aspekt wird die Betrachtung in Form einer Kosten-Nutzen-Analyse für den Einsatz von Datenloggern im Bereich des Grundwassermonitorings darstellen.

### 1.4 Technischer Ausblick

Großes Potenzial zur Weiterentwicklung basiert auf der umfangreichen, einheitlichen und aktuellen Informationsgrundlage, die nach der Importierung bzw. herkömmlichen Erfassung aller Grundwasserinformationen in der Zieldatenbank vorliegt. Hier ließen sich insbesondere Grundwassermodellierungs- und Prognosewerkzeuge anbinden. In Kombination mit einem regelbasierten System könnten so frühzeitig kritische Grundwassersituationen erkannt, die Entscheidungsträger unmittelbar über das System informiert und mit Vorschlägen für geeignete Gegenmaßnahmen versorgt werden [Rüppel, 2004].

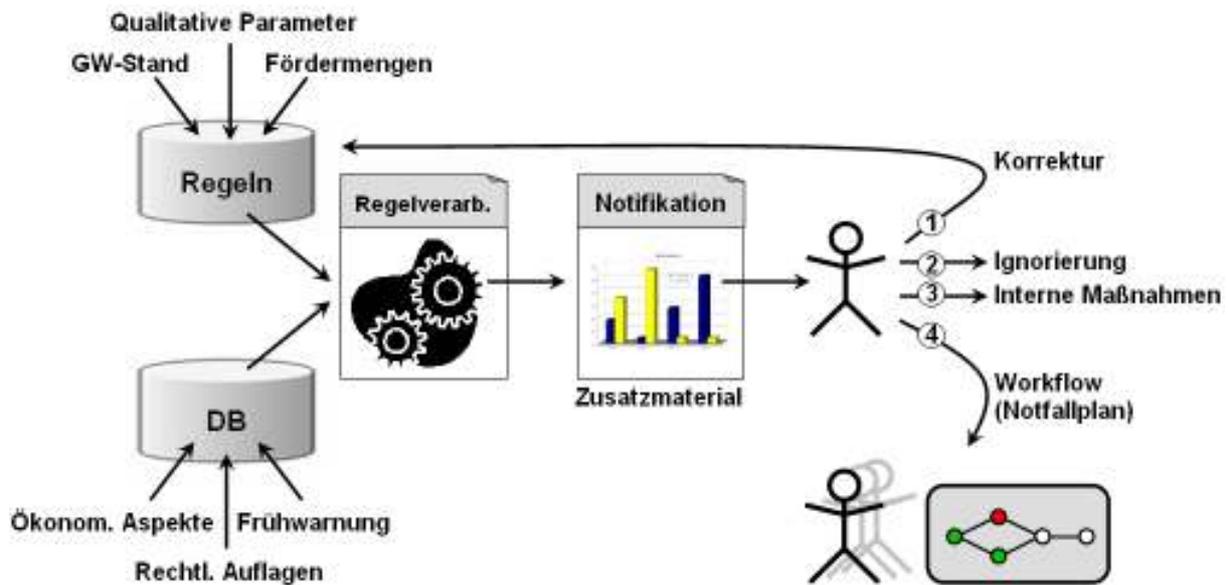


Abb. 6: Konzept für ein regelbasiertes Frühwarnsystem

Ein besonderer Schwerpunkt der weiteren Forschungsarbeiten (<http://www.dfg-spp1103.de>) wird die Modellierung der Notifikationsprozesse einnehmen, so dass die beteiligten Entscheider zeitnah informiert werden können.

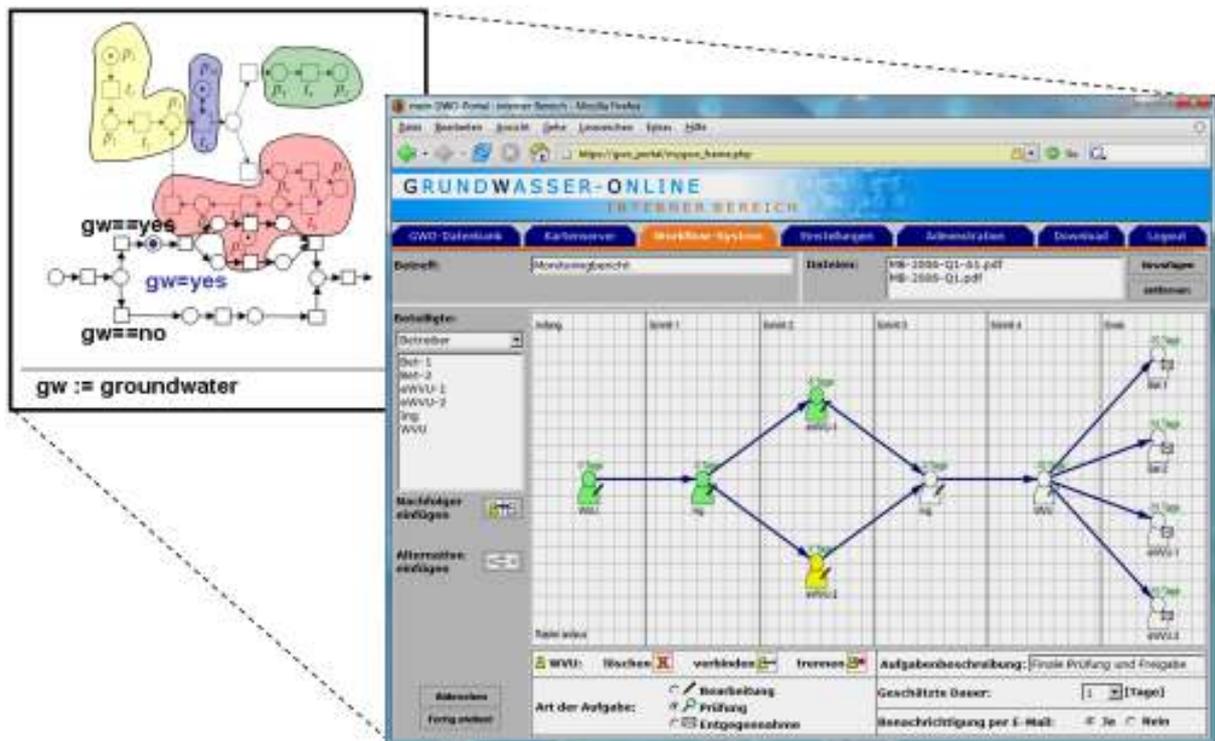


Abb. 7: Prozessbasiertes Notifikationsmanagement

Seitens der Hersteller von Datenlogger wird derzeit an Prozeduren der Echtzeit-Notifikation gearbeitet, die über die derzeit existierenden SMS-Benachrichtigungen bei Grenzwertverletzungen hinausgehen. Auch hinsichtlich der Einbindung qualitativer Sensoren ist eine rege Weiterentwicklung zu beobachten.



# **Data Warehouse Wasser – ein integratives Recherche- und Informationssystem innerhalb des bayerischen Informationssystems Wasserwirtschaft**

Anja Reineke

[anja.reineke@lfu.bayern.de](mailto:anja.reineke@lfu.bayern.de)

Claus Hofmann

[hofmann@disy.net](mailto:hofmann@disy.net)

## **Abstract / Einleitung**

Das Data Warehouse Wasser (DWW) ist Teil des vom Bayerischen Landesamt für Umwelt (LfU) entwickelten und betriebenen Informationssystems Wasserwirtschaft (INFO-Was) und bietet eine umfassende Sammlung von wasserwirtschaftlichen Daten zur Lösung von fachübergreifenden Fragestellungen. Der Datenpool enthält ca. 80.000 Objekte (inklusive 34.000 Messstellen) und ca. 13 Mio. Messwerte sowie eine Vielzahl von Geobasis- und Geofachdaten. Das DWW hat mit dem wasserwirtschaftlichen Datenpool die Aufgabe, Fachleute verschiedenster Disziplinen bei fachübergreifenden und unterschiedlich motivierten Auswertungen zu unterstützen. Es soll insbesondere als gemeinsame Datenbasis einfache und performante Zugriffe auf Sach- und Geodaten für Visualisierungen und Analysen sowie auf vorgefertigte Berichte mit aktuellen Daten ermöglichen. Zur Zielgruppe des DWW gehören primär die staatlichen Fachbehörden innerhalb des bayerischen Behördennetzes. Zu einem späteren Zeitpunkt soll der Anwenderkreis auch auf externe Nutzer erweitert werden (Internet-Lösung).

## **2 DWW als Teil des Informationssystems Wasserwirtschaft**

Das Informationssystem Wasserwirtschaft (INFO-Was) ist innerhalb des Geschäftsbereichs des Bayerischen Staatsministeriums für Umwelt, Gesundheit und Verbraucherschutz (StMUGV) das integrale Fachinformationssystem für den Bereich Wasser. Das gesamte System umfasst folgende verschiedene Bestandteile: Fachsysteme (Anwendungen und Datenbanken zur Pflege von wasserwirtschaftlich relevanten Fachdaten und deren Auswertung im fachlichen Kontext), das Geographische Informationssystem Wasserwirtschaft (GIS-Was) zur Erfassung raumbezogener Fachdaten und zur gemeinsamen Darstellung mit anderen Geodaten (Geobasisdaten und Fachdaten Dritter), das DWW sowie zukünftig das sich in der Konzeption befindliche Metainformationssystem Wasser (MiWas) für den integralen Datennachweis. Die Erfassung und primäre Speicherung von Stamm-, Mess- und Geofachdaten der Wasserwirtschaft finden innerhalb des INFO-Was grundsätzlich auf der Ebene der Fachsysteme und des GIS-Was durch die verschiedenen Dienststellen der Wasserwirtschaftsverwaltung (Wasserwirtschaftsämter,

Regierungen, LfU) statt. Dies bedeutet, dass alle wasserwirtschaftlichen Daten, die von verschiedenen Stellen genutzt und/oder für eine Veröffentlichung im Intra- oder Internet benötigt werden, in den Fachsystemen bzw. im GIS erhoben, gepflegt, auf Plausibilität geprüft und gespeichert werden. Die Daten können dort auch in ihrem fachlichen Kontext ausgewertet bzw. in Reports etc. präsentiert werden.

Zu den wichtigsten Fachsystemen, die als Quellsysteme für das DWW agieren, gehören zur Zeit das Fachanwendungspaket Grundwasser (Fachanwendungen Grundwasser, Wasserversorgung, Abfallentsorgung) sowie die Fachanwendungen Abwasser/Emissionen, Wasserkraftanlagen und Qualitative Hydrologie oberirdischer Gewässer. Technisch gesehen sind diese Anwendungen Client-Server-Systeme, installiert auf einer Terminalserverfarm mit einer Client-Anbindung auf Basis eines Citrix Presentation Server.

### 3 Ziele und Aufgabenstellung des DWW

Die Idee eines Data Warehouse Wasser ist aufgrund der verstärkten Nachfrage nach einem einfachen und performanten Zugriff auf die wesentlichen wasserwirtschaftlichen Objekte mit zugehörigen Messwerten zur fachübergreifenden Datenauswertung bzw. -präsentation entstanden. Durch den Aufbau einer gemeinsamen möglichst umfassenden Datenbasis zum Thema Wasser in einer zentralen Datenbank sollte einer möglichst breiten Anwendergruppe Möglichkeiten zur Recherche, Analyse und Präsentation von fachlichen Daten ermöglicht werden. Hierbei war die Integration von Fach- und Geodaten sowohl inhaltlich als auch technisch die zentrale Herausforderung.

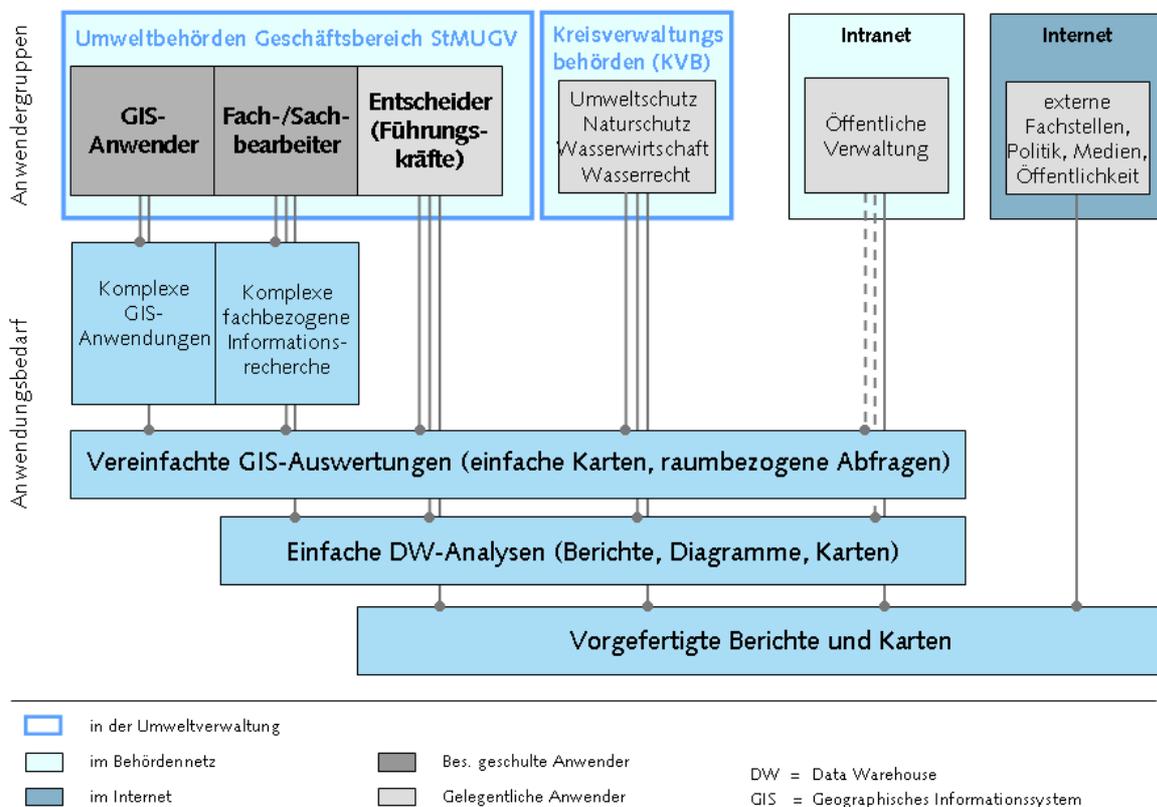


Abbildung1: Zielgruppenorientierte Lösungen im DWW

Die Abbildung 1 zeigt die Kategorisierung der verschiedenen Anwendungen und deren zugehörige Zielgruppen. Die Anforderungen an das DWW beginnen bei der Bereitstellung von Daten, Auswertungen und Berichten und setzen sich fort bis zur EDV-technischen Unterstützung von EU-Berichtspflichten.

## 4 Systemarchitektur

Das DWW besteht konzeptionell aus drei Ebenen (siehe Abbildung 2): Die unterste Ebene (Datenintegration) beinhaltet den Prozess des Datenimports von den verschiedenen Datenbanken der Fachsysteme und eines umfangreichen Geodatenbestandes in die Datenbank des DWW. Das Herzstück bildet die Ebene der integralen Datenhaltung in der DWW-Datenbank (Oracle, Version 9.2, mit Oracle Spatial) mit auswertebezogenem Datenmodell, das sowohl Sach- als auch Geometriedaten modelliert. Die DWW-Datenbank wird über sogenannte ETL-Prozesse (Extrahieren – Transformieren – Laden) aktuell gehalten. Datenauswertungen (Ebene 3) finden entweder auf Basis der Softwareplattform disy Cadenza oder mit GIS Werkzeugen statt.

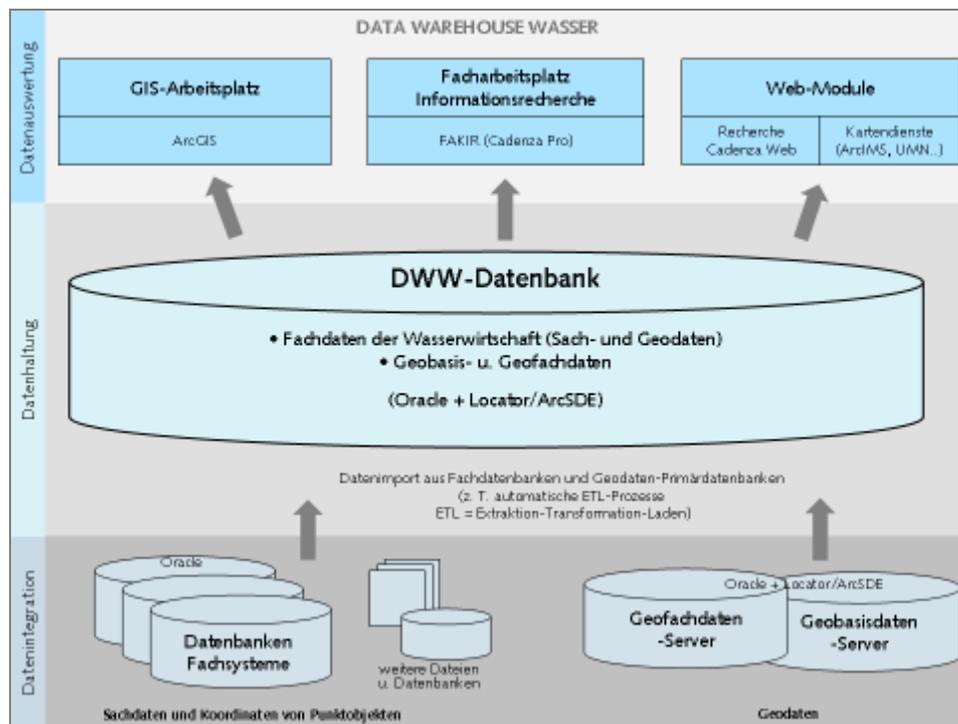


Abbildung 2: Systemarchitektur des DWW

### 4.1 DWW-Datenbank

Die Daten des DWW sind ausschließlich **Sekundärdaten**, die aus verschiedensten Datenquellen stammen können. Hierzu werden ausgewählte und bereinigte, z.T. aggregierte und normierte Daten aus den **Produktionssystemen** in die Datenbank überführt und dort redundant abgespeichert. Die Komplexität des Datenmodells ist gegenüber den Datenmodellen in den Fachsystemen wesentlich reduziert.

Im DWW befindet sich also eine reine Auswertedatenbank. Es findet keine Pflege von Daten statt. Die Einlagerung und Aktualisierung der Daten erfolgt ausschließlich

automatisiert durch ETL-Prozesse, die mit Hilfe des Oracle Warehouse Builder entwickelt werden. Einmal wöchentlich wird der gesamte Datenbestand aus den Fachsystemen aktualisiert.

Technische Voraussetzung für die Datenübernahme ist, dass die relevanten Daten zu den wasserwirtschaftlichen Objekten in einer geregelten Datenerfassung nach einheitlichen Vorgaben produziert und in definierten Strukturen im Primärsystem abgelegt wurden. Konsistenz und Qualität der Daten müssen auf Seiten der Primärdatenhaltung bzw. -pflege sichergestellt sein.

Die Auswahl der in den Datenpool aufzunehmenden Daten (Objekte, zugehörige Attribute bzw. Stammdaten, zugehörige Messdaten) erfolgt nach fachlichen Gesichtspunkten unter besonderer Berücksichtigung des Aspekts der raumbezogenen und fachübergreifenden Datenauswertung. Durch die gemeinsame Bereitstellung von Sach- und Geodaten sind kombinierte Raum-, Zeit- und Fachauswertungen auch direkt in der Datenbank möglich.

## 4.2 Facharbeitsplatz Komplexe Informationsrecherche

Eine wichtige Komponente des DWW ist der „**Facharbeitsplatz Komplexe Informationsrecherche**“ (FAKIR), der seit Juli 2006 in Produktionsbetrieb ist. Seit diesem Zeitpunkt sind ca. 60 Anwender geschult und die Anwendung wird täglich von ca. 5 bis 10 Usern genutzt.

Technologische Basis dieses Werkzeuges ist die Software Cadenza Professional (disy Informationssysteme GmbH), die auf eine Oracle-Datenbank mit Spatial- bzw. ArcSDE-Erweiterung zugreift. Des Weiteren werden andere Datenbanken und Kartendienste (WMS, ArcIMS) mit eingebunden. Das Verteilen der Client Software erfolgt über Java WebStart Technologie. Hierdurch erfolgen eine automatische Versionskontrolle und Aktualisierung der Client- Arbeitsplätze.

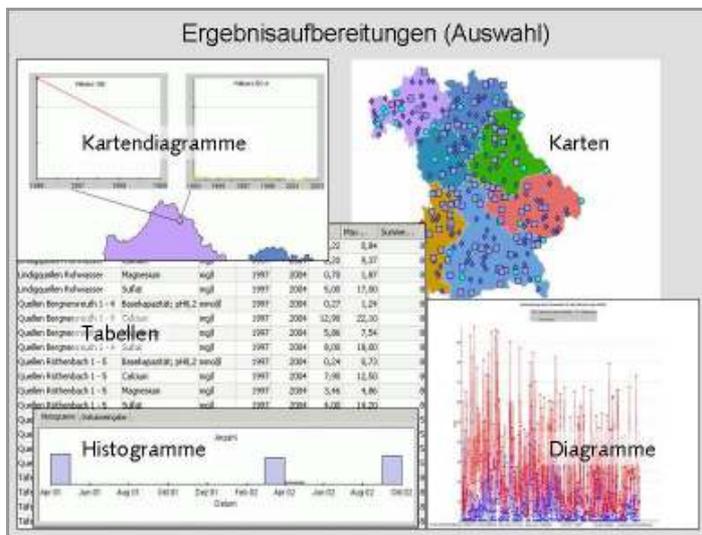


Abbildung 3: Ergebnisaufbereitungen im FAKIR (Beispiele)

Die vornehmliche Zielstellung dieses Facharbeitsplatzes ist der einfache, performante und fachübergreifende Datenzugriff auf Objekte aus den verschiedenen Fachbereichen. Neben der Wasserwirtschaft wird das zukünftig auch andere Bereiche der Umwelt und der Geologie betreffen. Weitere Zielstellungen sind:

- komfortable Navigation und Datenauswahl für einen breiten Anwenderkreis
- Datenzugriff auf einen beliebig auszuwählenden Datenumgriff zur Visualisierung und Analyse
- Datenrecherchen, statistische Auswertungen und Informationsbereitstellung zu Eckdaten für Übersichten, Führungsinformationen und strategische Entscheidungen
- Erfüllung von nationalen und internationalen Berichtspflichten
- Zugriff auf vorgefertigte Berichte mit jeweils aktuellem Dateninhalt

Eine Auswahl von typischen Ergebnissen zeigt die Abbildung 3. Neben beliebig auszuführenden Recherchen oder Auswertungen und deren Darstellung in Form von Tabellen und Karten bietet der FAKIR auch vorgefertigte Berichte, wobei jeweils der aktuelle Datenstand wiedergegeben werden kann. Exemplarisch zeigt Abbildung 4 das Rechercheformular für den Pflanzenschutzmittel-Bericht (PSM) nach Vorgaben des Umweltbundesamtes (UBA) und deren Ergebnisaufbereitung in Form eines Diagramms.

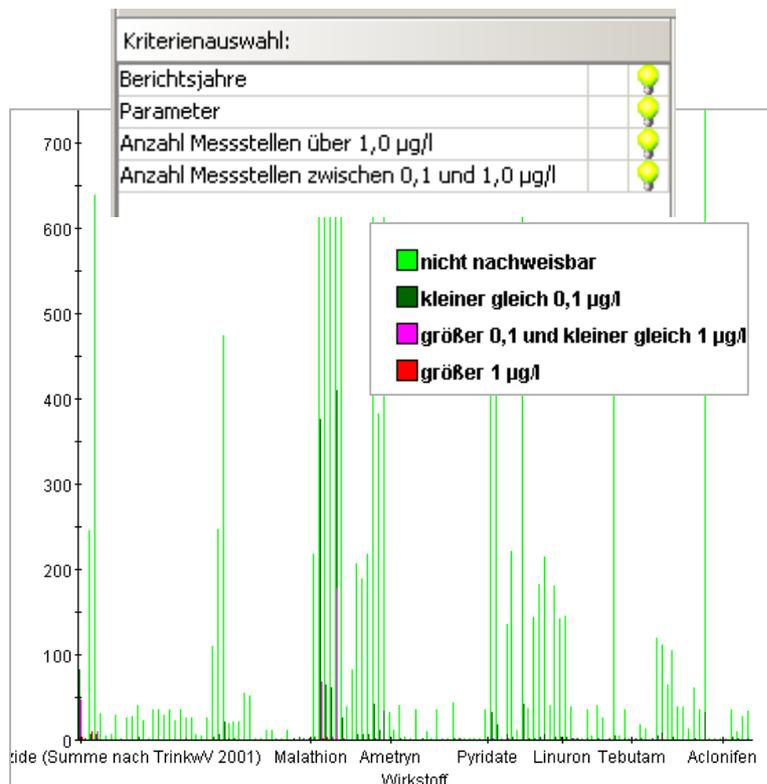


Abbildung 4: Anwendungsbeispiel Pflanzenschutzmittel-Bericht

Die Anwender des DWW sind eine sehr heterogene Gemeinschaft. Entsprechend sollten die Auswerteangebote den jeweiligen Zielgruppen gemäß angepasst werden. Eine der Hauptzielgruppen sind Führungskräfte und "Nicht-Spezialisten", die die Auswertungen ohne besondere Vorkenntnisse ausführen können sollten. Dies bedingt eine besonders sorgfältige Aufbereitung der Daten. Deshalb wird auf eine möglichst intuitive Benutzerführung, eine umfassende Dokumentation (Online-Hilfe, Online-Informationssdienst) sowie auf den Betrieb von Anwenderforen und einen anwendergerechten Support großen Wert gelegt.

## 5 Ausblick

Die Herausforderungen der Zukunft sind sowohl im technischen als auch im fachlichen Bereich zu sehen. Zum einen werden die zu integrierenden Datenmengen immer größer, d.h. die Datenmodellierung und Anwendungsentwicklung immer komplexer. Zum anderen gewinnt eine zentrale fachliche Koordinierung stärker an Bedeutung. Eine standardisierte Dokumentation in Form von Metadaten ist dabei dringend erforderlich.

## 6 Literaturverzeichnis

Beer, Michaela; Hofmann, Claus: Data Warehouse Wasser. In: move moderne  
verwaltung. Tübingen, September 2005

Hofmann, Claus et al.: disy Cadenza Übersicht und ausgewählte Lösungsbeispiele  
für Berichts- und Auswertesysteme. In: R. Mayer-Föll, A. Keitel, W. Geiger  
(Hrsg.): Projekt AJA Anwendung JAVA-basierter und anderer leistungsfähiger  
Lösungen in den Bereichen Umwelt, Verkehr und Verwaltung Phase IV – 2003.  
Karlsruhe, 2003

Landesamt für Wasserwirtschaft; disy Informationssysteme GmbH:  
 Projektdokumentation zur Erstellung von Pilotanwendungen für das Data  
Warehouse Wasser (DWW). Technische Infrastruktur und Fachszenarien.  
München 2004 (unveröffentlicht)

Landesamt für Umwelt: Gesamtkonzept Informationssystem Wasserwirtschaft –  
INFO-Was. München 2005 (unveröffentlicht)

# **Dezentrale EKVO-Datenerfassung mit einer Rich-Client-Anwendung**

## **- Praktische Erfahrungen mit Offline-Datenbanken -**

Ingo Mohr, Tilman Fischer, Hui Gao, Jo van Nouhuys

[im@condat.de](mailto:im@condat.de); [tf@condat.de](mailto:tf@condat.de); [gao@condat.de](mailto:gao@condat.de); [jvn@condat.de](mailto:jvn@condat.de)

Wolfgang Müller

[w.mueller@hlug.de](mailto:w.mueller@hlug.de)

Hubertus Pfaff

[hu.pfaff@rpu-wz.hessen.de](mailto:hu.pfaff@rpu-wz.hessen.de)

### **Abstract / Einleitung**

Betreiber von Abwasseranlagen, die gemäß der Eigenkontrollverordnung des Landes Hessen Jahresberichte erstellen müssen, und die zuständigen Wasserbehörden werden seit 2007 durch eine völlig neue Software unterstützt. Mit ihrer Hilfe werden Jahresberichte erstellt, geprüft, gedruckt und versendet und durch die Wasserbehörden in die Fachanwendung „Hessische Abwasseranlagen“ übernommen. Um möglichst viele Betreiber zur Nutzung zu motivieren, wurde eine Rich-Client-Anwendung realisiert, die auch ohne Internet-Zugang nutzbar ist. Jeder Betreiber verfügt mit dieser Software über eine eigene Datenbank, aus der Berichte in die zentrale Datenbank der Fachanwendung übertragen werden.

Im Beitrag wird erläutert, wie die Anforderungen auf Basis der Eclipse-Rich-Client-Plattform umgesetzt wurden. Es werden einige Lösungen vorgestellt, die das reibungslose Zusammenwirken einer Vielzahl von Anwendungen mit eigener Datenbank zu ermöglichen.

### **1 Rich-Client-Anwendung statt Web-Anwendung**

Betreiber von Abwasseranlagen sind verpflichtet, gemäß der Eigenkontrollverordnung des Landes Hessen [EKVO, 2000] Jahresberichte zu erstellen. Diese Berichte enthalten Informationen zu Größe, Zustand und Betrieb von Kläranlagen und Kanalnetzen, sowie Angaben zu Schadstoffen und Entsorgung. Um den Arbeitsaufwand in der Verwaltung gering zu halten, sollen möglichst viele der derzeit ca. 550 und zukünftig nach der Einbeziehung von Kleinkläranlagen ca. 7000 berichtspflichtigen Betreiber motiviert werden, die Berichte per Software zu erfassen,

statt sie in Papierform an die zuständigen Wasserbehörden zu übermitteln. Insbesondere sollte das Erstellen und Versenden von Berichten auch ohne Internet-Zugang möglich sein.

Aufgrund von Gesetzesänderungen war kurzfristig eine neue Software-Lösung notwendig. Die Fachanwendung „Hessisches Abwasseranlagenkataster“ [HAA] der zuständigen Wasserbehörden wurde außerdem gerade im Zuge der Zentralisierung des Rechenbetriebs in das Windows-Terminal-Server-System (WTS) übernommen. Kurzfristig stand keine entsprechende Web-Server-Infrastruktur zur Verfügung. Deshalb sollte die Übernahme und Verwaltung der Berichte in den Behörden per Windows-Anwendung erfolgen.

Somit lag es nahe, die Anwendungen für die Betreiber und für die Wasserbehörden einheitlich als Rich-Client-Anwendung zu realisieren. Dies würde es auch erlauben, viele Komponenten sowohl in der Fachanwendung als auch in der Erfassungsanwendung zu nutzen.

Die beschriebenen EKVO-Anwendungen wurden im Auftrage des Hessischen Landesamtes für Umwelt und Geologie (HLUG) [HLUG] von der Condat AG Berlin realisiert und 2007 zur Benutzung bereitgestellt [HLUG-EKVO].

## **2 Anwendungen basierend auf Eclipse-RCP-Komponenten**

Die EKVO wird durch zwei Anwendungen unterstützt, die hauptsächlich folgende Aufgaben erfüllen:

Die Erfassungsanwendung ermöglicht das Erstellen und Verwalten von Berichten und den Versand der Berichte an die zuständige Wasserbehörde.

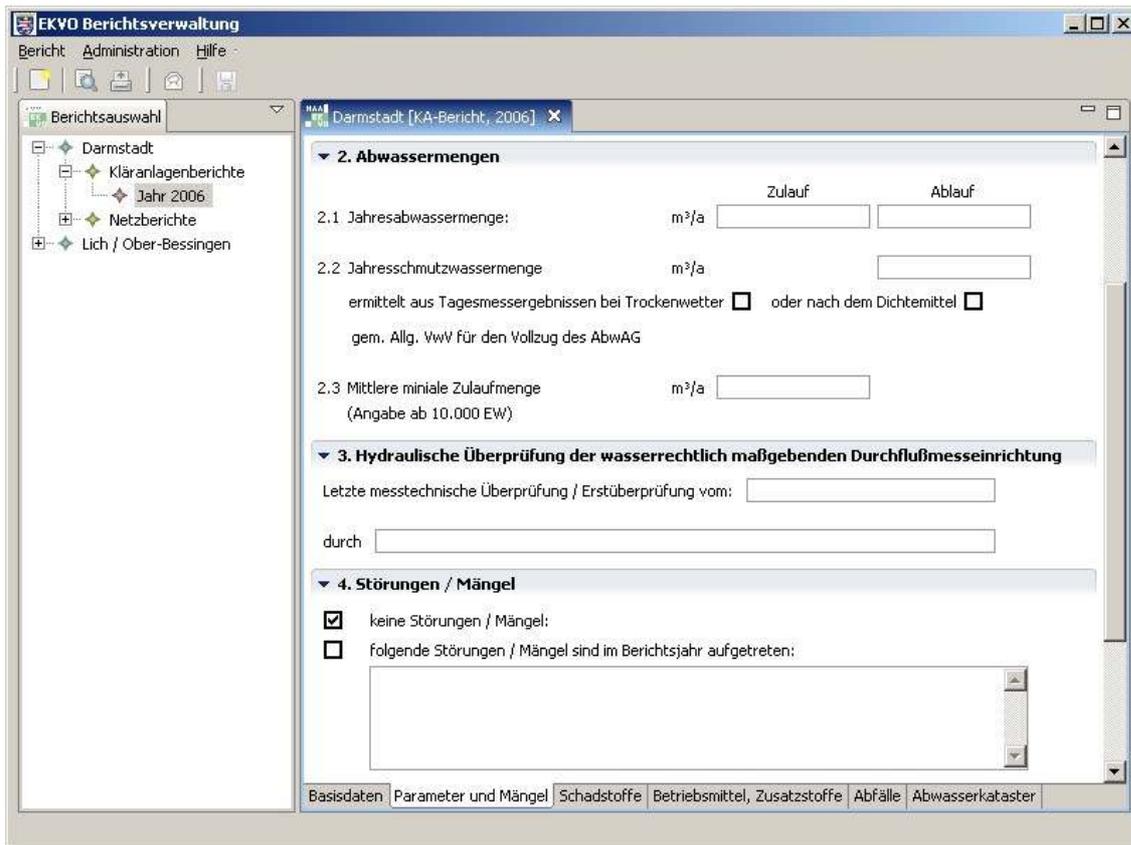


Abbildung 1: Erfassungsanwendung für Betreiber

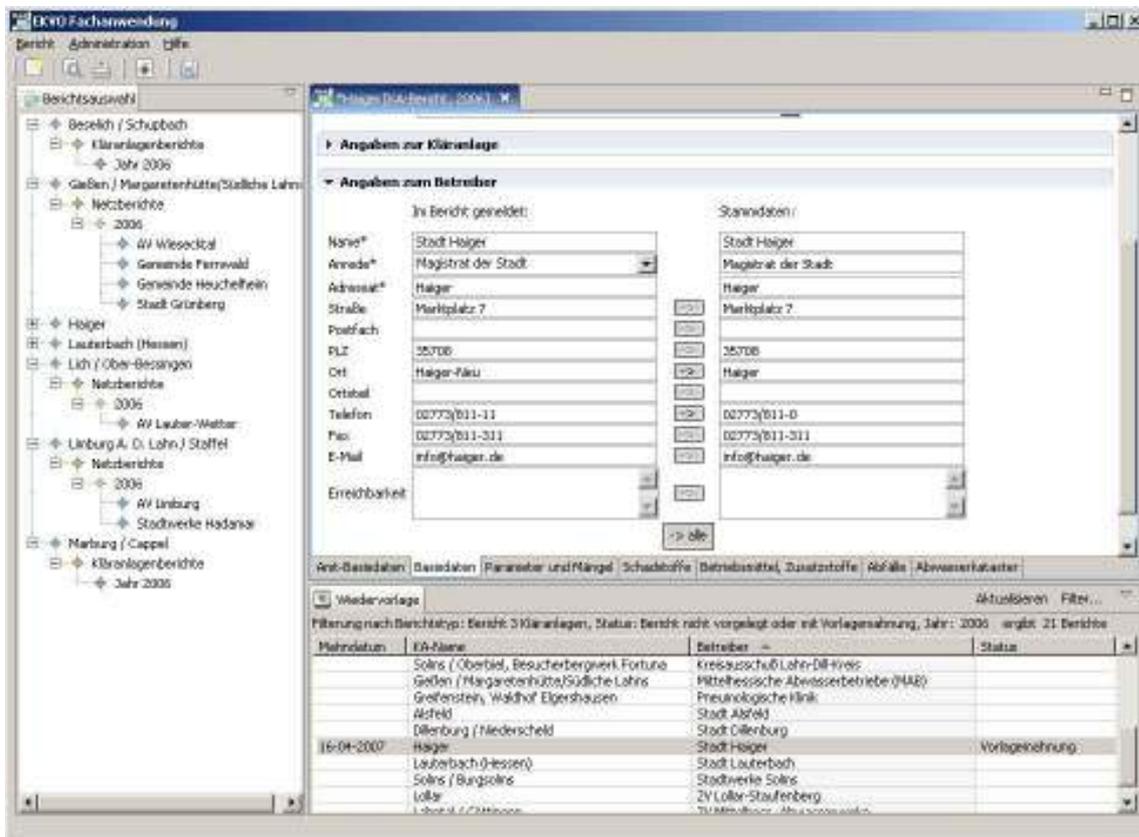


Abbildung 2: Fachanwendungskomponente der Wasserbehörden

- Die Fachanwendungskomponente ist eine eigene Anwendung im Rahmen der Fachanwendung „Hessisches Abwasseranlagenkataster“. Sie ermöglicht die Übernahme der von den Betreibern versendeten Berichte und deren Verwaltung. Hinzu kommt das Erfassen von Berichten, die in Papierform geliefert wurden. Außerdem gibt es eine Funktion für die Wiedervorlage und das Erstellen und Verwalten von Mahnungen. Siehe

Die Aufzählung läßt erkennen, dass ein Teil der Aufgaben identisch oder ähnlich sind. Nicht zuletzt aus diesem Grunde war es geboten, mehrfach verwendbare Komponenten zu realisieren.

Die Anwendung für die Betreiber sollte einerseits auf allen verbreiteten Betriebssystemplattformen lauffähig sein und sich andererseits unter Windows wie eine ganz normale Windows-Anwendung verhalten.

Die Eclipse-Rich-Client-Plattform (Eclipse-RCP) [Eclipse RCP] erfüllt diese Anforderungen hervorragend. Java als Programmiersprache bietet die Plattformunabhängigkeit. Eclipse-RCP-Anwendungen bieten sich dem Windows-Anwender als ganz normale Windows-Exe dar.

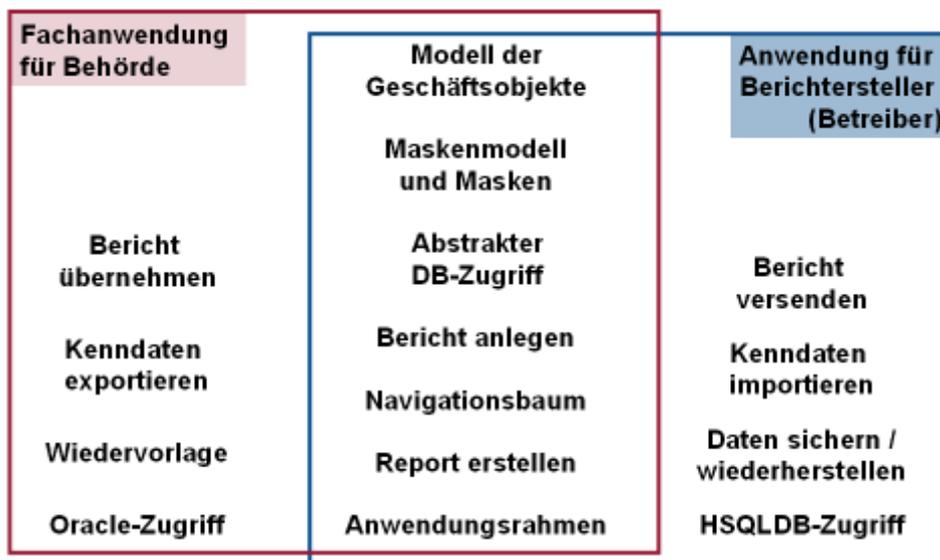


Abbildung 3: Anwendungen, zusammengestellt aus Komponenten

Eine Java-Laufzeitumgebung läßt sich in eine RCP-Anwendung integrieren, so dass man nicht von einem installierten Java ausgehen muss und davon abhängig ist. Und nicht zuletzt bietet diese Plattform eine hervorragend für unsere Zwecke geeignete Komponententechnologie.

Aus einem Pool von Komponenten für einzelne Teilaufgaben werden die Anwendungen per Konfiguration zusammengestellt. Einige Komponenten sind in beiden Anwendungen vorhanden, andere nur in einer (siehe Abbildung 3).

### **3 Eine einfache, robuste Datenbank für Berichtsersteller**

Für die Verwaltung der Berichte mit der Erfassungsanwendung für die Anlagenbetreiber gab es zwei wichtige Anforderungen:

- Einerseits sollten Fachanwendung und Erfassungsanwendung möglichst viele gemeinsame Software-Komponenten gemeinsam nutzen können. Dies erforderte auch für die Erfassungsanwendung eine datenbankgestützte Lösung.
- Andererseits sollten die Anwender der Erfassungsanwendung gar nicht merken, dass sie mit einer Datenbank arbeiten. Sie sollten in keiner Weise mit dem Management einer Datenbank belastet werden.

Um den zweiten Punkt gerecht zu werden, schieden alle Lösungen aus, die in irgendeiner Form einen Datenbank-Server mit eigenem Prozeß erforderten, um Wechselwirkungen mit anderen Programmen auszuschließen.

Die Fachanwendung speichert ihre Daten in einer Oracle-Datenbank [Oracle]. Die Datenbank für die Erfassungsanwendung sollte natürlich über die gleiche Schnittstelle angesprochen werden können wie in der Fachanwendung, um gleiche Funktionen nicht doppelt implementieren zu müssen.

Notwendig war also eine Datenbank, die in die Anwendung integrierbar ist und sich über die gleiche Schnittstelle ansprechen läßt wie die Oracle-Datenbank.

Beides erreichten wir, indem wir HSQLDB [HSQLDB] einsetzten, eine vollständig in Java realisierte relationale SQL-Datenbank mit Open-Source-Lizenz. Sie läßt sich nicht nur völlig in eine Java-Anwendung integrieren, sondern ist zudem auch sehr klein (ca. 650 kByte) und für die zu verwaltenden Datenmengen äußerst effizient. Und sie findet Verwendung in verbreiteten Software-Produkten wie z.B. OpenOffice, Mathematica, Jira und JBoss, was Sicherheit bezüglich Qualität und langfristiger Unterstützung gibt.

Statt einen Datenbank-Server zu starten, muß sich unsere Anwendung lediglich mit der Datenbankdatei verbinden. Danach gibt es keinen Unterschied mehr zwischen dem Zugriff auf die Oracle-Datenbank und auf die HSQLDB. Und auf dem Rechner des Anwenders läuft trotzdem kein eigener Datenbankprozeß, ein Port ist ebenfalls nicht belegt.

### **4 Einheitliches Fachmodell im Zentrum der Anwendungen**

Im Zentrum der Anwendungen steht das Fachmodell. Der Ursprung dafür ist ein UML-Modell, aus dem die software-technische Umsetzung der Fachobjekte einschließlich ihrer Speicherung in der Datenbank generiert wird.

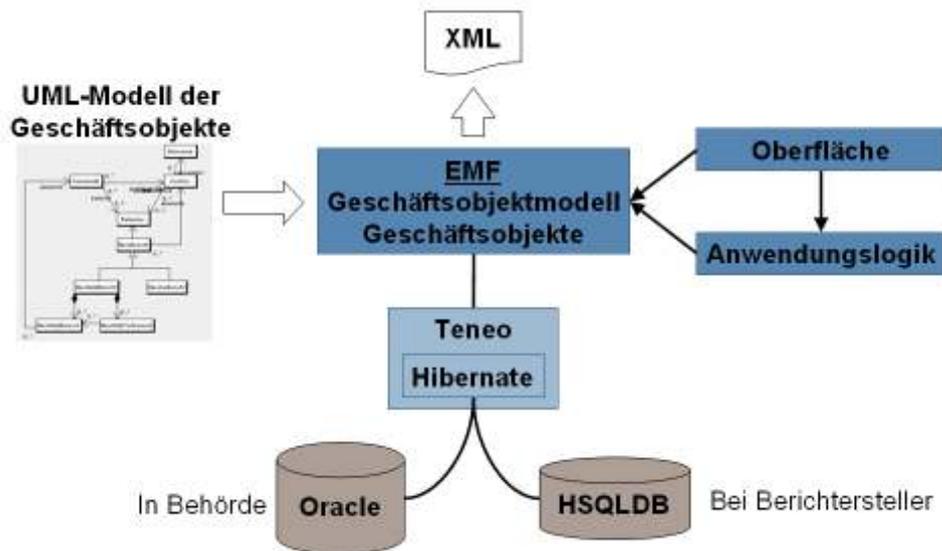


Abbildung 4: Geschäftsmodell im Zentrum der Anwendung

Bei der Entwicklung einer Anwendung auf Basis der Eclipse-Rich-Client-Plattform ist es nur konsequent, auch das Fachmodell auf der Grundlage des Eclipse Modeling Framework (EMF) [Budinsky, 2003] zu realisieren. Es bietet eine abstrakte Beschreibung von Objekten, auf die auch zur Laufzeit zugegriffen werden kann. Dies ist bei der Realisierung von generische Algorithmen sehr nützlich, z.B. um die gleiche Operation auf alle Objekte anzuwenden.

Aus dem Fachmodell wird eine hocheffiziente Java-Implementierung der Objekte generiert. Diese Objekte lassen sich ohne weiteres Zutun nach XML serialisieren und aus diesem Format wieder laden. Außerdem ist es möglich, einen für das Modell spezifischen Editor zu generieren, der sich bei Bedarf anpassen lässt und der es gestattet, Objekte des Fachmodells zu erfassen und in einer Datei abzuspeichern.

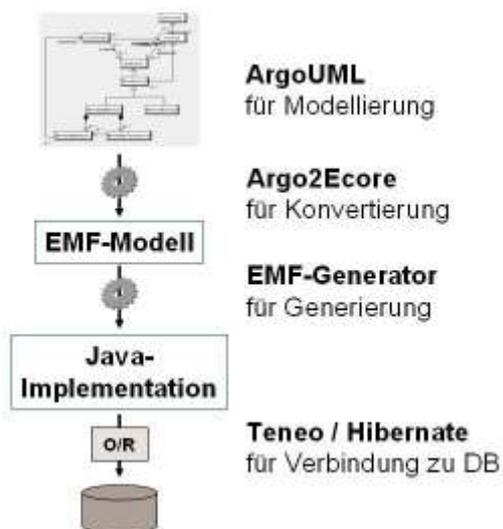


Abbildung 5: Werkzeugkette vom Modell zur Datenbank

Das Abspeichern der EMF-Objekte in eine Datenbank übernimmt eine RCP-Komponente namens Teneo [Teneo]. Diese nutzt wiederum das weit verbreitete Hibernate [Hibernate], um die Objektstrukturen auf Tabellen einer relationalen Datenbank abzubilden.

Alle Komponenten der Anwendungen, die Funktionen der Oberfläche oder der Anwendungslogik implementieren, sind nur vom Fachmodell abhängig. In welcher Datenbank die Speicherung erfolgt, ist ausschließlich in zwei datenbankspezifischen Komponenten implementiert, je eine für Oracle und für die HSQLDB.

Für die Erfassung der Berichte ist nur ein Teil des gesamten Fachmodells notwendig. Das Fachmodell hätte deshalb in zwei Teilmodelle aufgeteilt und der für die Behörden spezifische Teil nur in die Fachanwendungskomponente eingebunden werden können. Dies würde die Anwendung jedoch unnötig kompliziert machen. Andererseits soll die Persistenzschicht für die Erfassungsanwendung nicht mit unnötigen Dingen belastet werden. Aus diesem Grunde haben wir uns dafür entschieden, zwar das gesamte Fachmodell in beiden Anwendungen zur Verfügung zu stellen, aber für die Erfassungsanwendung nur einen Teil davon persistent zu machen. Dies lief darauf hinaus, für beide Anwendungen unterschiedliche Hibernate-Mapping-Dateien zu generieren, die die Speicherung der Objekte in der zugehörigen Datenbank beschreiben.

Insgesamt war es sehr einfach mit der Eclipse-Rich-Client-Plattform in Verbindung mit EMF und Hibernate, Anwendungen auf dem Fachmodell aufzubauen und die datenbankspezifischen Teile sauber von den anderen Komponenten zu trennen.

## 5 Datenaustausch mit Off-line-Datenbanken

Da eine Internet-Anbindung nicht vorausgesetzt werden konnte, mußte der gesamte Datenaustausch mit einfachsten Mitteln erfolgen.



Abbildung 6: Datenaustausch mit den off-line-Datenbanken

Für das Erstellen der Berichte sind bestimmte Stammdaten notwendig, die in der Fachanwendung HAA gepflegt werden. Beispielsweise Kläranlagennummern, Kanalnetze, Informationen zu den Betreibern oder eine Liste der Schadstoffe. Aktualisierungen dieser Informationen erhält der Betreiber, indem er sich eine sogenannte Kenndatendatei über den HLUG-Internet-Auftritt herunterlädt oder diese auf einem Datenträger anfordert. Diese Kenndaten enthalten alle diejenigen Stammdaten, die für die Berichterstellung mit der Erfassungsanwendung notwendig sind.

Der Versand der fertigen Berichte erfolgt normalerweise per Email mit einer XML-Datei als Anhang. Es kann jedoch auch ein Datenträger (z.B. CD oder Diskette) mit einer exportierten Berichtsdatei an die zuständige Behörde verschickt werden.

Alle Daten in der Fachanwendungsdatenbank, in den versendeten Dateien und in den off-line-Datenbanken der Erfassungsanwendung basieren auf dem gleichen EMF-Datenmodell. In Kenndatendatei und Berichtsdatei liegen die EMF-Objekte in XML-Form (genauer XMI) serialisiert vor. Die Verwendung eines einheitlichen EMF-Datenmodells und der XMI-Serialisierung macht den Datenaustausch zwischen der Fachanwendungsdatenbank und den off-line-Datenbanken technisch sehr einfach.

## **6 Berichtsversand per Email**

Zum Versand eines Berichts an die zuständige Behörde wird dieser mit Hilfe der EMF-nach-XMI-Serialisierung in eine XML-Datei exportiert. Die XML-Datei wird entweder einer Email als Anlage beigefügt oder auf einem Datenträger gespeichert und dann versendet.

Die XML-Datei mit dem Bericht wird in die Fachanwendung geladen und dabei durch den EMF-Deserialisierer aus dem XML-Format in die interne EMF-Struktur überführt. Überprüfungen erfolgen auf dem EMF-Objekt, bevor es in die Datenbank gespeichert wird.

In der Fachanwendung muß man davon ausgehen, dass Berichte möglicherweise mit älteren Versionen der Erfassungsanwendung erstellt oder dass nicht aktuelle Stammdatenstände verwendet wurden. Beides muß in ähnlicher Art und Weise berücksichtigt werden wie beim Transport von Berichtsdateien per Export und Import (siehe Abschnitt 8). Hier ist es notwendig, ein auf die spezielle Situation zugeschnittenes Verfahren zu finden, welches technisch funktioniert und fachlich korrekt ist. Beispielsweise gab es die Situation, dass Berichte durch nicht korrekte Stammdaten Informationen zu Kanalnetzen enthielten, die garnicht existierten. Hier wurde entschieden, diese Teile des Berichts bei der Übernahme in die Fachdatenbank zu löschen und die Bearbeiter entsprechend zu informieren.

Auch unerwartete Schwierigkeiten mußten überwunden werden. In manchen Konfigurationen von Email-Servern werden lange, sprechende Dateinamen durch generierte, nichtssagende Namen ersetzt. Die langen, sprechenden Namen sollen es den Sachbearbeitern bei der Übernahme jedoch erleichtern, die Berichtsdateien zu identifizieren, was nun nicht mehr möglich war. Abhilfe schaffte eine Funktion, die es gestattet, die ursprünglichen Dateinamen aus den darin enthaltenen Berichtsdaten wiederherzustellen.

## **7 Änderung und Synchronisation der Stammdaten**

In den Stammdaten zu den Betreibern und Anlagen sind im Laufe der Zeit Änderungen notwendig. Beispielsweise Name, Adresse, Ansprechpartner oder Telefonnummer. Die Pflege dieser Daten erfolgt zwar in der Fachanwendung HAA. Von Änderungen haben jedoch oft nur die Betreiber Kenntnis und müssen diese an die zuständigen Behörden übermitteln.

Um den Betreibern die Arbeit möglichst einfach zu machen, wurde entschieden, dass solche Stammdatenänderungen als Bestandteil der Berichte an die Behörde übermittelt werden. Dort können die Sachbearbeiter die neuen Werte in die Stammdaten der Fachanwendungsdatenbank übernehmen.

Haben sich die Stammdaten in der Fachanwendungsdatenbank geändert, kann eine neue Kenndatendatei erstellt und den Betreibern zur Verfügung gestellt werden, die sie in die Erfassungsanwendung laden können.

Möglicherweise stimmen jedoch vorhandene Stammdatenänderungen in der Erfassungsanwendung und neue Kenndaten nicht überein. Beispielsweise wenn Stammdatenänderungen bereits erfaßt, aber noch nicht mit dem Bericht versendet wurden. Deshalb hat der Anwender die Möglichkeit, bei Konflikten zwischen neuen Kenndaten und den lokalen Stammdaten zu entscheiden, ob die vorhandenen Daten erhalten oder durch die neuen Kenndaten überschrieben werden sollen.

Für einige Stammdaten ist es notwendig, sie mit einer Gültigkeitsinformation zu versehen. Beispielsweise wenn Kanalnetze umorganisiert werden, d.h. ab einem bestimmten Zeitpunkt andere Betreiber verantwortlich für sie sind. Bei der Kenndatenübernahme in die Erfassungsanwendung reicht es normalerweise aus, dass bestehende Berichte bleiben, wie sie sind und neue Berichte diese Informationen nicht mehr verwenden. Damit zum Beispiel keine Daten mehr zu einem bereits stillgelegten Kanalnetz erfaßt werden.

Wurde jedoch bereits ein neuer Bericht auf der Basis alter Stammdaten erfaßt, könnte ein Löschen der zugehörigen Daten unerwünscht sein. Hier wird (zukünftig) der Betreiber informiert, dass die entsprechenden Daten, z.B. ein Kanalnetz, ungültig sind. Er erhält die Möglichkeit die entsprechenden Datensätze zu löschen, kann aber bei Bedarf bereits erfaßte Daten vorher noch in neue Datensätze übertragen.

## **8 Berichtsverwaltung außerhalb der Datenbank**

Ursprünglich erfolgte die Verwaltung der Berichte beim Anlagenbetreiber ausschließlich in der Datenbank. Schon bald stellte sich jedoch heraus, dass Anwender auch Berichte in Dateien verwalten wollten. Dafür gab es verschiedene Gründe:

- Einige Anwender wollten ihre Daten unabhängig von der Anwendung sichern.
- Bei einigen Anwendern bearbeiteten mehrere Personen einen Bericht. Einige Betreiber beauftragen z.B. Ingenieurbüros mit der Berichterstellung. Sie wollen die Berichte jedoch auch selbst ansehen, prüfen und versenden können.

Es war also notwendig, dass man Berichte auch in eine Datei exportieren und daraus wieder importieren kann. Hier kam wie beim Berichtsversand der XML-Serialisierungsmechanismus von EMF zum Einsatz.

Eine weitgehende Unabhängigkeit der Berichtsdateien von der EKVO-Anwendung erreichten wir durch die Verwendung von XML und dadurch, dass von Berichten referenzierte Stammdatenobjekte vollständig mit in die Berichtsdatei geschrieben werden. So enthält eine solche XML-Datei nicht nur den eindeutigen Schlüssel einer Kläranlage, sondern auch deren Bezeichnung, die Anschrift usw. Neben dem Schlüssel eines Schadstoffes ist auch die vollständige Bezeichnung und die

Maßeinheit mit angegeben. Diese Redundanz erlaubt es, den Bericht auch unabhängig von den Stammdatentabellen eindeutig zu interpretieren.

Mit der Speicherung von Berichten in Dateien ergibt sich natürlich die Anforderung, dass diese Berichte auch nach Änderungen am Datenmodell, in der Anwendung oder bei den Stammdaten erkannt und geladen werden können. Spätere Software-Versionen müssen demzufolge ggf. das aktuelle Datenmodell und den aktuellen Lademechanismus für alte Berichte in einer zusätzlichen Komponente bereithalten.

## 9 Modellbasierte Entwicklung des Berichteditors

Um dem Anwender eine übersichtliche und ergonomische Bedienoberfläche zur Verfügung zu stellen, ist bei umfangreichen und/oder komplexen Daten normalerweise Handarbeit notwendig, um die Erfassungsmasken zu gestalten. Andererseits gibt es bei der Implementation solcher Masken viel Routinearbeit. Aus diesem Grunde haben wir ein Verfahren entwickelt, welches einerseits Routinearbeit durch Verwendung eines Editormodells reduziert und andererseits speziell gestaltete Masken ermöglicht.

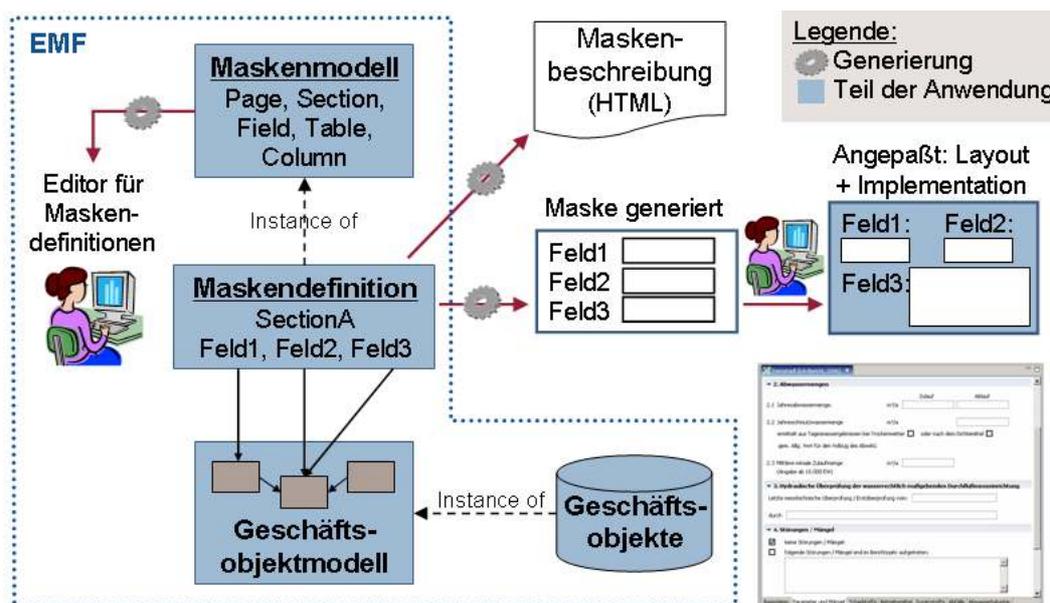


Abbildung 7: Modellbasierte Masken

Grundlage ist das abstrakte Maskenmodell, welches als UML-Modell realisiert ist [ArgoUML]. Daraus generieren wir ein EMF-Modell und erzeugen daraus wiederum die Java-Implementation des Modells und einen EMF-Editor zum Erfassen von Maskendefinitionen für die Masken des Berichtseditors.

Mit Hilfe dieses Maskendefinitions-Editors erfassen wir, in welche Seiten (Reiter) und Abschnitte der Berichtseditor aufgeteilt werden soll und welche Eingabefelder nötig sind. Für die Eingabefelder wird über Referenzen in das Fachmodell festgelegt, woher sie die Daten aus den Fachobjekten beziehen.

Für jede Maskendefinition wird eine Java-Klasse generiert, welche die Eingabemaske prototypisch implementiert. Diese Klasse muß mit Hilfe eines

Masken-Editors (Visual Editor von Eclipse [Visual Editor]) manuell im Layout optimiert werden. Außerdem ist die Implementation für spezielle Fälle anzupassen. Zum Beispiel bei speziellen Abhängigkeiten zwischen den Eingabefeldern.

Das geschilderte Vorgehen hat sich als ein sehr gute Balance zwischen spezifischer Implementierung und modellhafter Beschreibung zur Einsparung von Routinearbeit erwiesen. Hinzu kommt, dass dadurch die Robustheit der Anwendungen steigt und die Einheitlichkeit innerhalb der Anwendungen und zwischen ihnen höher ist als bei einer rein manuellen Implementierung.

## 10 Fazit und Ausblick

Im Gegensatz zu einer Anwendung mit zentraler Datenbank sind bei der beschriebenen Anwendungsstruktur mit einer Vielzahl eigenständiger Datenbanken – wie erwartet – einige spezielle Lösungen notwendig. Auf der anderen Seite konnte durch die Realisierung als Rich-Client-Anwendung bereits eine Erfassungsanwendung für EKVO-Berichte realisiert werden, ohne dass eine entsprechende Internet-basierte Infrastruktur für die Wasserbehörden und bei den Betreibern vorhanden ist. Hinzu kommt, dass eine sehr benutzungsfreundliche und schnell reagierende Anwendung entstanden ist. Dass sie sehr gut angenommen wird, zeigen entsprechende Reaktionen der Betreiber.

Von der Seite der Software-Entwicklung her hat sich gezeigt, dass die Eclipse-Rich-Client-Plattform hervorragend geeignet ist, um sehr ergonomische Anwendungen zu realisieren, die aus wiederverwendbaren, klar strukturierten Komponenten aufgebaut sind. Das Eclipse Modelling Framework ist eine sehr gute Grundlage für eine modellbasierte Entwicklung. Beide gemeinsam ersparen viel Entwicklungsaufwand und erhöhen die Robustheit der Anwendung, da an vielen Stellen händische und spezifische Implementierungen vermieden werden.

## 11 Literaturverzeichnis

[Argo2Ecore] <http://argo2ecore.sourceforge.net/>

[ArgoUML] <http://argouml.tigris.org/>

[Budinsky, 2003] Budinsky, Frank; Steinberg, David; Merks, Ed; Ellersick Raymond: Eclipse Modeling Framework. Prentice Hall, 2003.

[Daum, 2006] Daum, Berthold: Rich-Client-Entwicklung mit Eclipse 3.2. Anwendungen entwickeln mit der Rich Client Platform. Dpunkt Verlag, 2006.

[Eclipse RCP] <http://www.eclipse.org/home/categories/rcp.php>

[EKVO, 2000] Verordnung über die Eigenkontrolle von Abwasseranlagen (Abwassereigenkontrollverordnung - EKVO), 21. Januar 2000, (GVBl. I S. 59)

[HAA] <http://www.hlug.de/medien/wasser/abwasser/haa.htm>

[Hibernate] <http://www.hibernate.org/>

[HLUG] <http://www.hlug.de/>

[HLUG-EKVO] [http://www.hlug.de/medien/wasser/abwasser/download\\_ekvo.php](http://www.hlug.de/medien/wasser/abwasser/download_ekvo.php)

[HSQLDB] <http://www.hsqldb.org/>

[Oracle] <http://www.oracle.com/global/de/index.html>

[Teneo] <http://www.elver.org/>

[Visual Editor] <http://www.eclipse.org/vep/>

# A Modeling Methodology for Scientific Processes

Stefan Jablonski, Bernhard Volz, M. Abdul Rehman

[stefan.jablonski@uni-bayreuth.de](mailto:stefan.jablonski@uni-bayreuth.de)

[bernhard.volz@uni-bayreuth.de](mailto:bernhard.volz@uni-bayreuth.de)

[abdul.rehman@uni-bayreuth.de](mailto:abdul.rehman@uni-bayreuth.de)

## Abstract

Scientists are more and more dealing with complex scenarios during their daily research work. Processes are being used to describe these scenarios and help to understand them – on the outside thanks to graphical notations, on the inside thanks to structured methods which are necessary to conquer the information flood. However many modeling systems and their underlying methodology are confronting problems during the development of a scientific workflow. We want to present our methodology which eliminates these pitfalls and demonstrates its usage in the scientific application domain with an example from ecological research.

## 1 Assessment of Scientific Workflow Systems

Processes are being used in many areas for the description of complex scenarios ranging from business management [Jablonski & Bussler, 1996] over quality management [Jablonski et al, 2006a] up to health and medical application domains [Meiler, 2005]. It is only consequent to also use process management in scientific applications. Thus many systems supporting process management have been developed in the past years, like Taverna [Hull et al., 2006], Triana [Taylor et al, 2007], Kepler [Altintas et al., 2004], Askalon [Qin et al, 2006] and Pegasus [Deelman et al., 2005]. These systems are mainly focusing on the underlying implementation and execution of workflows (we want to use the terms "process" and "workflow" synonymously within this section) and are following a bottom-up approach. Their abstract workflow model is directly influenced and determined by its underlying implementation, i.e. that scientist must know technical details about things to use. For instance, they must know about the existence of a concrete service and about how to access it (e.g. as a web or grid service) or about the concrete parameters to be passed to it. This renders these systems unusable for the earlier phases of the development of such a workflow since, at that stage, scientist should more be supported in their contemplation about principle solution strategies. Instead, we offer a top-down approach for workflow management. In the first phase, scientists model the processes on a very abstract level without knowing implementation details. Therefore, we pursue a perspective oriented approach. By tuning any perspective of a process, any change in implementation can be reflected without changing the abstract workflow model.

Furthermore, the above systems do not support separate data models which define the structure and the semantics of data by using ontologies, terminologies and format

descriptions. Having a separate data model and therewith denoting the flow of data in the process model would help to increase correctness and adequateness of scientific processes. It can easily be seen from the data model where data is transformed and how it is manipulated.

Also responsibilities during the execution cannot be declared in most of the above mentioned systems. To explicitly assign responsibilities to workflows will help in auditing and tracking workflow executions.

Last but not least there is no possibility for the interaction between the users of a workflow and running workflow instances. This is needed when a process cannot be executed automatically but needs the assistance of a person (i.e manually).

Current systems for scientific workflows aim at the orchestration of existing services to form new services/processes. They provide good possibilities for the execution of workflows but they do not support the conceptual development phases of scientific workflow applications. With the introduction of our perspective oriented workflow model, we want to contribute to systematic workflow development.

## **2 Design methodology for scientific workflows**

Our contribution to process modeling for the scientific domain is based on a top-down approach. First of all we want to distinguish processes from workflows. Processes are abstract and cannot be executed because they usually lack information which is needed for execution (e.g. concrete location of data files). A workflow – that can be derived from a process – is executable, i.e. these implementation specific details are added (e.g. the exact location of a web service implementation).

Our modeling tool iPM (integrated Process Manager [ProDatO, 2005]) supports the perspective oriented process modeling (POPM) approach which we already applied successfully to various application domains (cf. Section 1). The idea behind POPM is that a modeling construct (such as a process) consists of several orthogonal building blocks, called perspectives. The five main building blocks are:

Functional perspective: Defines the steps of a process

Data and data flow perspective: Defines data used in the process and the flow of the data in between the single process steps. Relates those data to (external) data models.

Operational perspective: Specifies which service is invoked in order to execute a step of the process. Relates processes to services from (external) service libraries.

Organizational perspective: Defines user roles in the system. Relates user roles to (external) organizational models/charts.

Control flow perspective: Is used to define causal dependencies in between modeling elements (e.g. step "B" may only be executed after step "A").

Further perspectives can be added easily as iPM is based on a flexible meta model. This extensibility is important since some application domains require specific functions which are not provided a priori in general. Furthermore we are supporting the provision of libraries for perspectives which fosters the creation of structured models. For instance, in scientific application domains it is appropriate to offer a

library with pre-defined service calls that allow the access of the most important databases of this domain. The abstract process specification of iPM fosters rapid communication and exchange between researchers as technical details are still neglected. Through the integration of a (external) data model the processes become more consistent. Data is described on an abstract level and it has to be specified which transformation an item is undergoing at what time.

Also, iPM allows to specify workflows, derived from the processes modeled before. Workflow models are passed to iPE (integrated process execution, [Jablonski et al, 2006b]) that is able to execute them. To generate domain specific execution models, iPE is implemented according as Process Driven Architecture (PDA) [Müller, 2007]. The PDA heavily is based on templates which alleviates the generation of domain specific execution models. The usage of such a PDA is necessary in order to cope with domain specific workflow models.

### 3 POPM for modeling ecological processes

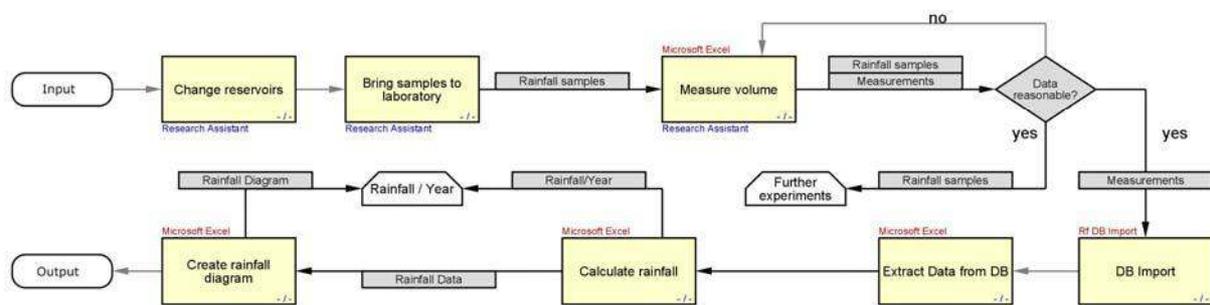


Figure 1. A process describing the calculation of the average rainfall per year

Figure 1 shows an example of a process describing the calculation of the average rainfall per year. It contains process steps which can be performed automatically, semi-automatically and manually. For example, the step "Change reservoir" can be only executed manually as no system exists yet that could change the probes on a test area (the same applies for "Bring samples to laboratory"). As a result a "Research Assistant" has to do this— here the value and necessity of the assignment of concrete persons to workflow steps becomes obvious. "Measure volume" could be performed semi-automatically depending on the laboratory of an institution. After the measurement of volume of the samples, these values must be somehow verified. In case they are not correct, the measurement has to be repeated. After a correct measurement, these samples are no longer required in this particular process so they are, from now, available for further research. The volume data of the samples is imported into the institution's database by a customized tool. This tool extracts the volume data from a Microsoft Excel sheet which has been created before. After the data being loaded into the database, it is imported again into a new Excel worksheet. This worksheet is then used to calculate the rainfall; finally the information is converted into a graphical representation and is available for further studies. In order not to forget: the structures of all documents mentioned in the above example are predefined in a common data model underlying the whole application domain. Modeling this process using iPM/iPE results in the following benefits.

- a) Interactive processes (processes which correspond with the user) can be modeled and also executed. The execution system also keeps track of these steps such that, later on, consistency checking becomes possible.
- b) The process is easy to read even for scientists which are not from the ecological application domain – so this process can be used as a communication medium.
- c) A separate data model exists which describes the data being passed in between the single process steps and does not center on the data being passed as parameters to methods during the execution of a process. In fact it is possible to convert the data items in between two steps depending on additional information provided such as ontologies.
- d) An organizational model can be specified to describe responsibilities. Users who do not have the proper rights cannot execute a process step.
- e) Single process steps can be reused in other processes: An example is the process step "DB Import"; its behaviour only depends on the algorithm used during the import.
- f) Changing the code which is executed in a workstep is easy as only the operational perspective has to be parameterized.

#### **4 Outlook and future work**

We are planning to support process libraries. A process library contains processes which have been developed earlier in order to increase the efficiency during modeling. Also advanced visualization techniques such as "where is my process currently" and "what data has been produced by step X out of what input data" still need to be integrated – the information is already stored in a database during execution but is not yet visible to the user.

In our presentation we also want to compare the two mentioned scientific workflow systems Taverna and Kepler with the iPM/iPE approach. Furthermore we will also present, in detail, the above mentioned example of process modeling in ecological research.

## 5 References

- [Altintas et al., 2004] Altintas, I. et al: Kepler: An Extensible System for Design and Execution of Scientific Workflows, system demonstration, 16th Intl. Conf. on Scientific and Statistical Database Management, 21-23 June 2004
- [Deelman et al., 2005] Deelman, E. et al: Pegasus: a Framework for Mapping Complex Scientific Workflows onto Distributed Systems, Scientific Programming Journal, Vol. 13(3), 2005
- [Hull et al., 2006] Hull, D. et al: MR: Taverna: a tool for building and running workflows of services, Nucleic Acids Research, Oxford University Press, 2006
- [Jablonski & Bussler, 1996] Jablonski, S; Bussler, C.: Workflow Management – Modeling Concepts, Architecture and Implementation. Intl. Thomson Computer Press, London, 1996
- [Jablonski et al, 2006a] Jablonski, S.; Faerber, M.; Schlundt, M.: Bridging the gap between SPICE Reference Processes and OU Processes: An iterative business process modeling approach. Proceedings of the 6th Int'l SPiCE Conference, Luxembourg, 2006
- [Jablonski et al, 2006b] Jablonski, S. et al: Configurable Execution Environments for Medical Processes, 4th Int'l Conf. on Business Process Management, Vienna, Austria, 2006
- [Meiler, 2005] Meiler, C.: Modellierung, Planung und Ausführung Klinischer Pfade. Eine integrierte modellbasierte Betrachtung evidenzbasierter, organisatorischer und betriebswirtschaftlicher Gesichtspunkte. ibidem, Stuttgart, Germany, 2005
- [Müller, 2007] Müller, S.: Modellbasierte IT-Unterstützung von wissensintensiven Prozessen, PhD Thesis, University of Erlangen-Nürnberg, Germany, 2007
- [ProDatO, 2005] ProDatO Integration Technology GmbH: Handbuch iPM – Integrated Process Manager. Documentation, Erlangen, Germany, 2005, <http://www.prodato.de>
- [Qin et al, 2006] Qin, J.; Fahringer, T.; Pillana, S.: UML Based Gird Workflow Modeling under ASKALON, 6th Austrian-Hungarian Workshop on Distributed and Parallel Systems, Innsbruck, Austria, September 21-23, 2006
- [Taylor et al, 2007] Taylor, I. et al: The Triana Workflow Environment: Architecture and Applications, Workflows for e-Science, Springer, New York, USA, 2007



# Gefahr in Verzug - Stoffinformationssysteme für die Bewertung von Chemikalien in Deutschland

Gerlinde Knetsch <sup>1</sup>

[gerlinde.knetsch@uba.de](mailto:gerlinde.knetsch@uba.de)

## Abstract

Chemikalien sind fester Bestandteil unseres Alltages. Die große Anzahl der auf dem Markt befindlichen Stoffe und Zubereitungen erfordern den Einsatz der Informationstechnologie, um schnell und zuverlässig Informationen zu den Eigenschaften und Anwendungsgebieten dieser Chemikalien zu finden. Damit verbunden ist ein hoher Anspruch an die Datenqualität, die Datensicherheit und das Datenmanagement.

Für verschiedene Aufgaben werden diese chemischen Stoff- Informationssysteme genutzt. Auskünfte zu gefährlichen Stoffen aufgrund eines Unfalls, Fragen zum Transport und der Lagerung von gefährlichen Chemikalien, die Einstufung einer von der Industrie in Verkehr gebrachten Chemikalie in eine Wassergefährdungsklasse oder die Anmeldung eines Pflanzenschutz- oder Arzneimittels zur Vermarktung benötigen valide Informationen. Dabei sind die Stoffe und Produkte auf ihre Risiken für die Ökosysteme zu untersuchen. Die Eigenschaften dieser chemischen Stoffe und ihr Verhalten in der Umwelt spielen dabei eine wesentliche Rolle. Verschiedene Stoffinformationssysteme des Umweltbundesamtes verwalten derartige Informationen und bieten für einen breiten Nutzerkreis derartige Daten und Informationen an.

## 1 Die Gefahrstoffschnellauskunft

Ein Tankerunglück in der Nordsee, bei einer Grenzkontrolle spürt der Zoll eine verdächtige LKW-Ladung Chemikalien auf oder die Feuerwehr ruft beim Löscheinsatz auf einem Industriegelände nach Informationen über vermutlich ausgetretene Stoffe. Meist müssen Helferinnen und Helfer in Notfällen sofort über Maßnahmen zum Schutz der Gesundheit oder der Umwelt entscheiden. Die notwendige Auskunft zu umweltgefährdenden Stoffen hält die Gefahrstoffschnellauskunft (GSA) des UBA bereit. Mit wenigen ‚Mausklicks‘ liefert die GSA ein genaues und fachlich gesichertes Bild der Gefahren etwa 30.000 vermarkteter Stoffe. Zugriff haben verschiedene Stellen des Bundes, der Länder sowie andere Institutionen mit öffentlich-rechtlichen Aufgaben. Ziele der Gefahrstoffschnellauskunft sind:

- schnelle und qualitativ hochwertige Auskunft zu gefährlichen Stoffen;
- Informationshilfe bei Störfällen und Unfällen mit umweltgefährdenden Stoffen;
- Überprüfen der Lagerung und des Transports umweltgefährdender Stoffe;

- Hinweise, wie sich Gefahren und Schäden durch umweltgefährdende Stoffe vermeiden lassen.

## 2 Chemische Stoffinformationen für die Öffentlichkeit

Interessierte Verbraucherinnen und Verbraucher können einen Teil der chemischen Stoffinformationen über Internetportale abrufen: Die zentrale Plattform <http://www.stoffdaten-deutschland.de/> vernetzt dazu verschiedene Anbieter.



Abb.1: Startseite der Stoff-Datenbanken der Bundesrepublik Deutschland

Dieses zentrale Portal bietet den Einstieg unter anderem in den Gemeinsamen Stoffdatenpool des Bundes und der Länder kurz GSBL an. Den Datenpool bauen Bund und Länder seit 1994 auf und pflegen deren Dateninhalte. Die aktuelle Version des GSBL (Stand: 2007) listet 40.000 Einzelinhaltsstoffe (Reinstoffe), 320.000 Komponentenstoffe (Gemische und Zubereitungen) und 190.000 Rechtsstoffklassen (rechtliche Regelungen).

Wer nutzt den Gemeinsamen Stoffdatenpool des Bundes und der Länder (GSBL) und für welche Zwecke ?

- Polizeibehörden, etwa die Wasserschutz- oder Autobahnpolizei, können mit dem GSBL das von einem Gefahrguttransport ausgehende Risiko schnell und sicher bewerten;
- Feuerwehreinsetzungskräften bietet der GSBL verlässliche Informationen bei Brand oder Freisetzung von Gefahrstoffen;
- Umweltbehörden auf allen Verwaltungsebenen nutzen die GSBL-Daten in der Gewerbeaufsicht, im Arbeitsschutz sowie beim Umgang mit und der Entsorgung gefährlicher Stoffe;

- Der politischen Ebene bietet der GSBL Informationen zur Bewertung bestehender und zur Erarbeitung neuer Rechtsvorschriften.

Der Gemeinsame zentrale Stoffdatenpool Bund/Länder ist ein Kooperationsprojekt zwischen dem BMU und den Länderministerien. Die Koordinierungsstelle liegt beim Umweltbundesamt. Unter Einsatz eines komplexen Softwaresystems werden Daten aus verschiedenen Quellen in- und ausländischer Partner verwaltet, gepflegt und für Recherchen aufbereitet. Unter [www.gsbl.de](http://www.gsbl.de) ist der öffentliche Zugang zu dem nicht lizenzbehafteten Datenbestand des GSBL möglich.

### **3 Informationssystem zu ökotoxikologischen Wirkungsdaten und Umweltqualitätskriterien**

Das Informationssystem WEB-ETOX dient als Grundlage zur Entwicklung von ökotoxikologisch begründeten Referenzwerten zum Schutz der aquatischen und terrestrischen Lebensgemeinschaften sowie zur Bewertung von Stoffen. Das Informationssystem ermöglicht, in einer übersichtlichen Form neben validierten Wirkungsdaten die vorhandenen Umweltqualitätskriterien, -ziele und -standards für Stoffe und Wirkparameter in den Medien Wasser, Boden, Sedimente, Biota und Nahrungsmittel zu dokumentieren. Die schnelle Verfügbarkeit dieser Daten ist eine wichtige Grundlage für die Bewertung von Umweltdaten im Rahmen der Umweltbeobachtung und bei der Weiterentwicklung von Qualitätskriterien.

WEB-ETOX ermöglicht den strukturierten Zugriff auf aufbereitete Wirkungsdaten zur aquatischen und terrestrischen Ökotoxikologie einschließlich von Details zu den verwendeten Untersuchungsmethoden. ETOX enthält auch Informationen zu nationalen und internationalen Qualitätszielen, Standards, Grenz-/ Richtwerte.

Die Webapplikation verwaltet ca. 19.000 Datensätze zur Wirkung auf aquatische Organismen, ca. 4000 Datensätze zur Wirkung auf terrestrische Organismen sowie ca. 3000 Datensätze zu Umweltqualitätskriterien für die Medien Wasser und Boden. Ein Ausbau des Datenbestandes erfolgt über projektbezogene (Stand: Dezember 2007).

Allgemeine Informationen zu chemischen Stoffen erhält das IS ETOX unter anderem durch eine Verknüpfung mit dem Gemeinsamen Stoffdatenpool Bund Länder (GSBL).



Abb. 2 WEB-ETOX - Datenbank für ökotoxikologische Wirkungsdaten und Qualitätsziele

Unter <http://webetox.uba.de/webETOX/index.do> ist der online-Zugriff für eine Recherche möglich. Eine link-Seite zu weiteren themenrelevanten Informationsanbietern gibt den Einstieg zu Projektpartnern auf nationaler und internationaler Ebene. Die US EPA bietet ein ähnliches System, aber auch nützliche Hintergrundinformationen zu den Canadian Environmental Quality Guidelines (CEQG) findet der Anwender.

Die Web-Applikation bietet ein online Erfassungsmodul zur projektgebundenen Dateneingabe. Somit können Daten für die Pflege und Fortschreibung des Datenbestandes interaktiv bereitgestellt werden. Der Zugang wird über ein Rechte- und Nutzerkonzept geregelt.

#### 4 Daten auf einen Blick: das Informationssystem Chemikaliensicherheit

Bevor neue chemische Stoffe auf den Markt kommen, bewertet das Umweltbundesamt die stofflichen Umweltrisiken. Hierzu legen die herstellenden Unternehmen Daten über die Wirkungen und das Verhalten ihrer Stoffe in der Umwelt vor. Auf dieser Basis prüft das UBA, ob es der Vermarktung, beispielsweise eines Pflanzenschutzmittels, Biozids oder Tierarzneimittels aus Umweltschutzsicht zustimmen kann. Was aber, falls alle drei genannten Produkte denselben Wirkstoff enthielten? Dies vereinfacht das Verfahren nicht, denn die drei Produkte werden nach unterschiedlichen Gesetzen (Pflanzenschutzgesetz, Biozidgebiet und Tierarzneimittelgesetz) geprüft, und die Bewertung des Umweltrisikos unterscheidet sich in diesen Gesetzen. Auch bleiben andere Umweltbelastungen des Wirkstoffs als die konkret beantragte Anwendung außer Acht.

Für die Bewertung der einzelnen Anmeldung lassen sich jedoch gesetzesübergreifende Informationen nutzen: Seit mehr als 15 Jahren verfolgt das UBA die Strategie, Informationen zu chemischen Stoffen in dem modular aufgebauten Informationssystem Chemikaliensicherheit (ICS) für die

gesetzesübergreifende Vollzugsarbeit aufzubereiten. Beantragt ein Unternehmen beispielsweise die Zulassung eines (sehr giftigen) synthetischen Pyrethroids als Biozidprodukt, so reicht den Fachleuten des Amtes ein Blick in die ICS-Datenbank, um das Ergebnis einer abgeschlossenen Pyrethroidbewertung mit dem Verwendungszweck als Pflanzenschutzmittel oder Tierarzneimittel zu finden. Diese Daten sind dann auch für die Bewertung des Biozidproduktes nutzbar. Für eine effektive und gründliche Stoffbewertung ist es sinnvoll, Daten und Bewertungsergebnisse verschiedener Produktarten auf einen Blick vollständig verfügbar zu machen. Im UBA hat sich das ICS mit der Zeit zu einem europaweit einzigartigen System entwickelt, das derzeit über 119.000 Stoffe (Stand: 03/2007) verwaltet. Dazu gehören:

- Altstoffe (davon 80 % mit Strukturformel)
- Neustoffe (Anmeldungen)
- Pflanzenschutzmittel (Präparate)
- Pflanzenschutzmittel-Wirkstoffe
- Pflanzenschutzmittel (EU-Prüfung)
- Biozid-Wirkstoffe
- Humanarzneimittel (Präparate)
- Tierarzneimittel (Präparate)
- Arzneimittelwirkstoffe
- Metabolite

Tabelle 1: Stoffe und Produkte des Informationssystems Chemikaliensicherheit  
Quelle: Umweltbundesamt

Dieses modular aufgebaute Informationssystem bietet einerseits die Aufnahme weiterer Vollzüge in das Informationssystem, andererseits unterstützt es die vollzugsübergreifende Recherche über einen chemischen Stoff. Dazu enthält das System einen gut strukturierten Identifikationsbereich zu den chemischen Stoffen. Dies umfasst u.a. die CAS-Nr. des registrierten Stoffes mit der Strukturformel, ein Mol-File sowie den Smiles-Code, ebenso Angaben zu internationalen Verzeichnissen wie das EINEC (European Inventory of Existing Commercial Chemical Substances) und EILINCs (European List of Notified Chemical Substances). Faktendaten geben Auskunft zu den chemischen Eigenschaften des Stoffes in der Umwelt, aber auch zum Abbauverhalten, der Bioakkumulation und möglicher ökotoxikologischer Effekte in der Umwelt.

Da dieses System in den Geschäftsprozess des Bewertungsverfahrens für die Zulassung eines chemischen Stoffes eingebunden ist, enthält das komplexe Vollzugsinstrument auch administrative Metadaten für die Registrierung eines Vorgangs wie Eingangsdatum, Aktenzeichen, Antragsteller etc.

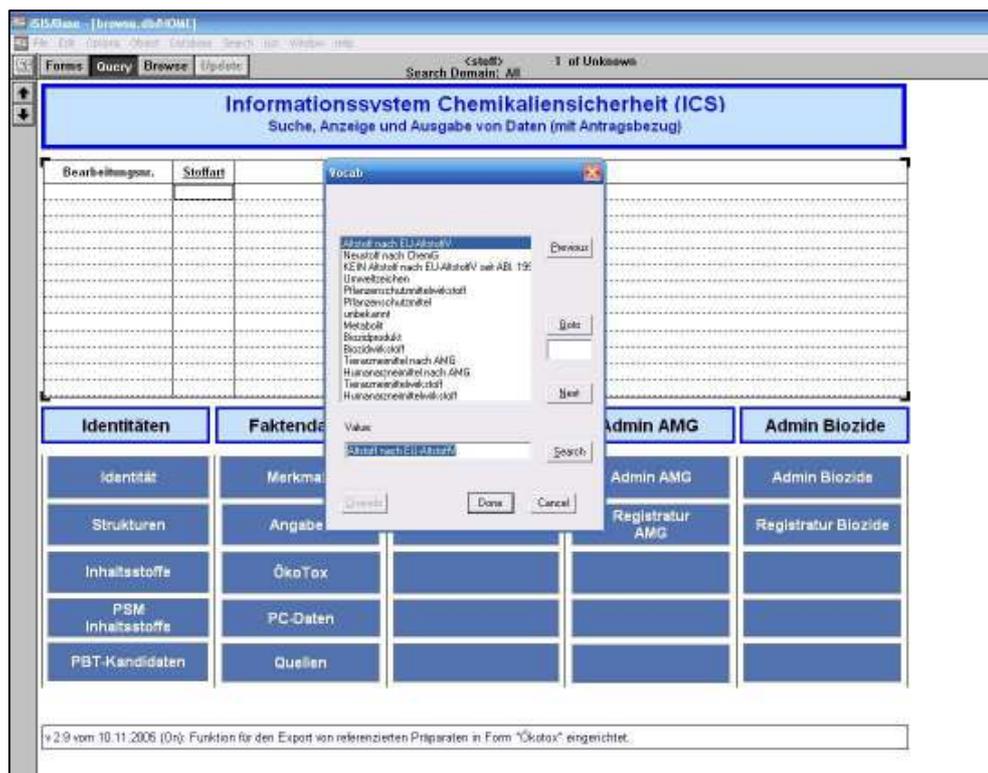


Abb. 3 Recherchemaske des ICS für die Anzeige und Ausgabe von Daten mit Antragsbezug

Darüber hinaus enthält das System Module für die Datenaktualisierung, Qualitätssicherung, Administrierung und Ausgabemöglichkeiten für standardisierte Templates für die Bewertungsberichte.

In einem 2007 begonnenen Projekt zur technischen und fachlichen Weiterentwicklung des Informationssystems Chemikaliensicherheit finden auch die Entwicklungen der Internationalen Chemikalienpolitik Berücksichtigung. Das Inkrafttreten von REACH am 1. Juli 2007 beeinflusst in seiner technischen sowie inhaltlichen Ausgestaltung die Prozesse des Datenmanagement von Stoffinformationssystemen auch auf nationaler Ebene erheblich. Mit der Ausgestaltung einer Service-orientierten Architektur des Informationssystems Chemikaliensicherheit soll diesen Anforderungen Rechnung getragen werden.

## 5 Ausblick

Europäische Richtlinien und Gesetzesvorschriften fordern zunehmend die Bereitstellung von chemischen Stoffinformationen für die Öffentlichkeit. Datenschutz und die gesicherte und schnelle Kommunikation spielen hierbei eine wesentliche Rolle.

Die Umsetzung von REACH – REACH steht für Registrierung, Evaluierung und Autorisierung von Chemikalien - ordnet das Chemikalienrecht in der Europäischen Union grundlegend neu. Die Flut von Daten, die in den kommenden Jahren die Europäische Chemikalienagentur ECHA in Helsinki erreichen wird, kann nur eine leistungsstarke und auf Prozessabläufe orientierte Informationstechnik in geordnete Bahnen lenken.

Gemeinsam entwickeln daher die zuständigen Behörden der EU-Mitgliedstaaten, Unternehmen der chemischen Industrie und die Europäische Kommission ein komplexes und vernetztes Informations- und Kommunikationssystem, das auf zwei Säulen ruht. REACH-IT ist das Webportal der Europäischen Chemikalienagentur zur Steuerung des Anmeldeverfahrens. Dort beantragen mehrere zehntausend Unternehmen die Registrierung ihrer Stoffe und binden die dafür notwendigen Daten in das Workflow-System der Chemikalienagentur ein.

IUCLID 5 (International Uniform Chemicals Information Database) heißt die zentrale Fachdatenbank für die Verwaltung der Stoffdaten, die die Unterlagenerstellung unterstützt. Eine Gefahrstoffdatenbank, die wir in Deutschland bereits mit dem GSBL kennen, wird es dann auch für Europa geben: Über das Portal REACH-IT erhält die Öffentlichkeit Informationen zur Gefährlichkeit der Chemikalien für den Menschen und die Umwelt. Alle Prozessschritte - von der Anmeldung bis zur Registrierung und der Stoffauskunft - erfolgen papierlos. Die EU-Chemikalienagentur ist damit Vorbild für andere Gesetzesvollzüge.

## 6 Literaturverzeichnis

[UBA, 2007] Umweltbundesamt: Jahresbericht des Umweltbundesamtes 2006

[UBA, 2007] Umweltbundesamt: Chemikalienpolitik und Schadstoffe, REACH - Gemeinsamer zentraler Stoffdatenpool Bund/Länder (GSBL)

Link: <http://www.umweltbundesamt.de/chemikalien/gefahrstoffe/gsbl.htm>

[UBA, 2004] Schudoma, D.: Fachliches Handbuch zur Recherche und Datenerfassung für die Datenbank ökotoxikologische Wirkungsdaten und Qualitätsziele, Stand: 12/2004

[ECHA 2007] Europäische Kommission, OECD, ECHA: IUCLID5 - International Uniform Chemical Information Database and REACH-IT, ECHA 2007



# PortalU®-Komponenten und -Schnittstellen zur Datenrecherche und -erfassung

Martin Klenke, Fred Kruse, Thomas Vögele

[kst@portal.de](mailto:kst@portal.de)

## Einleitung

Am 29.05.2006 wurde PortalU® als neues Umweltportal des Bundes und der Länder freigeschaltet. Seitdem sind unter <http://www.portal.de> sowohl Webseiten als auch Datenbankinhalte und Metadatenkataloge der beteiligten öffentlichen Institutionen und Organisationen recherchierbar. Die Zahl der Informationsanbieter steigt kontinuierlich: Derzeit stellen mehr als 150 Partner (Stand Mai 2007) Angebote für PortalU® zur Verfügung.

In unserem Beitrag zum Workshop Umweltdatenbanken geben wir, vor dem Hintergrund der Themenkreise "Internationale und globale Netzwerke" sowie "Methodische Umsetzung", einen Überblick über das InGrid®-Architekturkonzept sowie über die existierenden und geplanten InGrid®-Schnittstellen und deren Arbeitsweise.

Den zweiten Themenschwerpunkt bildet die derzeit in Entwicklung befindliche, web-basierte Metadatenerfassungskomponente („InGrid® 1.1“) von PortalU®.

## 1 Anforderungen an die Software

Die Vielzahl der potentiellen Betreiber von PortalU®-Komponenten in Bund und Ländern legte, um Lizenz- und andere Folgekosten zu minimieren, bei der Anwendungsentwicklung die Nutzung von Open-Source-Software nahe. Die hinter PortalU® arbeitende Software „InGrid®“ (Kurzform für „Information-Grid“) wurde daher komplett auf der Basis von Open Source Komponenten realisiert. Tabelle 1 fasst die wesentlichen Bausteine zusammen.

Suchmaschine, Indexierung	Nutch	<a href="http://lucene.apache.org/nutch">http://lucene.apache.org/nutch</a>
	Apache Lucene	<a href="http://lucene.apache.org">http://lucene.apache.org</a>
Portal	Jetspeed	<a href="http://portals.apache.org/jetspeed-2">http://portals.apache.org/jetspeed-2</a>
WMS	UMN Mapserver	<a href="http://mapserver.gis.umn.edu">http://mapserver.gis.umn.edu</a>
WMS Viewer	Mapbender	<a href="http://www.mapbender.org">http://www.mapbender.org</a>
Datenbank	MySQL	<a href="http://www.mysql.com">http://www.mysql.com</a>
Kommunikation	JXTA	<a href="http://www.jxta.org">http://www.jxta.org</a>

Tabelle 1: PortalU®/InGrid®-Softwarekomponenten

## 1.1 Anforderungen aus Sicht der Datenanbieter

Für die Anbieter war u. a. der Wunsch nach Aktualität der zur Verfügung gestellten Daten sowie nach Singularität der Datenbestände zu berücksichtigen. Ein Data-Warehouse-Ansatz widerspricht diesen Anforderungen durch redundante Datenhaltung und wirft in der föderalen Struktur, in der die Datenbestände der deutschen Umweltverwaltung eingebettet sind, sowohl rechtliche als auch organisatorische und technische Probleme auf. InGrid<sup>®</sup> wurde daher modular aufgebaut und folgt einem dezentralisierten Konzept der Abfrage und Zusammenführung von Ergebnissen verteilter Datenquellen.

Besonderes Gewicht wurde bei der Entwicklung demgemäß auf ein flexibles Schnittstellenkonzept gelegt. Dieses hat, neben der erforderlichen Eigenkommunikation der InGrid<sup>®</sup>-Komponenten, u. a. sicherzustellen, dass PortalU<sup>®</sup>-Datenquellen kurzfristig und unaufwendig in Datennetze einzubinden sind, die auf Grundlage von aktuellen nationalen und internationalen Richtlinien realisiert werden (EU-Umweltinformationsrichtlinie, INSPIRE-Richtlinie, etc.). Die Berücksichtigung aktueller ISO- und OGC-Standards und -Empfehlungen ist obligatorisch.

## 1.2 Anforderungen aus Nutzersicht

Aus Nutzersicht wurden, neben der Qualität der Ergebnisse im Hinblick auf die durchgeführte Suche, zwei Kernanforderungen formuliert. Beide resultieren mittelbar aus der Arbeitsweise und Performanz aktueller Suchmaschinen, die durch ihre Verbreitung und ihren hohen Nutzungsgrad einen Quasi-Standard für internet-basierte Suchen setzen: Um vom Nutzer akzeptiert zu werden muss eine Suche (i) zeitnah ein erstes aggregiertes Ergebnis in Form einer Ergebnisliste erzeugen und (ii) muss die Ergebnisliste, bedingt durch die potentiell große Treffermenge, im Hinblick auf die Suchanfrage hin optimiert sein („Ranking“), d.h. sie muss für die Treffer eine Relevanz berechnen und dieser folgend die Ergebnisliste aufbauen.

## 2 Architektur

Ein zentrales Problem von verteilten Dateninfrastrukturen sind mehrstufige Kaskaden von An-/Abfragesystemen, die Dubletten erzeugen können und häufig wenig performant arbeiten. Letzteres fällt besonders stark ins Gewicht, wenn die Kommunikation zusätzlich über wenig performante Protokolle, wie z. B. XML/SOAP, erfolgt. Dies kann zu Antwortzeiten im Minutenbereich führen, die Nutzer in Zeiten von Google und anderen performanten, index-basierten Suchmaschinen kaum akzeptieren.

Um den in Abschnitt 1 formulierten Anforderungen gerecht zu werden, wurde InGrid<sup>®</sup> unter Nutzung von P2P-Technologie modular aufgebaut. Die Komponenten sind flexibel in unterschiedlichen Szenarien einzusetzen: Beispielsweise, wie bei PortalU<sup>®</sup>, als Portal-Komplettlösung zur Bereitstellung unterschiedlichster Datenquellen oder im Minimalfall als reine Metadatenrecherche-Anwendung, z.B. als Ersatz für den WWW-UDK. Das InGrid<sup>®</sup>-Architekturkonzept basiert maßgeblich auf der Abfrage verteilter Indizes und erlaubt den Verzicht auf mehrstufige Datenanbieter-Kaskaden,

da jede Datenquelle direkt an beliebig viele InGrid-Abfrage-Knoten angeschlossen werden kann.

## 2.1 System-Komponenten

Die schraffiert hinterlegten Komponenten in Abbildung 1 geben einen Überblick über die PortalU®/InGrid®-Softwarearchitektur. Als zentraler Broker fungiert der iBus („Informations-Bus“), an den über „iPlugs“ externe Datenquellen, z. B. Fachinformationssysteme, Datenbanken und Suchmaschinenindizes angeschlossen werden können. In der Abhandlung bezeichnen wir die Kommunikationsmodule „nach oben“ (Portal, ECS, CSW, OpenSearch) als Schnittstellen, die „unter“ dem iBus positionierten Module zum Anschluss von Datenquellen als iPlugs.

Der iBus vermittelt zwischen den über iPlugs angeschlossenen Datenquellen und den Systemen, die auf diese via Schnittstellen zugreifen möchten: Intern das Portal, über das ein Nutzer Suchanfragen stellen und, entsprechende Berechtigungen vorausgesetzt, administrative Aufgaben wahrnehmen kann, sowie externe Systeme, die PortalU® abfragen: In der Abbildung z. B. GeoMIS.Bund<sup>8</sup> über die CSW-Schnittstelle („Catalog Service Web“), die Metasuchmaschine MetaGer<sup>9</sup> über die OpenSearch-Schnittstelle und ein externer iBus über die ECS-Schnittstelle („Environmental Catalog Service“). Die verschiedenen iPlugs und Schnittstellen werden in Abschnitt 3 eingehender beschrieben.

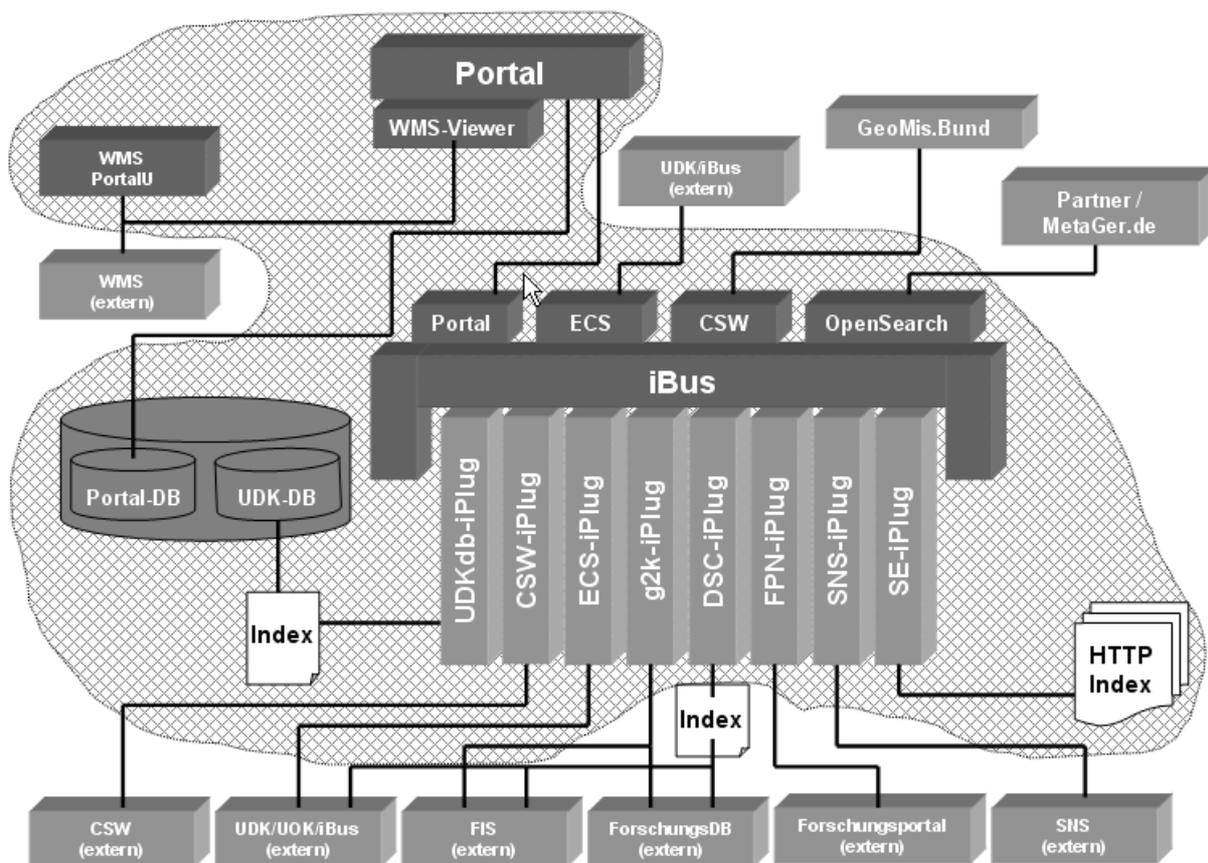


Abbildung 1: PortalU®/InGrid®-Softwarearchitektur.

<sup>8</sup> <http://www.geoportal.bund.de/> - 04.05.2007

<sup>9</sup> <http://www.metager.de/> - 04.05.2007

## 2.2 Kommunikation

Als Kommunikationsplattform dient das jxta-P2P-Framework (vgl. Tab. 1). Es ermöglicht die physikalische Verteilung der InGrid<sup>®</sup>-Komponenten und damit den Aufbau einer verteilten Infrastruktur. Zum Anschluss einer Datenquelle kann der verbindende iPlug sowohl lokal am iBus als auch extern an der Datenquelle eingerichtet werden. Die Kommunikation kann firewall-freundlich über http oder, wenn möglich, über das performantere tcp-Protokoll erfolgen. Die P2P-Technologie automatisiert dabei über einen Rendezvous-Server die Kontaktaufnahme sowie die Kommunikation zwischen iBus und iPlugs weitgehend.

## 2.3 An- und Abfragekonzepte

Um der Nutzeranforderung einer "schnellen Antwort" Rechnung zu tragen und die Treffermenge bezogen auf die Anfrage nach Relevanz sortieren zu können, wurde der Ansatz verfolgt, so viele Datenquellen wie möglich zu indexieren. Treffer indexierter Quellen werden zunächst ohne Rückgriff auf die eigentlichen Ressourcen aus dem Index generiert und zeitnah in einer gemischten, nach Relevanz sortierten Trefferliste zur Verfügung gestellt. Erst bei Anforderung der Detailansicht eines Treffers wird das datenhaltende System kontaktiert. Die Indexierungs-Technologie basiert auf dem Apache Lucene Framework (vgl. Tab. 1). Beispiele für indexierte Datenquellen in PortalU<sup>®</sup> sind die Webseiten und im Internet stehende, d. h. über einen Crawler erreichbare, Dokumente der Partner (pdf, doc, etc.), deren Umweltdatenkataloge sowie Datenbanken, z. B. die Literaturdatenbank für Naturschutz und Landschaftspflege DNL-Online<sup>10</sup> des Bundesamtes für Naturschutz (BfN), das Informationssystem Pflanzen und Vegetation FloraWeb<sup>11</sup> (BfN) und die Umweltforschungsdatenbank Ufordat<sup>12</sup> des Umweltbundesamtes (UBA) in Dessau.

Der Tatsache, dass eine Indexierung nicht für alle angeschlossenen Systeme infrage kommt bzw. gewünscht ist, wurde im Portal durch das Anlegen einer zweiten „Neben-Trefferliste“ Rechnung getragen (vgl. Abb. 2). Diese baut sich, je nach genutzter Schnittstelle und system-spezifischer Antwortzeit, gegenüber der Haupttrefferliste verzögert auf und ist nach Datenquellen gruppiert. Eine Mischung der Treffer unterschiedlicher Systeme bzw. ein Ranking der Ergebnisse ist an dieser Stelle systembedingt nicht möglich. Aktuell finden sich hier z.B. Angebote wie die Umwelt-Literaturdatenbank Ulidat<sup>5</sup>, die Umweltprobenbank UPB<sup>13</sup> und der Gemeinsame Stoffdatenpool des Bundes und der Länder GSBL<sup>14</sup>, die über die, vom Vorgängersystem GEIN („German Environmental Information Network“) [Bilo, 2000] übernommene, g2k-Schnittstelle angeschlossen sind. Zukünftig werden an dieser Stelle auch Ergebnisse von Metadatenkatalogen eingebunden, die über eine CSW 2.0-Schnittstelle verfügen.

Abbildung 2 zeigt eine PortalU<sup>®</sup>-Ergebnisliste mit nach Relevanz sortierten Treffern (links, unter „1-10 von 4968 bewerteten Treffern“) und der Liste der nicht index-basiert angeschlossenen Datenquellen unter „4 weitere Datenquellen“. Die Umweltforschungsdatenbank „Ufordat“ erscheint sowohl in der Haupt- als auch in der

<sup>10</sup> <http://www.dnl-online.de/> - 04.05.2007

<sup>11</sup> <http://www.floraweb.de/> - 04.05.2007

<sup>12</sup> <http://doku.uba.de/> - 04.05.2007

<sup>13</sup> <http://193.174.169.38/wwwupb/servlet/upb> - 04.05.2007

<sup>14</sup> <http://www.gsbl.de/> - 04.05.2007

Nebenergebnisliste, da sie über Index und direkt (g2k, siehe Abschnitt 3) an PortalU<sup>®</sup> angeschlossen ist.

Die Mischung der Datenquellen in der Hauptergebnisliste ist in Abbildung 3 erkennbar. Dem Nutzer wird sowohl grafisch, durch ein Symbol vor jedem Ergebniseintrag, als auch textuell („Quelle:“) die Datenherkunft transportiert.

The screenshot shows the PortalU search interface. At the top, there is a navigation bar with the PortalU logo and the text 'Umweltportal Deutschland'. Below this is a search bar containing the text 'Ozon' and a 'PortalU Suche' button. To the right of the search bar is a box titled 'INFO ZUR SUCHE' containing text about the search process. Below the search bar is a section for 'Ähnliche Begriffe: Suchen Sie nach ...'. The main search results are displayed in a list, starting with '1 - 10 von 4968 bewerteten Treffern'. The first result is 'Ozon-Bulletins' from 'Ozon-Bulletins', with a source link to 'http://www.dwd.de/...'. The second result is 'LUBW: Aktuelle Immissionsdaten' from 'LUBW', with a source link to 'http://www.umeg.de/...'. On the right side, there are two additional result boxes: '4 weitere Datenquellen' and 'UFOODAT (1320 TREFFER)' from 'Umweltforschungsdatenbank des Umweltbundesamtes Dessau'. Below that is 'UBA: Umweltforschungs-Datenbank (UFOODAT)' and 'UMWELTOBJEKTKATALOG BAYERN (67 TREFFER)' from 'Informationssystem für Planung und Umwelt'. At the bottom right, there is a link for 'Trefferliste vom Portal des UOK Bayern'.

Abbildung 2: PortalU<sup>®</sup>-Trefferlisten.

### 3 Schnittstellen und iPlugs

#### 3.1 Data Source Client (DSC)

PortalU<sup>®</sup> verfügt mit dem "Data Source Client"-iPlug über ein generisches Modul das es ermöglicht, Datenbanken und Fachinformationssysteme komfortabel und ohne Eigenentwicklung auf Anbieterseite an das Portal anzubinden. Datenbankinhalte können dabei sowohl in der PortalU<sup>®</sup>-Detailansicht dargestellt werden, als auch aus der PortalU<sup>®</sup>-Ergebnisliste heraus direkt per URL in die Ergebnisansicht des Fremdsystems verlinken (soweit vorhanden). Somit können per DSC auch Teile des sogenannten „Hidden Web“ einfach und kostengünstig für PortalU<sup>®</sup>/InGrid<sup>®</sup> erschlossen werden.

## **Rechtsprechung zu Nutzungsbeschränkungen in Wasserschutzgebieten**

Anbieter: Bundesamt für Naturschutz

Quelle: DNL-online

<http://www.dnl-online.de/1905.html?id=26588>

---

## **Fachgebiet Wasser**

... Hydrogeologie Messwerte Niederschlag Retentionskataster Wasserrahmenrichtlinie Wasserschutzgebiete Wasserstände/Pegel Gewässerkundlicher Dienst GEOLOGIE ...

Anbieter: Hessisches Landesamt für Umwelt und Geologie

Quelle: PortalU-Internet-Suchmaschine Suchkomponente

<http://www.hlug.de/medien/wasser/index.htm>

---

## **Wasserschutzgebiete (Sachsen)**

Sachdaten der Trinkwasser- und Heilquellenschutzgebiete

Anbieter: Sächsisches Landesamt für Umwelt und Geologie

Quelle: Sächsischer Umweltdatenkatalog

---

## **Wasserschutzgebiete (Sachsen)**

Festgesetzte Wasserschutzgebiete (Schutzzone I, II, II A, II B, III, III A, III B) - Trinkwasserschutzgebiete für Grundwasserfassungen, Fließgewässer und Talsperren - Heilquellenschutzgebiete

Anbieter: Sächsisches Landesamt für Umwelt und Geologie

Quelle: Sächsischer Umweltdatenkatalog

---

## **Verordnungen zur Festsetzung von Wasserschutzgebieten**

Hier finden sich Verordnungstexte, Karten und sonstige Informationen zu Wasserschutzgebieten, die der Niedersächsische Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWK) seit 2005 festgesetzt hat. Wasserschutzgebiete sind für die öffentliche Trinkwasserversorgung von größter Bedeutung. In Wasserschutzgebieten sind bestimmte Handlungen und Nutzungen verboten oder nur...

Anbieter: Niedersächsisches Umweltministerium

Quelle: Umweltdatenkatalog Niedersachsen

---

## **[PDF] Wasserschutzgebiet Billstedt Freie und Hansestadt Hamburg Umweltbehörde Sie wohnen oder arbeiten...**

... Wasserschutzgebiet in Billstedt festgesetzt? Wasserschutzgebiete werden dort ausgewiesen, wo das ...

Anbieter: Behörde für Stadtentwicklung und Umwelt Hamburg

Quelle: PortalU-Internet-Suchmaschine Suchkomponente

[http://fhh.hamburg.de/.../faltblatt\\_\\_wsg\\_\\_billstedt-pdf,property=source.pdf](http://fhh.hamburg.de/.../faltblatt__wsg__billstedt-pdf,property=source.pdf)

---

## **[DOC] Gliederung der Erläuterung für Wasserversorgungsanlagen 1 Vorhabensträger Name und Sitz...**

Gliederung der Erläuterung für Wasserversorgungsanlagen 1 Vorhabensträger Name und Sitz Tag der Antragstellung, Antrag Satzung, Verträge, Zweckvereinbarungen 2 Zweck ...

Anbieter: Bayerisches Staatsministerium für Umwelt, Gesundheit und Verbraucherschutz

Quelle: PortalU-Internet-Suchmaschine Suchkomponente

<http://www.bayern.de/lfw/service/rewas/a07.doc>

Abbildung 3: „Gemischte“ Trefferliste.

Der DSC erstellt zudem einen Index der gemappten Datenbankinhalte, so dass diese in der PortalU<sup>®</sup>-Trefferliste suchanfragen-spezifisch und datenquellen-übergreifend nach Relevanz sortiert werden können. Zum Suchanfragezeitpunkt wird, um die Ergebnisliste aufzubauen, lediglich auf den Index zugegriffen, was kurze Antwortzeiten ermöglicht. Erst bei der Anforderung eines Detailergebnisses wird auf die Quelldatenbank durchgegriffen.

## 3.2 XML-basierte Schnittstellen und iPlugs

Neben dem DSC verfügt InGrid<sup>®</sup> über verschiedene XML-basierte Schnittstellen und iPlugs, die im Folgenden kurz beschrieben werden.

### 3.2.1 Catalog Service Web (CSW)

Über die CSW 2.0-Schnittstelle wird PortalU<sup>®</sup> die Inhalte der angeschlossenen Umweltdatenkataloge für GeoPortal.Bund<sup>15</sup> (und mittelfristig möglicherweise INSPIRE<sup>16</sup>) verfügbar machen. Andererseits sollen über das CSW-iPlug zukünftig externe Geodaten-Kataloge an PortalU<sup>®</sup> angebunden werden.

### 3.2.2 Environmental Catalog Service (ECS)

Das ECS-iPlug bzw. die ECS-Schnittstelle ist eine Weiterentwicklung der UDK-SOAP-Schnittstelle und bildet das gesamte fachliche UDK-Datenmodell ab. Sie soll sich zukünftig zu einer Standardschnittstelle für den Austausch von Umwelt-Metadaten entwickeln. Um dies zu unterstützen wird derzeit versucht, die Schnittstelle und das UDK-Datenmodell im Rahmen des Deutschland-Online Vorhabens „Standardisierung“ als XÖV-Projekt<sup>17</sup> zu verankern. XÖV steht für "XML-basierte fachliche Standards für den elektronischen Datenaustausch innerhalb und mit der öffentlichen Verwaltung".

### 3.2.3 Semantic Network Service (SNS)

Das SNS-iPlug dient zur Integration von Funktionen des Semantic-Network-Service<sup>18</sup> (SNS) des UBA [Angrick, 2002]. Der SNS wird mit seinen thesaurus-basierten Dienstleistungen u. a. zur automatischen Verschlagwortung von Webseiten (fachlich und räumlich) im Rahmen der Indexierung genutzt.

### 3.2.4 g2k

Das g2k-iPlug wurde, wie weiter oben bereits angemerkt, vom Vorgängersystem GEIN übernommen und bislang weiter unterstützt. Es ist geplant, die g2k-Funktionalitäten mittelfristig in das ECS-Format zu integrieren.

### 3.2.5 forschungsportal.net

Das FPN-iPlug dient zur Abfrage und Aufbereitung der Treffer des Angebots „forschungsportal.net“<sup>19</sup> vom Bundesministerium für Bildung und Forschung, dass unter dem Suchraum „Forschungsprojekte“ in PortalU<sup>®</sup> integriert ist.

### 3.2.6 OpenSearch

Die OpenSearch-Schnittstelle<sup>20</sup> ermöglicht die Aufbereitung von InGrid<sup>®</sup>-Suchergebnissen für andere Suchmaschinen und Clienten. Das OpenSearch-Format

---

<sup>15</sup> [http://geoportal.bkg.bund.de/DE/Home/homepage\\_node.html](http://geoportal.bkg.bund.de/DE/Home/homepage_node.html) - 04.05.2007

<sup>16</sup> <http://www.ec-gis.org/inspire/home.html> - 04.05.2007

<sup>17</sup> [http://www.deutschland-online.de/DOL\\_Internet/broker.jsp?uMen=fae10482-4b88-e011-4fbf-1b1ac0c2f214](http://www.deutschland-online.de/DOL_Internet/broker.jsp?uMen=fae10482-4b88-e011-4fbf-1b1ac0c2f214) - 04.05.2007

<sup>18</sup> <http://www.semantic-network.de> - 04.05.2007

<sup>19</sup> <http://www.forschungsportal.net/> - 04.05.2007

wurde ursprünglich von A9/Amazon.com entwickelt und wird mittlerweile von vielen Suchmaschinen und Suchanwendungen unterstützt bzw. genutzt.

## **4 Metadatenerfassung**

### **4.1 Umweltdatenkatalog**

In Deutschland spielt der Umweltdatenkatalog (UDK) eine zentrale Rolle bei der Referenzierung und Beschreibung von Daten, Karten und Dokumenten im Umweltbereich [Swoboda, 2000]. Seit mehr als 10 Jahren ist der UDK bei den Umweltbehörden des Bundes und der Länder im Einsatz und hat sich in diesem Umfeld zu einem Quasi-Standard zur Beschreibung von Umweltinformationen entwickelt. Das Datenmodell des UDK berücksichtigt die internationalen Standards ISO 19115 und ISO 19119 [Karschnick, 2003].

Als technisches System wurde der UDK ursprünglich mit einer Online-Komponente für Recherche und Visualisierung (WWW-UDK), sowie einer Desktop-Applikation (Windows-UDK) für die Eingabe von Metadaten und für das Katalogmanagement betrieben. Seit Mai 2006 ist die Online-Recherche-Komponente des UDKs in PortalU<sup>®</sup> integriert. Seither können unter dem gemeinsamen Dach von PortalU<sup>®</sup> nicht nur Webseiten und Fachinformationssysteme von Behörden, sondern auch die UDKs des Bundes und der Länder durchsucht werden.

### **4.2 InGrid<sup>®</sup> 1.1: Web-basierte Erfassung von UDK-Metadaten**

Um die Dateneingabe und den kompletten Workflow, von der Metadatenerfassung und Pflege über die Qualitätssicherung (QS) bis hin zur Datenrecherche, über das Internet abwickeln zu können, wird derzeit eine neue Softwarekomponente entwickelt: InGrid<sup>®</sup> 1.1. Diese wird ab 2008 die bisherige Desktop-Erfassungssaplikation ersetzen. Aktuell (Mai 2007) sind die technischen Feinkonzepte für InGrid<sup>®</sup> 1.1 erstellt und für erste Masken und Funktionsabläufe existieren Prototypen. Die Implementierungsphase steht unmittelbar bevor.

Die Entscheidung, zukünftig ein internet-basiertes Tool für Erfassung, Pflege und QS zu verwenden, basierte u. a. auf der langjährigen Erfahrung mit einem UDK-Katalogmanagement, in dessen Rahmen verteilt gepflegte Kataloge mit Hilfe eines definierten Aktualisierungszyklus' in einen zentralen Katalog überführt wurden und dieser dann wieder in die nachgeordneten Installationen zurück verteilt wurde. Dieses Vorgehen hat sich in der Praxis sowohl technisch als auch organisatorisch als wenig robust erwiesen. Das neue Werkzeug wird daher mit Ajax-Technologie<sup>21</sup> als benutzerfreundliche und browserbasierte Internetapplikation umgesetzt, die unabhängig vom Standort des Metadatenautors das Schreiben in einen gemeinsamen Katalog ermöglicht. Bei der Entscheidung für Ajax, das aktiviertes Javascript im Browser bedingt, setzten sich Gesichtspunkte der Ergonomie gegen die der Barrierefreiheit (im Sinne der BITV) durch: Die Metadatenautoren erwarten nach langjähriger Erfahrung mit einer komfortablen, performanten Desktopanwendung ein ähnlich flüssiges Arbeiten auch unter einer Webapplikation. Da die Erfassung/Qualitätssicherung lediglich einen autorisierten Nutzerkreis betrifft, wurden einige BITV-Belange zugunsten der Usability zurück gestellt.

---

<sup>20</sup> <http://www.opensearch.org> - 04.05.2007

<sup>21</sup> [http://de.wikipedia.org/wiki/Ajax\\_\(Programmierung\)](http://de.wikipedia.org/wiki/Ajax_(Programmierung)) - 04.05.2007

Damit entfallen zukünftig bei den Partnern sowohl das aufwendige und teilweise fehleranfällige UDK-Katalogmanagement als auch die Notwendigkeit einer Eigeninstallation von UDK-Komponenten: Kataloge können zukünftig bei Bedarf auch auf der MySQL-Datenbank der PortalU<sup>®</sup>-Zentralinstallation betrieben und über das Internet gepflegt werden.

Erforderliche Anpassungen am aktuellen UDK-Datenmodell beschränken sich auf ein technisch erforderliches und sinnvolles Minimum. Die Möglichkeit der Überführung eines UDK 5.0-Kataloges in das InGrid<sup>®</sup> 1.1-Metadatenmodell wird über eine Altdatenübernahme unaufwendig möglich sein.

### **4.3 Vorschau: Die InGrid<sup>®</sup> 1.1 Metadaten-Erfassungskomponente**

Im Folgenden werden einige Aspekte der neuen Erfassungsanwendung vorgestellt. Die Bildschirmabzüge basieren auf dem aktuellen (04.05.2007) Oberflächenprototyp der Anwendung.

#### **4.3.1 Konfigurierbare Benutzeroberfläche**

Das UDK-Datenmodell berücksichtigt, wie oben angesprochen, die ISO 19115 und 19119 und geht in Teilen darüber hinaus. Insgesamt existieren mehr als 450 Eingabefelder. In der Praxis ist allerdings meist nur ein geringer Teil der vorhandenen Metadatenelemente erforderlich, um ein Datenobjekt im Sinne der Standards ausreichend zu beschreiben. Die Mindestanforderungen des ISO-Standards, der ISO-Core, umfassen daher weniger als 20 Metadatenelemente.

Um die Usability zu unterstützen und unnötige Komplexität zu vermeiden besteht in InGrid<sup>®</sup> 1.1 grundsätzlich die Möglichkeit der Wahl zwischen einer vollständigen und einer auf die Core-Elemente reduzierten Darstellung. Zusätzlich kann der Katalogadministrator festlegen, welche Elemente über die Eingabeoberfläche per Voreinstellung sichtbar sind. Das "Ausfalten" der Ansichten ist bei Bedarf für einzelne Elementbereiche oder global möglich. Damit bietet sich die Möglichkeit, die Metadatenerfassungsfläche katalogweise an die Bedürfnisse spezifischer Landes-Metadatenprofile bzw. Nutzergruppen anzupassen.

Abbildung 4 zeigt einen Ausschnitt aus der Erfassungsmaske für Metadaten. Der Abschnitt „Raumbezug“ bietet hier zunächst nur die Auswahl einer administrativen Einheit an, über die der Raumbezug in Form einer Bounding-Box erstellt wird. Alternativ können Bounding-Box-Koordinaten in die Tabelle eingegeben werden. Die administrative Einheit kann aus einer Liste oder anhand einer Karte, die über den PortalU<sup>®</sup>-Web-Map-Service bereitgestellt wird, ausgewählt werden. Per Voreinstellung wird in der Liste bei der Neuerfassung eines Objekts das Land ausgewählt, auf das der Katalog referenziert.

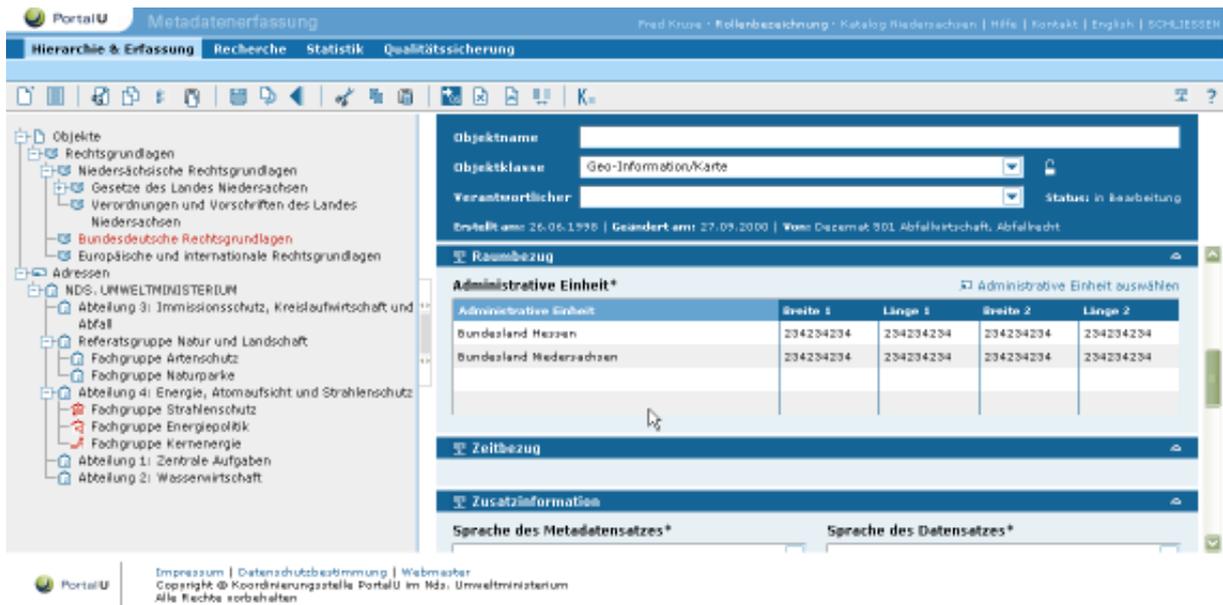


Abbildung 4: InGrid® 1.1 Metadatenerfassungskomponente.

Abbildung 5 zeigt den „ausgefalteten“ Bereich für den Raumbezug, der alle vorhandenen UDK-Raumbezugs-Elemente enthält.

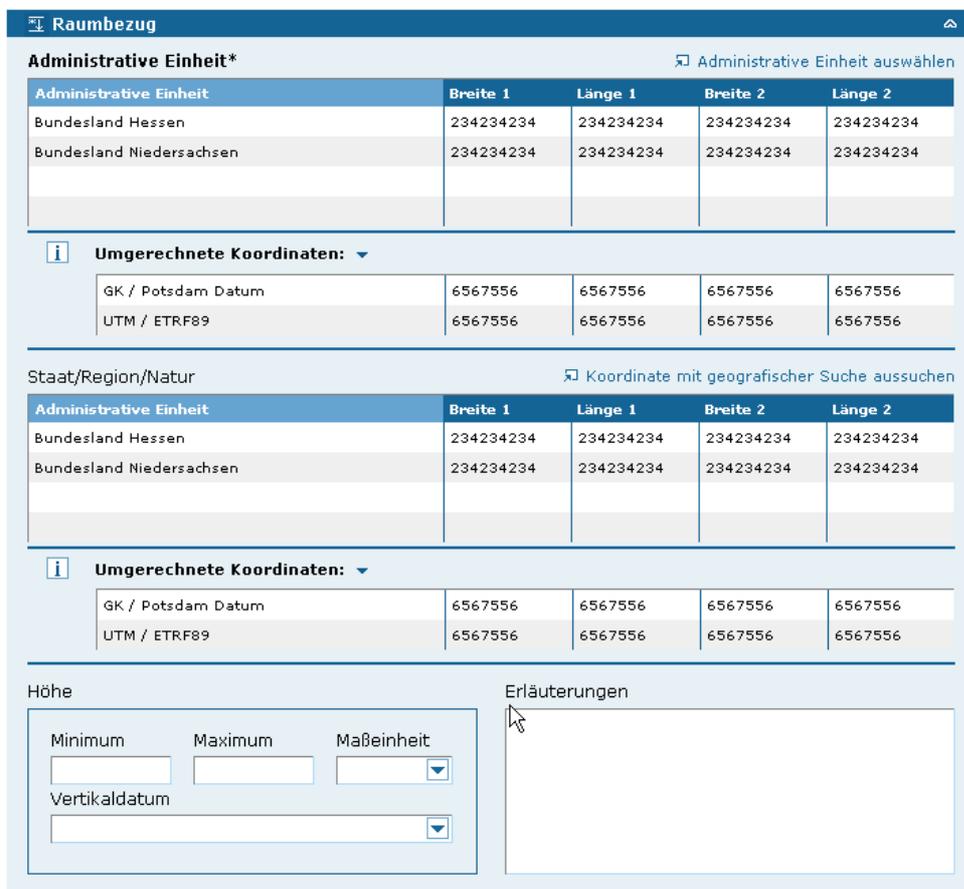


Abbildung 5: Raumbezugsinformation eines UDK-Objekts.

### 4.3.2 Erfassungsassistenten

Konsistente Fach-, Raum- und Zeitbezüge sind für die inhaltliche Beschreibung und Kategorisierung von Datensätzen erforderlich und für den Erfolg einer Datenrecherche mitentscheidend, insbesondere wenn über vernetzte Kataloginfrastrukturen gesucht wird. Um den Nutzer bei der Erfassung zu unterstützen wird ein Assistent eingesetzt, der über den SNS (vgl. Abschnitt 3) beispielsweise Internetseiten analysiert und auf dieser Grundlage passende Fach-, Raum- und Zeitbezüge vorschlägt. Aus den Vorschlägen kann ein Bearbeiter die passenden Bezüge auswählen. Die dem SNS zugrunde liegenden Fach- und Geothesauri unterstützen die inhaltliche Qualität der vergebenen Schlagworte und fördern die semantische Homogenität des Gesamtdatenbestands. Abbildung 6 zeigt einen Ausschnitt aus dem Verschlagwortungsassistent für Internet-Seiten. Über SNS wird eine URL analysiert und Deskriptoren, Raumbezüge und Zeitbezüge sowie optional eine Kurzbeschreibung der Seite zurück geliefert.

URL der Internetseite

Anzahl an Worten für Beschreibung

---

Ergebnis

Beschreibung

Deskriptoren

<input type="checkbox"/> Deskriptor
<input type="checkbox"/> Deskriptor 01
<input type="checkbox"/> Deskriptor 02
<input type="checkbox"/> Deskriptor 03
<input type="checkbox"/> Deskriptor 04

Raumbezug

<input type="checkbox"/> Administrative Einheit	Breite 1	Länge 1	Breite 2	Länge 2
<input type="checkbox"/> LK Aurich	234234234	234334543	343433333	982323223
<input type="checkbox"/> LK Leer	234234234	234334543	343433333	982323223
<input type="checkbox"/> LK Wittmund	234234234	234334543	343433333	982323223
<input type="checkbox"/> LK Emden	234234234	234334543	343433333	982323223

Zeitbezug

Auswahl	Bezeichnung	Datum	Typ
<input type="radio"/>	<Titel>	tt.mm.jj - tt.mm.jj	von - bis
<input type="radio"/>	<Titel>	tt.mm.jj	am
<input type="radio"/>	<Titel>	tt.mm.jj	seit

Abbildung 6: Erfassungsassistent.

Dieselbe Funktionalität wird bei der Erfassung von Objekten genutzt, um auf Basis eines Beschreibungstextes Vorschläge für Keywords vom SNS anzufordern. Abbildung 7 zeigt den Verschlagwortungsbereich der Erfassungskomponente mit der Option, zum einen den Verschlagwortungsassistenten aufzurufen und zum anderen

den Thesaurus-Navigator zu nutzen, mit dessen Hilfe man in der Hierarchie des Fachthesaurus „blättern“ kann.

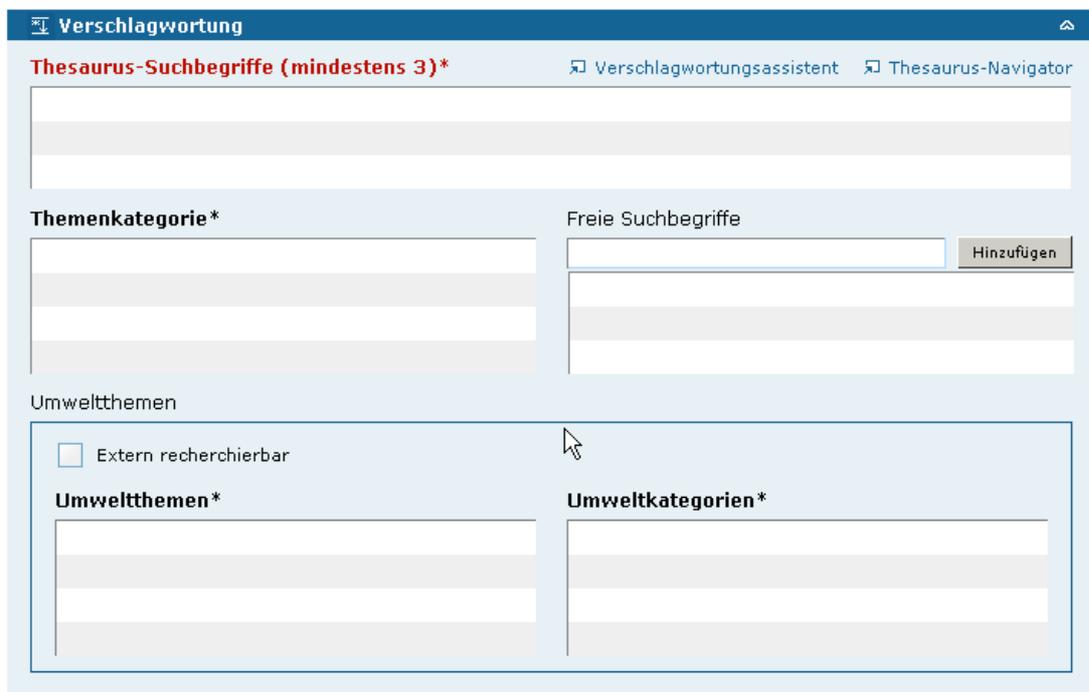


Abbildung 7: Verschlagwortung.

Den Thesaurus-Navigator zeigt Abbildung 8.

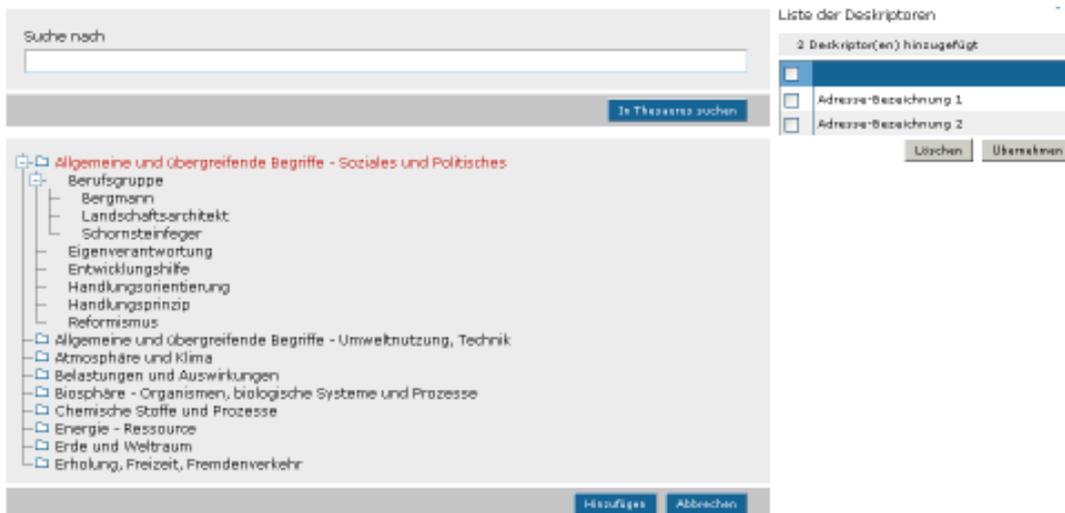


Abbildung 8: Thesaurus-Navigator.

Für die UDK-Klasse Dienst/Anwendung/Informationssystem, die auf die ISO 19119 abgebildet ist, existiert ein weiterer Assistent, der es ermöglicht, Service-Metadaten automatisch über eine Analyse der Capabilities des Services zu extrahieren. Die Maske des Assistenten zeigt Abbildung 9. Der Assistent kann Capabilities-Dokumente von WMS in den Versionen 1.1.0 und 1.1.1 und WFS in der Version

1.1.0 verarbeiten und trägt die Rückgaben direkt in die zugehörigen Felder des Service-Metadatenatzes ein.

Abbildung 9: Capabilities-Assistent.

### 4.3.3 Qualitätssicherung

Die Qualitätssicherung (QS) neu eingepflegter sowie die Kontrolle bereits vorhandener Metadaten ist erforderlich um sicherzustellen, dass ein Datenkatalog in sich konsistent und homogen ist und bleibt. InGrid® 1.1 sieht hierfür eine zweistufige Qualitätskontrolle vor: Neu eingepflegte Metadaten werden, bevor sie im Produktions-Datenkatalog abgespeichert werden können, einer manuellen Qualitätskontrolle unterzogen. Dabei werden die Metadaten automatisch an eine mit der Qualitätssicherung beauftragte Person weitergeleitet. Erst nach Durchlaufen der Qualitätssicherung können die Daten endgültig abgespeichert und der externen Recherche zur Verfügung gestellt werden.

Als weitere QS-Maßnahme werden, nach einer für jeden Katalog frei zu definierenden Verfallszeitspanne, alle Metadaten automatisch dem jeweiligen Datenverantwortlichen zur erneuten Kontrolle und Aktualitätsprüfung vorgelegt. Interne Mechanismen stellen sicher, dass die ISO-Pflichtfelder bereits vor der Weiterleitung eines Objekts an die QS befüllt wurden. Der QS-Zyklus ist für einen Katalog optional und nicht verpflichtend.

Die Nutzeroberfläche bietet sowohl Datenverantwortlichen als auch Qualitätssichernden eine Übersicht über aktuell in QS befindliche Dokumente und deren Status. Abbildung 10 zeigt beispielhaft die Zusammenstellung für einen Bearbeiter.

Name	Zugeordnet an	Typ	Datum
Umweltdatenkatalog (UDK)	Verantwortlicher	Verfallene Raumbestände	11.07.2006
Umweltdatenkatalog (UDK) des Saarlands	Bearbeiter/Verantwortlicher	Verfalldatum erreicht	11.07.2006
Umweltdatenkatalog (UDK) von Hessen	Qualitätssichernder	A	11.07.2006
Umweltdatenkatalog (UDK) von Hessen	Qualitätssichernder	L	11.07.2006

Abbildung 10: Übersicht der aktuell in QS befindlichen Objekte für den Datenverantwortlichen.

Der Lebenszyklus eines Objektes wird im Bereich Hierarchie & Erfassung gesteuert. Die wesentlichen Menüelemente sind in Abbildung 11 rot gekennzeichnet: : Objekt an QS überweisen, : Objekt als gelöscht markieren, : Löschmarkierung aufheben, : Änderungen zwischen dem letzten qualitätsgesicherten Stand und dem aktuellen Stand anzeigen.

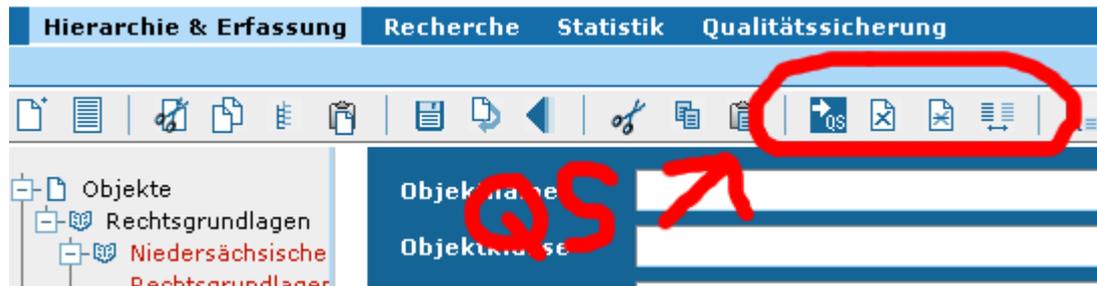


Abbildung 11: QS-Optionen im Bereich Hierarchie & Erfassung.

#### 4.3.4 ArcGIS-Metadatenimport

Zahlreiche Institutionen der öffentlichen Verwaltung nutzen Komponenten aus ESRI's ArcGIS-Produktfamilie um Geodaten zu prozessieren und zu dokumentieren. Für diesen Anwenderkreis wird eine XML-Schnittstelle zur Verfügung gestellt, über die Metadaten, die mit dem ArcGIS-Metadateneditor gepflegt werden, direkt in einen UDK überführt werden können. Die Befüllung der ISO-Pflichtfelder in den zu importierenden Datensätzen wird dabei über einen internen QS-Mechanismus sichergestellt.

Die Schnittstelle wird folgende Anwendungsfälle unterstützen:

- (1) Neu Anlegen eines Objektes in einem InGrid®-Katalog: Erstmaliges Überführen von Metadaten eines ArcGIS-Datensatzes.
- (2) Fortführen: Aufdatieren der Inhalte eines im InGrid®-Katalog existierenden Objektes mit aktualisierten Metadaten eines ArcGIS-Datensatzes.

Die Variante (2) ist nur möglich, wenn über eine ID die Zuordnung eines ESRI-Metadatenatzes zu einem InGrid-Metadatenobjekt möglich ist. Eine UUID wird bei Variante (1) automatisch erzeugt. Anschließend ist für das neue InGrid-Metadatenobjekt Variante (2) möglich.

## 5 Zusammenfassung

In der nächsten Ausbaustufe wird PortalU® ab 2008 mit InGrid 1.1® über eine web-basierte Metadaten-Erfassungskomponente verfügen, die sich nahtlos in die bestehende InGrid®-Architektur einfügt. Damit wird die angestrebte Zusammenführung der Vorgängersysteme GEIN, Windows-UDK und WWW-UDK abgeschlossen.

Die InGrid®-Software steht über eine Verwaltungsvereinbarung dem Bund und den Ländern kostenfrei zur Verfügung und stellt die deutsche Umweltverwaltung, durch innovative Architektur- und Schnittstellenkonzepte, im Hinblick auf die Anforderungen der Umweltinformationsgesetze und die geplanten nationalen und internationalen Dateninfrastrukturen, zukunftssicher auf.

## 6 Literaturverzeichnis

- [Angrick, 2002] Angrick, M. et al.: Semantic Network Services (SNS). Proceedings of the 16th Conference "Environmental Informatics 2002" (EnviroInfo'2002). Vienna, 2002.
- [Bilo, 2000] Bilo, M.; Streuff, H.: Das Umweltinformationsnetz Deutschland - GEIN2000 - Fachliche Anforderungen an ein Forschungs- und Entwicklungsvorhaben. 3rd Workshop Hypermedia im Umweltschutz. Ulm, 2000.
- [Karschnick, 2003] Karschnick, O. et al.: The UDK and ISO 19115 Standard. Proceedings of the 17th International Conference "Informatics for Environmental Protection". Cottbus, 2003.
- [Swoboda, 2000] Swoboda, W. et al.: Harmonisierter Zugang zu Umweltinformationen für Öffentlichkeit, Politik und Planung: Der Umweltdatenkatalog UDK im Einsatz. In: Tochtermann, K.; Riekert, W.-F. (Eds): Proceedings of the 14th International Symposium "Computer Science for Environmental Protection". Bonn, 2000.



# Ein Konzept zum Fachdokumenten-Management und zum Einsatz von InGrid im Umweltinformationssystem Baden-Württemberg (UIS BW)

Rainer Weidemann<sup>22</sup>,

Renate Ebel<sup>23</sup>,

Thorsten Schlachter<sup>1</sup>

[Rainer.Weidemann@iai.fzk.de](mailto:Rainer.Weidemann@iai.fzk.de)

[renate.ebel@lubw.bwl.de](mailto:renate.ebel@lubw.bwl.de)

[thorsten.schlachter@iai.fzk.de](mailto:thorsten.schlachter@iai.fzk.de)

## Abstract / Einleitung

Die Zusammenführung und übergreifende Erschließung der heterogenen und meist verteilt vorliegenden Informationen im Umweltbereich stellt nach wie vor eine Herausforderung dar. Mit PortalU<sup>®</sup> und der diesem zu Grunde liegender InGrid<sup>®</sup>-Software wurde auf Bundesebene ein anspruchsvoller Ansatz zur Vernetzung von Umweltinformationen umgesetzt [Klenke, 2006]. Die Bundesländer können nun zum einen ihre Informationsquellen über die vorgesehenen Schnittstellen (iPlugs) an PortalU anbinden und zum anderen für eigene Portallösungen die InGrid-Software nutzen. Für das UIS BW wurden die für PortalU relevanten Informationsangebote identifiziert und ein Konzept zur Zusammenführung und Einbindung der Informationen entwickelt. Weiterhin wurden mögliche Einsatzszenarien für die InGrid-Software im UIS BW analysiert und bewertet. Mit den ersten Umsetzungsschritten wurde begonnen.

## 1 Informationsbereitstellung für PortalU

In PortalU können sowohl schwach strukturierte Informationsquellen eingebunden werden, wie z.B. Webseiten oder Fachdokumente, als auch Kartendienste und Objekte mit definierter Struktur (UDK-Objekte, Forschungsprojekte). Im UIS BW sind die für PortalU relevanten Informationsquellen über eine Vielzahl von Websites und Fachsysteme verteilt. Die modulare Struktur von InGrid, insbesondere die adaptierbaren Schnittstellenkomponenten (iPlugs) erleichtern die Zusammenführung der verteilt vorliegenden Informationen. Trotzdem kann es sich lohnen, in einer vorgelagerten Stufe Informationen zu verdichten oder anzureichern und erst mit dem dabei entstehenden Zwischenprodukt die Schnittstellen von PortalU zu versorgen. Dies kann dann zum Tragen kommen, wenn komplexere Verarbeitungsschritte notwendig sind, um z.B. in Datenbanken oder Content Management Systemen abgelegte Daten zu extrahieren, zu ergänzen und in geeignete Strukturen zu

---

<sup>22</sup> Forschungszentrum Karlsruhe, IAI, Rainer.Weidemann|Thorsten.Schlachter@iai.fzk.de

<sup>23</sup> Landesanstalt für Umwelt, Messungen und Naturschutz BW, Renate.Ebel@lubw.bwl.de

bringen. Als konkretes Beispiel dazu wird nachfolgend der geplante Umgang mit Fachdokumenten im UIS BW beschrieben.

## **1.1 Fachdokumenten-Browser**

Auch unabhängig von PortalU besteht im UIS BW der Bedarf, transparent auf alle, verteilt in verschiedenen Fachsystemen gepflegten Fachdokumente zuzugreifen [Weidemann, 2006]. Als Fachdokument wird dabei grob eine mit einem Metadatensatz beschriebene Datei bezeichnet, die Fachinformationen enthält und eine längere Lebensdauer besitzt. Diese Definition erlaubt es, nicht nur Dateien mit textuellen Informationen einzubeziehen, sondern prinzipiell Dateien beliebigen Typs, also insbesondere auch Bilder, Audiodateien und Videos. Eine Zentralisierung der Dokumentenverwaltung erscheint aus organisatorischen Gründen nicht geeignet und ist aus technischer Sicht auch nicht erforderlich. Für eine übergreifende Sicht werden die Fachsysteme im Sinne einer Service-Architektur verpflichtet, bestimmte Dienste (Webservices) anzubieten, und zwar mindestens eine Auskunft, welche Fachdokumente vorliegen und welche Metadaten zu diesen Fachdokumenten vorhanden sind. Die Schnittstelle erlaubt dem Fachsystem, die Sichtbarkeit seiner Fachdokumente zu kontrollieren, und verbirgt deren interne Repräsentation.

Auf dieser Basis wurde ein höherwertiger Dienst konzipiert (Fachdokumenten-Browser, FDB), der im Hintergrund wie ein Volltext-Crawler die Fachsysteme abfragt und Metadaten zu allen UIS-Fachdokumenten pflegt. Der momentan in der Realisierungsphase befindliche FDB wird vereinheitlichte Sichten auf die einzelnen Fachdokumente bereitstellen, sowohl in Form einer HTML-Seite (Metadaten plus Link auf Fachdokument-Datei) als auch als Webservice. Ebenso werden übergreifende Such- und Navigationsfunktionen als Webservices bereitgestellt, die z.B. in Portalen eine transparente Einbindung von Fachdokumenten erlauben. Es bietet sich an, den FDB zusätzlich dazu zu nutzen, die Indizierung der Fachdokumente durch Internet-Suchmaschinen zu optimieren, zumal ein Großteil der Fachdokumente direkt in den Fachsystemen für die Crawler der Suchmaschinen sonst nicht erreichbar wäre. Zu diesem Zweck soll die von Google, Yahoo und Microsoft forcierte XML-Sitemap [Sitemaps, 2007] eingesetzt werden.

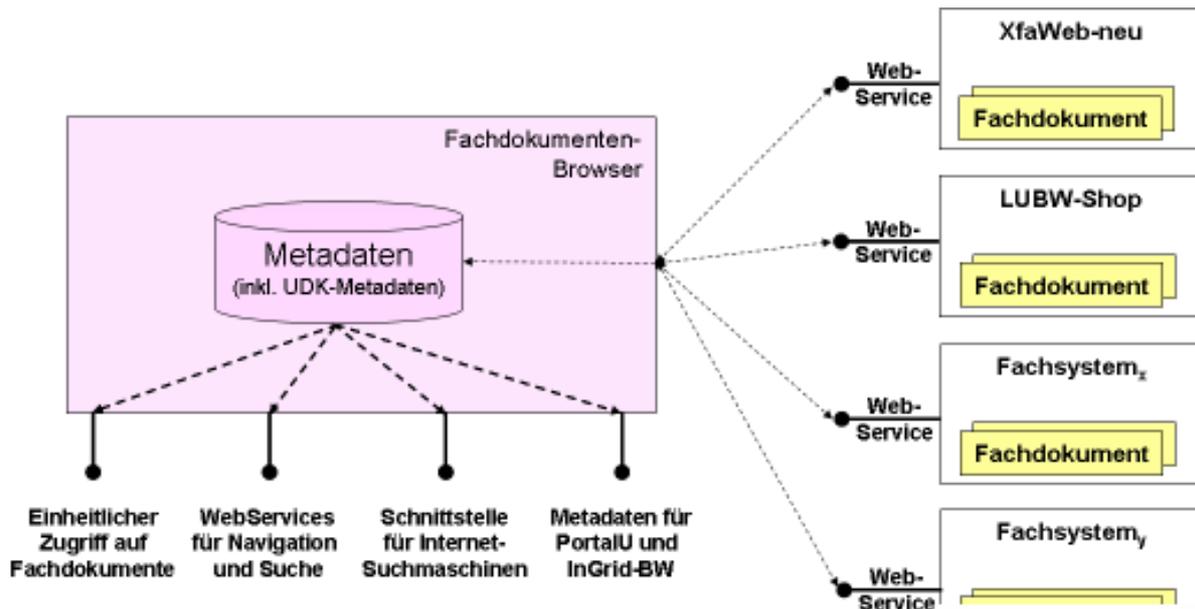


Abbildung 1: Der Fachdokumenten-Browser (FDB) als zentrale Zugangskomponente für Fachdokumente im UIS BW (XfaWeb-neu und LUBW-Shop stehen beispielhaft als Fachsysteme, die Fachdokumente enthalten).

Bezüglich PortalU kann der FDB nicht nur die Einbindung der Fachdokumente in die Volltextsuche unterstützen, sondern diese Fachdokumente auch als UDK-Objekte beschreiben. Hierbei kommt zum Tragen, dass PortalU bzw. InGrid den Umweltdatenkatalog und die zugehörigen Pflegewerkzeuge ablösen werden. Auf dieser Basis wird es dann leichter möglich sein, die UDK-Metadaten dezentral zu pflegen und in PortalU / InGrid lediglich zusammenzuführen. Im UIS BW soll künftig die Strategie verfolgt werden, die Pflege der Metadaten näher an den Ort ihrer Entstehung heranzuführen. Die entsprechende Analyse für das UIS BW ergab, dass abhängig von der UDK-Klasse unterschiedlich vorgegangen werden sollte. Der FDB wird zusätzlich UDK-Metadaten für Fachdokumente aus den Fachsystemen ermitteln, bei Bedarf über Heuristiken ergänzen und geschlossen weitergeben. Für Geodatenbeschreibungen und teilweise auch für Fachdaten ist der Einsatz einer eigenen Metadatenverwaltung auf Basis von Disy Preludio geplant. Diese soll, ebenso wie der FDB, über eine geeignete Schnittstelle (DSC-iPlug, s. Kap. 5) an InGrid angeschlossen werden. Objekte der übrigen UDK-Klassen sollen weiter zentral mit InGrid verwaltet werden.

## 2 InGrid-BW

Seitens der Umweltverwaltung Baden-Württembergs besteht die Anforderung, weiterhin ein eigenes Umweltportal für die Landes- und zunehmend auch für die kommunalen Informationsangebote zu betreiben. Neben Performance-Aspekten steht der Wunsch im Vordergrund, die Sortierung / Gewichtung der Suchergebnisse individuell regeln und bei Bedarf auch in PortalU nicht integrierte Inhalte mit aufnehmen zu können. Das bisherige Umweltinformationsnetz Baden-Württemberg (UINBW) mit seinem Frontend, dem Portal Umwelt-BW, ist seit November 2004 im Produktionsbetrieb [Schlachter, 2006]. In einer Studie wurden Alternativen zur Weiterentwicklung des bisherigen, mit dem Content Management System

WebGenesis und dem Suchwerkzeug [ht://Dig](http://Dig) realisierten UINBW untersucht, insbesondere Möglichkeiten Teile des Systems durch InGrid-Komponenten zu ersetzen. Die Service-Architektur von InGrid erlaubt eine flexible Kombination der Komponenten, sodass zwischen einer eigenen InGrid-Vollinstallation und der weitgehenden Nutzung von PortalU ein breites Spektrum an Szenarien möglich wird. Als zusätzliche Randbedingungen an das Lösungskonzept waren zu beachten, dass bisherige Funktionalitäten erhalten, beispielsweise die Einbindung bestimmter Informationsquellen über speziell zugeschnittene Schnittstellen, und bereits vorhandene oder geplante Suchebenen reproduziert werden sollten. Als Suchebenen seien hier Suchfunktionen bezeichnet, die auf einen vorgegeben Teil des gesamten Suchraums beschränkt bleiben, wie z.B. die LUBW-Suche, welche nur unmittelbare Inhalte des Informationsangebots der Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg (LUBW) finden soll und die FDB-Suche, die auf Fachdokumente beschränkt bleiben soll.

Das aus der Analyse und Bewertung möglicher Szenarien abgeleitete Lösungskonzept (InGrid-BW) geht von einer eigenen Installation des iBus, der Kernkomponente der InGrid-Software, im UIS BW aus. Die Integration von Informationsquellen erfolgt über entsprechende iPlugs, die in der Regel gleichzeitig an PortalU und InGrid-BW angebunden werden. Um die verschiedenen Suchebenen zu realisieren, werden für jede Ebene eigene SE-iPlugs (Crawler für die Volltextsuche) verwendet. Auch für die Zusammenführung der UDK-Metadaten werden mehrere iPlugs benötigt, welche die Datenstrukturen der dezentralen Metainformationssysteme auf das UDK-Profil abbilden.

Für die Einbindung von UDK-Metadaten gibt es allerdings die Alternative, die iPlugs nicht zusätzlich direkt an PortalU anzubinden, sondern PortalU indirekt mit Daten zu versorgen, genauer über einen ECS-iPlug der an die entsprechende Schnittstelle von InGrid-BW (iBus) ankoppelt. Das hätte zur Folge, dass die UDK-Objekte aus Baden-Württemberg nicht verteilt in der gerankten („linken“) Ergebnisliste von PortalU erscheinen, sondern zusammengefasst bleiben („rechte“ Ergebnisliste in PortalU). In der linken Ergebnisliste erscheinen Informationen, die durch PortalU indiziert wurden und bezüglich einer aktuellen Anfrage bewertet und sortiert (Ranking) werden können. Die rechte Ergebnisliste enthält zu PortalU externe Informationsquellen, an die Suchanfragen zur Laufzeit weitergereicht werden und für die erst einmal nur ein summarisches Suchergebnis angezeigt wird. Unter den Anwendern wird durchaus lebhaft diskutiert, wo die eigenen Informationen erscheinen sollen, stark beeinflusst durch den Ranking-Algorithmus bzw. die durch diesen typischerweise erreichte Positionierung der eigenen Angebote in der linken Liste.

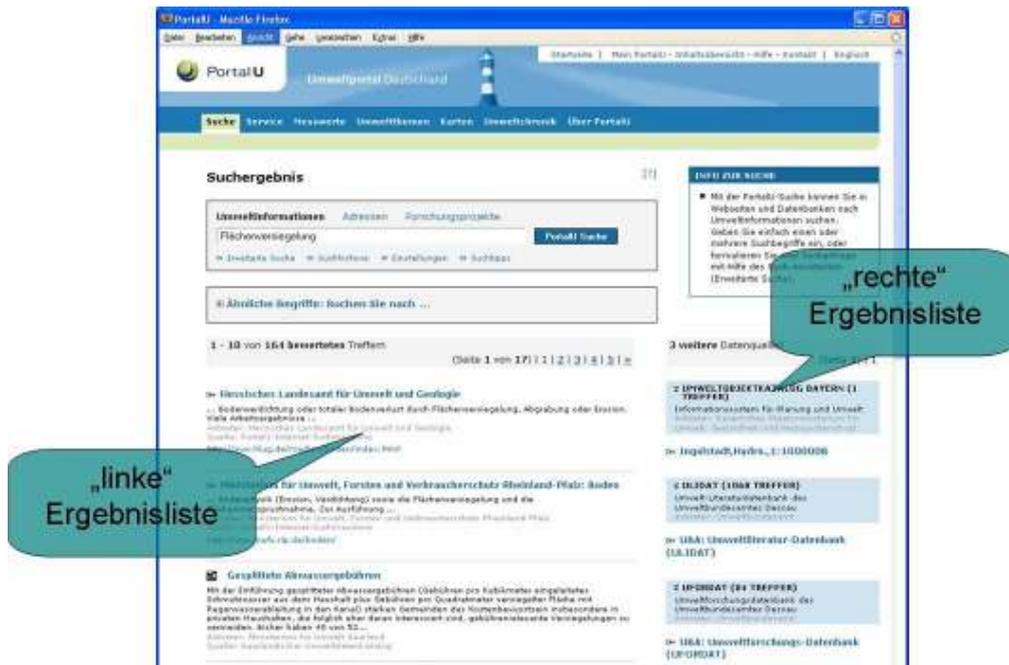


Abbildung 2: Beispielhafte Ergebnisliste von PortalU

Das Portal-Modul von InGrid kann in Layout und Funktionalität nur eingeschränkt modifiziert werden. Schon allein um die Designvorgaben des Landes erfüllen zu können, muss deshalb die OpenSearch-Schnittstelle des iBus verwendet werden [OpenSearch, 2007]. Damit kann die Suchfunktionalität von InGrid-BW flexibel in Portale und Dienste eingebunden werden, z.B. in den Fachdokumenten-Browser, die Suche auf der Website der LUBW und insbesondere in UINBW. InGrid-BW vereinheitlicht somit auch die Suche im UIS BW und es entfällt die bisherige Mehrfachindizierung von Informationsquellen. Die Administration der Suche kann an einer Stelle zusammengefasst werden (Portal von UINBW).

Die Nutzung der OpenSearch-Schnittstelle hat den weiteren Vorteil, dass sich zusätzliche Suchmöglichkeiten realisieren lassen, da auch Attribute angesprochen werden können, die im Portal-Modul nicht zugänglich sind. Erkauft wird dies allerdings damit, dass ohne das Portal-Modul die kartenbasierte Suche nicht verfügbar ist und derzeit über OpenSearch nur die gerankte Ergebnisliste geliefert wird.

Die folgenden beiden Abbildungen zeigen die wesentlichen Teile des Konzepts im Überblick.

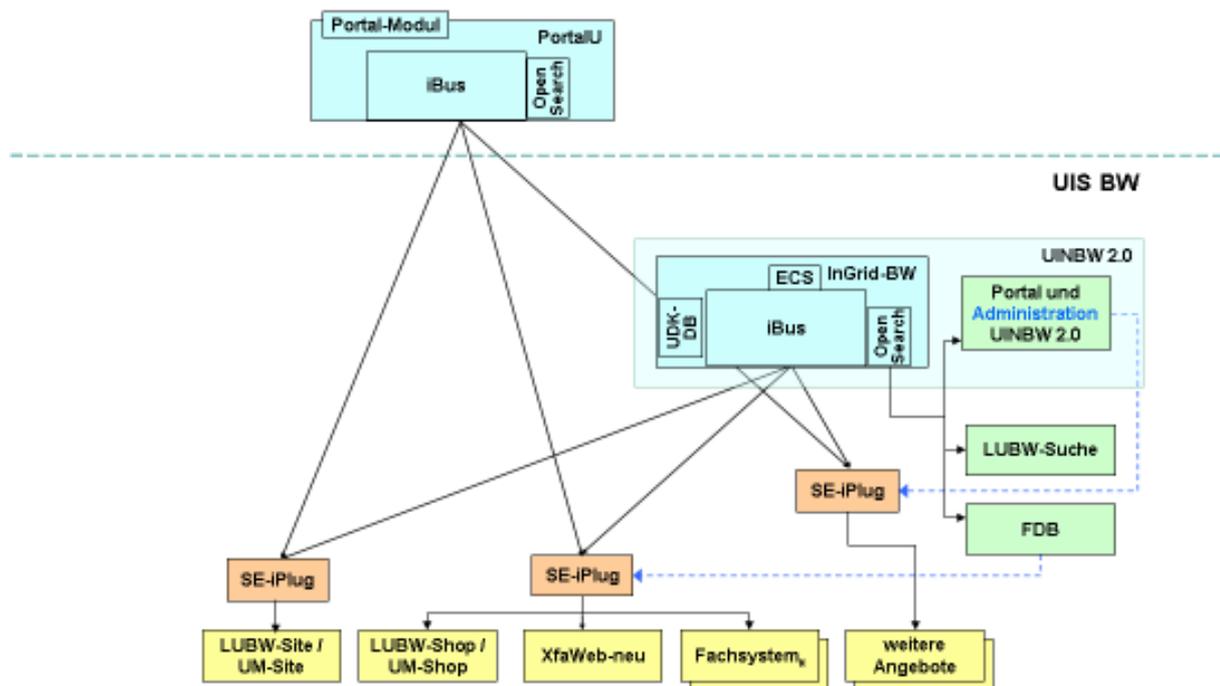


Abbildung 3: Konzept für den Einsatz von InGrid-Komponenten im UIS BW (Volltextsuche)

Die nächste Generation des Umweltinformationsnetzes Baden-Württemberg (UINBW 2.0) soll auf einer landeseigenen Installation der InGrid-Software basieren (InGrid-BW). Es wird jedoch ein separates Portal verwendet, das die OpenSearch-Schnittstelle für Abfragen benutzt. Das Portal soll auch eine Administrations-Schnittstelle zur Konfiguration der Volltextsuche anbieten (Abbildung 3). Die Indizierung der Informationsquellen für die Volltextsuche erfolgt verteilt über mehrere SE-iPlugs, um weitere Suchebenen (LUBW-Suche, FDB) zu realisieren – bei Suchanfragen kann die Id des iPlugs mit als Suchkriterium verwendet werden.

Abbildung 4 schließlich beschreibt den Umgang mit UDK-Metadaten, wobei davon ausgegangen wurde, dass PortalU diese Metadaten indirekt über InGrid-BW erhält (für die rechte Ergebnisliste, s.o.). Die als „DSC-iPlug Projekte“ bezeichnete Schnittstellenkomponente transportiert spezielle Metadatenätze zu Forschungsprojekten, für die PortalU / InGrid eigene Suchmöglichkeiten anbietet.



Verkehr in neuen Verwaltungsstrukturen, Phase I 2005/06, Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 7250, S. 53 - 62, 2006  
<http://www2.lubw.baden-wuerttemberg.de/lfu/uis/kewa1/08-fzk-uis/kewa1-fzk-uis.html>

[Sitemaps, 2007]

Sitemaps.org, <http://www.sitemaps.org/>, besucht am 25.04.2007

[Weidemann, 2006]

Weidemann, Rainer et al.: FADO BW – Konzept für das Fachdokumentenmanagement im Umweltinformationssystem Baden-Württemberg und zur Weiterentwicklung der XfaWeb-Fachinformationssystemfamilie. In Mayer-Föll, R.; Keitel, A.; Geiger, W. (Hrsg.): Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt und Verkehr in neuen Verwaltungsstrukturen, Phase I 2005/06, Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 7250, S. 7 - 17, 2006  
<http://www2.lubw.baden-wuerttemberg.de/lfu/uis/kewa1/03-fzk-fado/kewa1-fzk-fado.html>

## 5 Abkürzungen / Glossar

DSC	Data Source Client (DSC-iPlug: Basis-Schnittstelle zu strukturierten Daten)
ECS	Environmental Catalogue Service (ECS-iPlug: UDK-Schnittstelle)
FDB	Fachdokumenten-Browser, in Entwicklung befindliches Metainformationssystem zum übergreifenden Zugriff auf Fachdokumente in unterschiedlichen Fachsystemen des UIS BW
iBus	Zentrale Komponente von InGrid
InGrid	Für PortalU realisiertes Softwarepaket, das auch für Umweltportale der Länder und anderer Stellen genutzt werden kann.
InGrid-BW	Geplante Installation der InGrid-Software im UIS BW als Teil von UINBW 2.0
iPlug	Auf bestimmte Art von Datenquellen spezialisierte Schnittstellen- komponente
LUBW	Landesanstalt für Umwelt, Messungen und Naturschutz Baden- Württemberg
PortalU	Das in einer Bund-Länder-Kooperation entwickelte Umweltportal Deutschland; in Betrieb seit Mai 2006, <a href="http://www.portalu.de">http://www.portalu.de</a>
Portal Umwelt-BW	Portal des Umweltinformationsnetzes Baden-Württemberg (UINBW, s.u.)
SE	Search Engine (SE-iPlug: Schnittstelle für Volltext-Crawler)
UDK	Umweltdatenkatalog
UINBW	Das Umweltinformationsnetz Baden-Württemberg erschließt text- orientierte behördliche und kommunale Informationsangebote aus dem UIS BW mit übergreifenden Zugangsfunktionen <a href="http://www.umwelt.baden-wuerttemberg.de/">http://www.umwelt.baden-wuerttemberg.de/</a> UINBW 1.x = aktuelle Versionen UINBW 2.0 = geplante, auf InGrid basierende Version
UINST	Umweltinformationsnetz Sachsen-Anhalt, basiert auf der gleichen technischen Grundlage wie UINBW <a href="http://www.umwelt.sachsen-anhalt.de/">http://www.umwelt.sachsen-anhalt.de/</a>
UIS BW	Im Umweltinformationssystem Baden-Württemberg wird die gesamte Verarbeitung von Umweltinformationen unter der Regie des Umweltministeriums Baden-Württemberg koordiniert <a href="http://www.lubw.baden-wuerttemberg.de/servlet/is/2886/">http://www.lubw.baden-wuerttemberg.de/servlet/is/2886/</a>
UM	Umweltministerium Baden-Württemberg
WIBAS-OK	Objektkatalog / Metadatenkatalog von WIBAS (vormals WAABIS: <a href="http://de.wikipedia.org/wiki/WAABIS">http://de.wikipedia.org/wiki/WAABIS</a> )
XfaWeb-neu	Neuentwicklung der XfaWeb-Systemfamilie [Weidemann, 2006]