

Entwicklung einer Umwelt- und Forschungsdatenbank für die Lausitzer Bergbaufolgelandschaften

Suzan Yigitbasi*, Bernhard Thalheim, Kati Seelig
und Sabine Radochla

Zusammenfassung

Im Umweltbereich und ähnlich heterogenen Anwendungsfeldern nehmen die Anforderungen an Client/Server-Datenbankanwendungen rasch recht komplexe Form an. Die Entwicklungsarbeiten zur Umwelt- und Forschungsdatenbank des BTUC Innovationskollegs „Bergbaufolgelandschaften,, geben ein Beispiel dafür. Dies soll im folgenden, anhand der wesentlichen Grundlagen und Anforderungen, skizziert werden. Im Vordergrund stehen dabei die Informations-Modellierung im Rahmen des Entity-Relationship Ansatzes und die Einarbeitung von Metainformationen in das Datenbankkonzept. Daneben werden Ansätze zur Benutzerunterstützung sowie Methoden zur Erweiterung der vorhandenen Datenmodelle vorgestellt, da ihnen vor allem in Bezug auf die Nutzbarkeit der Datenbankanwendung große Bedeutung zukommt.

1. Einleitung

Aufgrund der schwindenden Zahl naturnaher Lebensräume rücken Schutz und Erhalt unserer Umwelt immer mehr in den Vordergrund des allgemeinen Interesses. In diesem Zusammenhang stellt sich die Frage, welche Möglichkeiten uns die Renaturierung von Industriebrachen wie den ehemaligen Tagebauen im Lausitzer Braunkohlerevier bietet, um zusätzliche und zu einem späteren Zeitpunkt hoffentlich wieder ökologisch wertvolle Lebensräume zurück zu gewinnen. Hier ist anzumerken, daß es sich um stark gestörte Ökosysteme handelt, deren internen Wechselwirkungen und Stoffverlagerungsprozesse erst wenig bekannt sind. Daher bietet es sich an, die Sanierungsmaßnahmen durch umfangreiche Forschungsprogramme, wie das von der DFG geförderte BTUC Innovationskolleg „Bergbaufolgelandschaften,, zu unterstützen. Ihre Aufgabe ist es Wissen zu sammeln und daraus Methoden für die Renaturierung abzuleiten und deren Grenzen aufzuzeigen.

Um ein möglichst vollständiges Bild vom Entwicklungspotential der Lausitzer Bergbaufolgelandschaften zu erhalten, werden die unterschiedlichen Teilbereiche

* BTU Cottbus, Lehrstuhl DBIS, Postfach 10 13 44, 03013 Cottbus.

und Subsysteme: Kippen, Tagebaurestseen, unbehandelte Freiflächen, meliorierte, rekultivierte Kiefern- und Eichenforste¹ durch derzeit 21 Teilprojekte des BTUC Innovationskollegs untersucht. Abbildung 1 zeigt die verschiedenen Untersuchungsthemen. Bei der Bearbeitung ihrer Aufgaben, der Abschätzung des ökologischen Entwicklungspotentials, sind die Fachwissenschaftler in den einzelnen Teilprojekten neben den eigenen Ergebnissen auch auf die Auswertungen ihrer Kollegen angewiesen. Dies setzt neben interdisziplinärer Zusammenarbeit auch einen einfachen Daten- und Informationsaustausch voraus. Zur Sicherung und Bereitstellung der gesammelten Daten wird im Rahmen des Teilprojekts DENDA (Dynamic Environmental Databases) eine relationale, Umwelt- und Forschungsdatenbank entwickelt.

Teilprojekte des Innovationskollegs Bergbaufolgelandschaften

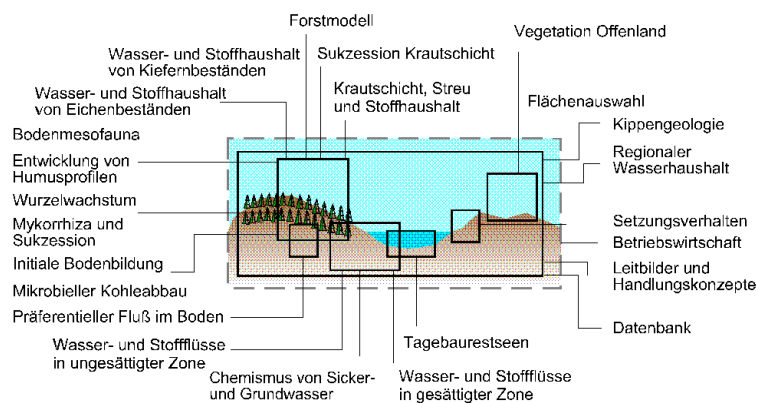


Abbildung 1:
Die Untersuchungsthemen innerhalb des BTUC Innovationskollegs.

2. Anforderungen an die Datenbank

Beim Innovationskolleg handelt es sich um ein langfristig angelegtes und von den Untersuchungsthemen her breit gefächertes Forschungsprojekt. Die Hauptaufgabe im Rahmen der Datenbankentwicklung besteht in der Integration der anfallenden, heterogenen Datenbestände, der Ermöglichung einer konsistenten Nutzung der Datenbank unter Berücksichtigung der an die Datenbank gestellten Qualitätsanforderungen:

¹ Die zu verschiedenen Zeiten aufgeforsteten Kiefern- und Eichenbestände werden dabei als Chronosequenz aufgefaßt.

Schemakonsistenz

In einigen Fällen liegen Überschneidungen hinsichtlich der Untersuchungsthemen und angewandten Verfahren vor. Hier ist es notwendig, auch bei voneinander abweichenden Meßmethoden ein konsistentes Gesamtschema zu finden, aus dem sich die unterschiedlichen externen Schemata ableiten lassen.

- *Flexibilität des Datenbankschemas – einfache Erweiterbarkeit:*
Der Schwerpunkt des Innovationskollegs liegt im Forschungsbereich, so daß die verwendeten Untersuchungsmethoden meist nicht unveränderlich sind, sondern den Umständen entsprechend abgewandelt oder durch Neue ersetzt werden.
- *Einbindung der Metainformationen:*
Die zu integrierenden Umweltinformationen werden von Arbeitsgruppen verschiedener Fachbereiche erhoben. Um allen Teilprojekten einen Einblick in die gesammelten Daten zu gewähren, müssen neben den Meßdaten auch die zur Interpretation notwendigen Metainformationen bereitgestellt werden (Kremer/Krasemann 1996 und Zamperoni et. al. 1996).
- *Schutz der Daten gegenüber unbefugten Zugriffen:*
Ein Großteil der Daten wird im Rahmen von Diplom- und Doktorarbeiten erhoben. Schon der konzeptionelle Entwurf muß es erlauben, daß der Erfasser Lese-rechte für von ihm gesammelte Informationen dem gesamten Innovationskolleg, einzelnen Teilprojekten oder nur bestimmten Nutzern einräumt.
- *Einfache Bedienbarkeit der Datenbankoberfläche*
Da es sich bei den Benutzern nicht um Datenbankexperten handelt, ist es notwendig einfach bedienbare, auf die Benutzerwünsche abgestimmte Arbeitsoberflächen zu erstellen.

3. Konzeption der Datenbank

Der derzeitige Prototypen stellt ein „einfaches,, Datenbankanwendungssystem dar, bestehend aus drei relationalen „Teildatenbanken,, . Die Datenbank selbst liegt auf einem Sybase SQL Server Version 11, vgl. Abbildung 2.

- Der Abschnitt *Metainformationen* enthält übergeordnete Informationen zu Teilprojekten, Aufgaben, Personen und Untersuchungsflächen sowie detaillierte Angaben zu den Untersuchungsmethoden und Meßgrößen. Dies gewährleistet eine langfristige Interpretierbarkeit der Meßdaten. Dabei verstehen wir unter Metainformationen:

- Im mittlerweile abgeschlossenen Teilprojekt *Flächenauswahl* fanden Voruntersuchungen auf verschiedenen ehemaligen Tagebauflächen statt, die hinsichtlich ihrer Eignung für die Chronosequenzstudien des Innovationskollegs beprobt wurden. Nicht alle Flächen werden weiter untersucht, weshalb diesen Informationen ein eigener Datenbankabschnitt gewidmet ist.
- Die *Meßdatenbank* enthält, nach Themen geordnet, die reinen Meßdaten aller derzeit beprobten Versuchsflächen (meliorierte, aufgeforstete Eichen- und Kiefernflächen, unbehandelte Offenlandflächen sowie Kippen und Restseen). Auch die Unterabschnitte sind weiter nach Themen bzw. Schlagwörtern unterteilt. So enthält die Teildatenbank Boden alle Auswertungen zu Bodenphysik (Korngrößen, Substratart, etc.) und Bodenchemie (z. B. Standortcharakteristik, Daten zu den chemischen Fraktionen bestimmter Elemente und deren Häufigkeit).

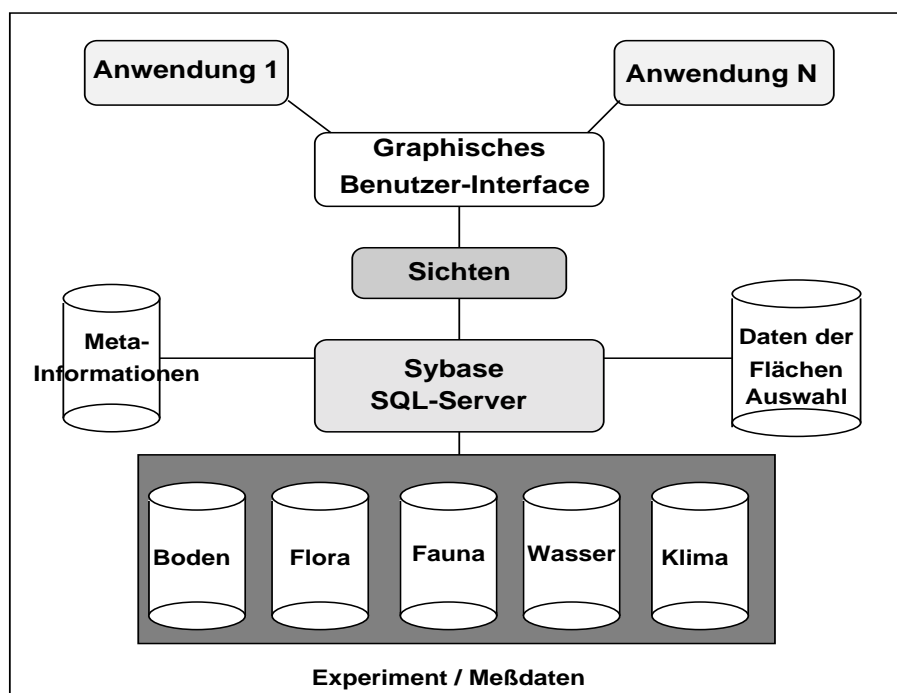


Abbildung 2
Schematischer Aufbau der Datenbank.

Die Abbildung zeigt auch Möglichkeiten für die Erweiterung der Datenbank. Sollen zu einem späteren Zeitpunkt neben den Meßdaten noch Simulationsergebnisse und Laborversuche in die Datenbank integriert werden, so kann in analoger Weise zur Meßdatenbank einfach eine Teildatenbank *Laborversuche / Simulation* eingefügt werden. Entsprechend ist dann der Abschnitt Metainformationen um Angaben zu den Versuchen und Simulationen zu ergänzen. Da in den nächsten Monaten die Einmessungs- und Kartierarbeiten zur räumlichen Referenzierung der Flächen samt darauf befindlichen Meßstellen abgeschlossen werden, ist das Datenbankschema auch um den Abschnitt *Koordinaten / Quadranten* zu erweitern. Er wird die Zuordnung der Meßpunktnamen bzw. Kürzel zu den jeweiligen Quadranten enthalten.

Die Kommunikation zwischen den Benutzern und der Datenbank erfolgt über graphische, ereignisgesteuerte Benutzungsschnittstellen, die Mithilfe der Oberflächenentwicklungsumgebung Powerbuilder 5 erstellt wurden.

4. Entwicklung und Bereitstellung der Datenbank

Die Entwicklungsarbeiten zur Realisierung des Datenbankanwendungssystems können in zwei große Abschnitte eingeteilt werden, den Aufbau der Datenbank und die anschließende Entwicklung der graphischen Oberflächen.

4.1 Datenbankentwicklung und Implementation

Im ersten Entwicklungsabschnitt, der Modellierung der verschiedenen Untersuchungsbereiche, wie Stoffhaushalt oder Sickerwasserchemismus, werden externe Datenmodelle entworfen. Wir können die betrachteten Modelle als Abstraktionen zur Beschreibung der jeweiligen Teilbereiche verstehen, vgl. Abbildung 3. In dieser Phase wird zusammen mit den Erfassern, ihren Versuchs- und Probenahmebeschreibungen sowie prototypischen Meßdaten der jeweilige Umwelt- / Informationsaustausch und die der Untersuchung zugrundeliegende Semantik analysiert.

Im Fall der Sickerwasserchemie beispielsweise werden als systemrelevante Zustandsgrößen Anionen- und Kationenkonzentrationen sowie der pH-Wert in Abhängigkeit von der Bodentiefe ermittelt. Diese Größen bilden zusammen mit den zugeordneten Hintergrundinformationen ein externes Modell zur Beschreibung der Sickerwasserchemie. Weil die Untersuchungen zum Sickerwasser in Eichenbeständen vom Teilprojekt 4, in Kiefernbeständen jedoch durch Teilprojekt 3 vorgenommen werden, treten geringfügige Abweichungen in Bezug auf die Anzahl der betrachteten Ionen und das Probenahmeverfahren auf. Es sind daher zwei externe Modelle zu berücksichtigen. Für die eingangs geforderte Konsistenz der Datenbank ist entscheidend, daß die beiden externen Modelle soweit mit einander kompatibel sind,

daß wir sie auf ein einziges konzeptionelles Modell abbilden können (Vinek et. al. 1982). An dieser Stelle zeigt sich, welche Bedeutung der methodischen Abstimmung zwischen den Teilprojekten zukommt. Erst durch sie können Probleme aufgrund unterschiedlicher Betrachtungsweise aufgelöst und ein einheitliches Modell entwickelt werden.

Im Fall der Sickerwasserchemie ist dies gewährleistet, da wir die zusätzlichen Informationen des Teilprojekts 4 zu den Schwermetallkonzentrationen, vgl. Tabelle 1, einfach als Ergänzungen eines externen Modells „Sickerwasser-Standard,“ auffassen können.

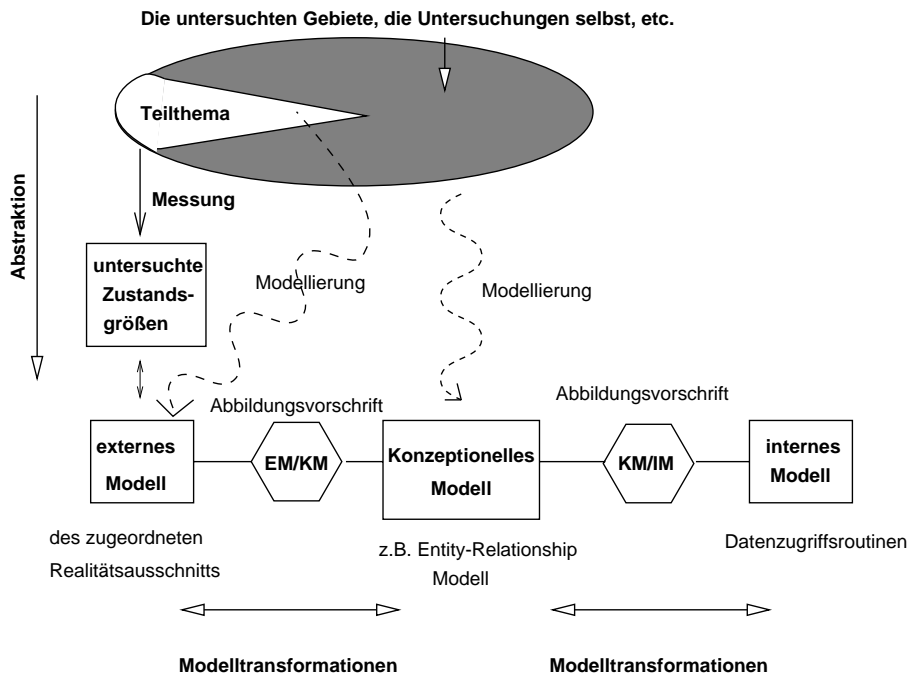


Abbildung 3
Skizze zur Datenbankmodellierung

Tabelle 1 liefert einen Überblick zu den Meß- und Metadaten aus dem Bereich der Sickerwasserchemie und zeigt, daß die eher deklarativen Metainformationen von den tatsächlichen Meßwerten getrennt werden müssen und ihnen ein eigener Datenbankabschnitt „Metainformationen,“ zu zuordnen ist. Das diesen Informationen zu-

geordnete Teilschema wird mit den Mitteln des Entity-Relationship Ansatzes (P. Chen 1976 und B. Thalheim 1993) entwickelt, vgl. Abbildung 4 und 5.

Tabelle 1: Zur Modellierung der Eingangsinformationen

	Teilprojekt 3	Teilprojekt 4
Tiefen [cm]	20, 40, 70, 100	20, 40, 80, 160
Probenart	Monatsmittelwert, Sickerwasser	Monatsmittelwert, Sickerwasser
Gerät	P80 Saugkerzen	P80 Saugkerzen
Verfahren	6 Kerzen pro Tiefe	8 Kerzen pro Tiefe
Behandlung der Proben	Filtration (Nylonfilter, Maschenweite 45µm)	Filtration (Nylonfilter, Maschenweite 45µm)
Parameter	pH, el. Leitf., DOC, SO ₄ , NO ₃ , Cl, Ca, Mg, Al, Fe, Mn, Na, K.	pH, el. Leitf., DOC, SO ₄ , NO ₃ , Cl, Ca, Mg, Al, Fe, Mn, Na, K. PO ₄ , BO ₃ , Cd, Zn, Pb, Cu.

Abbildung 4 entnehmen wir, daß jede Sickerwasseranalyse durch die Angabe des Meßpunktes **pk-Messpkt**¹, er liefert uns auch die Versuchsfläche, der Meßtiefe **pk-Messtiefe** und dem Meßdatum **pk-Messdat** von allen anderen Sickerwasseranalysen unterscheidbar ist. Wird noch die Probenart berücksichtigt, so können alle Analysen eindeutig identifiziert werden. Wir können daher die Standardanalyse als Entitätstyp *sw_std* behandeln. Abhängig davon, ob die Analyse durch das Teilprojekt 3 oder 4 erfolgt, dies ist sowohl über den Erfasser und das Datum als auch den Meßpunkt ersichtlich, ist eine Ergänzungstabelle *sw_erw* zu berücksichtigen. Sie wird wieder als Entity angesetzt und durch den Beziehungstypen *hat* mit den Standardanalysen verknüpft. Bei übereinstimmenden Schlüsselattributen **pk-Messpkt**, **pk-Messdat**, **pk-Messtiefe** und **pk-Messart** sind Standard- und Schwermetallanalysen verfügbar. Die Beschriftung der Verbindungspfeile gibt die Kardinalität an. Zu einem Eintrag in der *sw_std* Tabelle gehört maximal ein Eintrag in der Ergänzungstabelle, jedem Ergänzungseintrag kann eindeutig ein Standardeintrag zugeordnet werden. Wir können das Ganze auch abkürzend, wie in der unteren Hälfte von Abbildung 4 dargestellt, schreiben. Unter Berücksichtigung der weiteren Analysen, wurden die für das Meßprogramm relevanten Informationen im Entity *messung* zusammengefaßt und vor die Sickerwasser spezifischen Informationen geschaltet.

In analoger Weise können die weiteren „punktuellen“, Beprobungen in die Meßdatenbank eingebettet werden. Treten zu einem späteren Zeitpunkt neue Untersu-

¹ Die Abkürzung **pk** steht für primary key, **fk** für foreign key.

son auf die Daten zugreifen darf.

Im Anschluß an die Schema-Generierung werden die physischen Tabellen definiert und den zu erwartenden Anfragen und Aufgaben entsprechend Indexe und benutzerspezifische Sichten festgelegt. Gleichzeitig werden für die Dateneingabe sowie zur Umsetzung von Integritätsbedingungen Prozeduren und Trigger geschrieben. Zum Abschluß der Implementationsphase wird anhand von prototypischen Eingabedaten die Funktionalität der Datenbank überprüft und gegebenenfalls modifiziert.

4.2 Graphische Benutzungsschnittstellen

Ein Anwender hat meist weder die Möglichkeit noch den Willen viel Einarbeitungszeit in die Benutzung einer Datenbankanwendung zu investieren. Um daher die Arbeit mit einer Datenbank so einfach wie möglich zu gestalten, bieten sich graphische, intuitiv bedienbare Benutzungsschnittstellen an. Die Oberfläche muß eine effiziente und umfassende Arbeit mit der Datenbank ermöglichen. Dies setzt vor allem eine integrierte Gestaltung von Datenbank und Oberfläche voraus. Schon allein aus Navigationsgründen müssen Oberflächenkonzept und Datenbankstruktur einander angepaßt sein. Der Übersichtlichkeit und einfachen Benutzbarkeit halber sollte den Fenstern eine klare und einheitliche Gestaltung zugrunde liegen. Als Richtlinie wurde ein Styleguide entwickelt, in dem die Größe der verschiedenen Fenster- und Objekttypen, ihre Anordnung und, in Bezug auf Nachnutzung und Modifikation der Datenbank, die Namensgebung festgelegt ist.

In Anlehnung an die Datenbankstruktur und zur Vereinfachung der Navigation auf den Oberflächen sind die einzelnen Fenster nach Themen / Schlagwörtern: *Boden, Wasser, Klima, Flora und Fauna* sowie *übergeordnete Informationen* und Funktionen: *Suchen, Einfügen, Ändern, Löschen* und *Hilfe* geordnet. Als Navigationshilfe dient das Einstiegsfenster aus Abbildung 6. Es bleibt während der gesamten Datenbanksitzung geöffnet und erlaubt daher auch multiple Datenbankdialoge, z.B. gleichzeitiges Eintragen und Suchen auf verschiedenen Datensätzen.

Die Aktivierung der verschiedenen Funktionen im Einstiegsmenü sowie in den nachfolgenden Fenstern wird durch Benutzerereignisse, Anklicken von Radio- und Controlbuttons, gesteuert. Damit kann der Anwender nur noch unter den aktuell für ihn verfügbaren Funktionen wählen, so daß bei einer abgestimmten und konsistenten Oberflächen / Datenbankkonzeption vom Prinzip her nur wohl definierte Dialogsituationen durchlaufen werden können. Insgesamt wurden die Dialoge so gestaltet, daß Anfragen zügig, mit wenig dazwischen geschalteten Fenstern, durchlaufen werden.

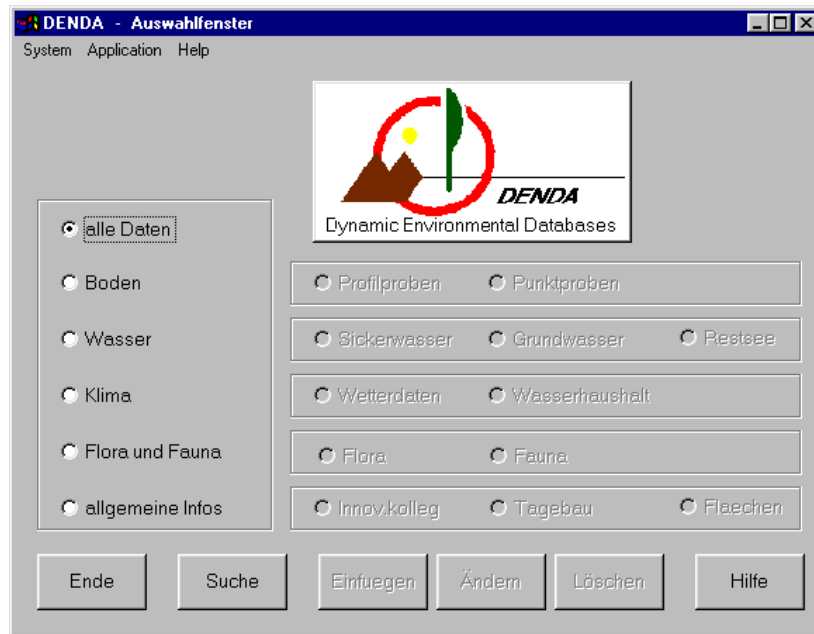


Abbildung 6
Das Einstiegsfenster der DENDA-Datenbank

Die Oberflächen für die Eingabe, Änderung sowie das Löschen von Daten wurden aus Gründen der Benutzerunterstützung in Anlehnung an die von den Erfassern eingereichten Beispieldateien (i.a. Excel-Dateien) entwickelt. Die Daten werden vom Benutzer entweder direkt eingetippt oder, insbesondere bei häufig aufgenommenen Daten, aus einem Eingabefile eingelesen. Funktionalität, Aussehen und Gliederung der einzelnen Fenster wurden in Zusammenarbeit mit den Benutzern festgelegt.

Im Fall der Datensuche und Auswertung basieren die Fenster auf Aufgaben- und Benutzerspezifischen Sichten, in denen die Leserechte des Anfragenden überprüft und die aus Modellierungsgründen über mehrere Tabellen verteilten Informationen wieder zusammengefaßt werden. Auch bei der Ausarbeitung von Suchfenstern müssen die Benutzer in die Entwicklung mit einbezogen werden, da sie letztendlich angeben müssen, welche Datendarstellungen und Zusammenstellungen einen Sinn ergeben und welche Darstellungen sie zur Auswertung benötigen.

5. Diskussion und Schlußfolgerungen

Ein wichtiger Aspekt unserer Entwicklungsarbeiten ist die Adaptierbarkeit der Datenbankanwendung, besonders in Hinblick auf neu hinzukommende Teilprojekte und Aufgabenstellungen. Dabei ist es nicht die Erweiterung des ER-Schemas um neue Teilbereiche, die diese Nachnutzungen aufwendig gestaltet, sondern vor allem die Integration der hinzutretenden Sichten und Oberflächen in ein konsistentes Gesamtschema.

Auch der Entwicklung von Frames zur Darstellung analoger Sichten sowie der schnellen Generierung von Sicht und Oberfläche kommt eine große Bedeutung zu. Den Vorteil, den der einfache Datenbankzugriff über graphische Benutzungsschnittstellen gegenüber Tools wie der interaktiven SQL-Schnittstelle *isql* bietet, hat natürlich auch den Nachteil fehlender Spontaneität und Dynamik zur Folge. Will der Benutzer beispielsweise die chemischen Eigenschaften von Boden und Sickerwasser miteinander vergleichen, so ist er darauf angewiesen, daß eine, die beiden Themen verknüpfende, Sicht und das darauf zugreifende Fenster schon vordefiniert wurden.

Insgesamt hat sich gezeigt, daß der Datenbankentwicklungsprozeß durch eine enge Zusammenarbeit zwischen Benutzern, Datenbank- und Oberflächenentwicklern unterstützt werden kann und besonders in Bezug auf die Anwendungsentwicklung neue Konzepte für einen integrativen Entwicklungsansatz (B. Thalheim 1997) notwendig sind, um komplexen Anfragesituationen gerecht zu werden.

Literatur

- Chen, P. (1976): The Entity-Relationship-Model – Toward a Unified View of Data, in: ACM Transactions on Databasesystems, Bd. 1, Nr. 1, S. 9-36
- Kremers, H., Krasemann, H. L. (1996): Umweltdaten verstehen durch Metainformationen, Metropolis Verlag
- Schmidt, J., Bräunig, A. (1994): Fachinformationssystem Boden und Geologie, Erfassungsvorschrift Flächendaten / Kippengutachten, Sächsisches Landesamt für Umwelt und Geologie
- Thalheim, B. (1993): Fundamentals of Entity-Relationship-Modelling, in: Annals of Mathematics and Artificial Intelligence, Vol. 7, S. 197-256
- Thalheim, B. (1997): Der Abstraktionsschichtenansatz für den Datenbankentwurf – Codesign von Struktur, Sichten, Prozessen und Dialogen, in: BTU Cottbus, Informatik-Bericht, I-05,
- Vinek, G., Rennert, P. F., Tjoa, A. M. (1982): Datenmodellierung, Theorie und Praxis des Datenbankentwurfs, Physica-Verlag
- Zamperoni, D., Behlig, G., Wenderoth, C. (1996): Dynamische Metadatenverwaltung von Umweltdaten, in: Umwelt-Informatik aktuell, Bd. 10, S. 79-88