

Das integrierte Datenmodell des Hessischen Erdinformationssystems HEISS

Willi Classen, Klaus Friedrich, Achim Stärk und Dominik Ulmer¹

Abstract

The Hessian Soil Information System HEISS is developed at the Hessian Geological Survey with the target to administer the large quantity of its geoscientific data in a system uniformly. For the implementation of the system a general object-relational data model for the central administration of the data was developed. For all areas of the system development client-server Intranetapplications were created, with whose assistance the data modelling can be made. Based on the central data model there are applications available to edit and search the geoscientific data as well. As a first step information systems for the areas of geology and soil science and soil protection are in development at present.

1 Einleitung

Das Hessische Landesamt für Bodenforschung (HLfB) ist die Fachbehörde für die geowissenschaftliche Landesaufnahme im Land Hessen. Zentrale Aufgaben sind die Erfassung, Erarbeitung und Bereitstellung geologischer, bodenkundlicher, hydrogeologischer, rohstoffgeologischer u.a. Daten und Grundlagen für die unterschiedlichen Bereiche der Geologie, die Raum- und Landesplanung sowie für den Umweltschutz. Eingebunden in die geologische und bodenkundliche Landeserkundung erhebt und verwaltet es eine große, ständig wachsende Anzahl von Bohr- und Bodenproben. Die Erstellung und Veröffentlichung von geowissenschaftlichen Kartenwerken und ihren Erläuterungen ist eines der anschaulichen Ergebnisse dieser Arbeiten.

Zur Bewältigung der vielfältigen Aufgaben nehmen auch am HLfB in immer stärker zunehmendem Maße EDV-basierte Informationssysteme zentrale Stellungen ein. Mit sich stetig verringernden personellen Ressourcen bei gleichzeitig weiter anschwellenden Datenmengen und rasch wachsenden Informationsflüssen zeigen moderne Informationssysteme Wege auf, komplexe Fragestellungen schnell, fundiert und umfassend zu lösen. Aus diesem Grund entwickelt das HLfB ein integriertes Datenverarbeitungs- und Informationssystem. Das Hessische Erdinformationssystem HEISS vereint die Aktivitäten der geowissenschaftlichen Fachbereiche des HLfB

¹ Hessisches Landesamt für Bodenforschung, Leberberg 9, D - 65193 Wiesbaden; w.classen@hlfb.de

von der Datenerfassung, über die Datenanalyse, Datenverwaltung bis zur Datenausgabe.

2 Zielsetzung

HEISS ist als langfristige und umfassende Lösung für die EDV-basierte Verwaltung der geowissenschaftlichen Daten, z.B. über Bohrungen, Aufschlüsse, Einzelfunde und Schichten, am HLfB in Entwicklung. Die Verwaltung der Daten ist zentral zu organisieren, wogegen die Zugriffe auf die Daten als Recherchen dezentral erfolgen. Der Betrieb des Systems, vor allem die Recherche auf den Daten, soll ohne spezielle systeminterne Kenntnisse vorgenommen werden können. Daten und Programmstrukturen müssen flexibel erweiterbar sein. Änderungen und Erweiterungen im Datenmodell dürfen sich nicht behindernd auf die Anwendungen auswirken. Die langfristige Nutzung muß durch die Wahl geeigneter Hard- und Softwarekomponenten gewährleistet werden. Eine plattformunabhängige Realisierung als Client-Server- und Browsertechnologie stellt die Voraussetzungen hierfür zur Verfügung. Die Entwicklung als Browseranwendung gewährleistet gleichzeitig dem Anwender eine bekannte Anwendungsumgebung.

3 Konzeption

Aufgrund der vielseitigen thematischen und aufgabenspezifischen Anforderungen ist ein modularer Aufbau zweckmäßig. Neben den unterschiedlichen Systemkomponenten zur Erfassung, Verarbeitung, Analyse und Ausgabe der Daten sind inhaltlich unterschiedliche Sachgebiete in eigenen Fachinformationssystemen (FIS) abzulegen. Das Gesamtsystem ist als eine zentral verwaltete Datenbankanwendung in einer Client-Server-Architektur konzipiert. Für die Datenmodellierung ist ein objekt-relationaler Ansatz gewählt worden. Die in Objekten beschriebenen Inhalte werden in einem relationalen Datenbanksystem zentral auf einem Server verwaltet. Über ein Netzwerk greifen Client- und Server-Anwendungen auf die Daten zu (Abbildung 1). Für die Entwicklung und Implementierung stehen ein Oracle Server-Datenbanksystem und ein Oracle WebServer zur Verfügung. Bei der Anwendungsentwicklung kommen Oracle Forms, PL/SQL und Java zum Einsatz.

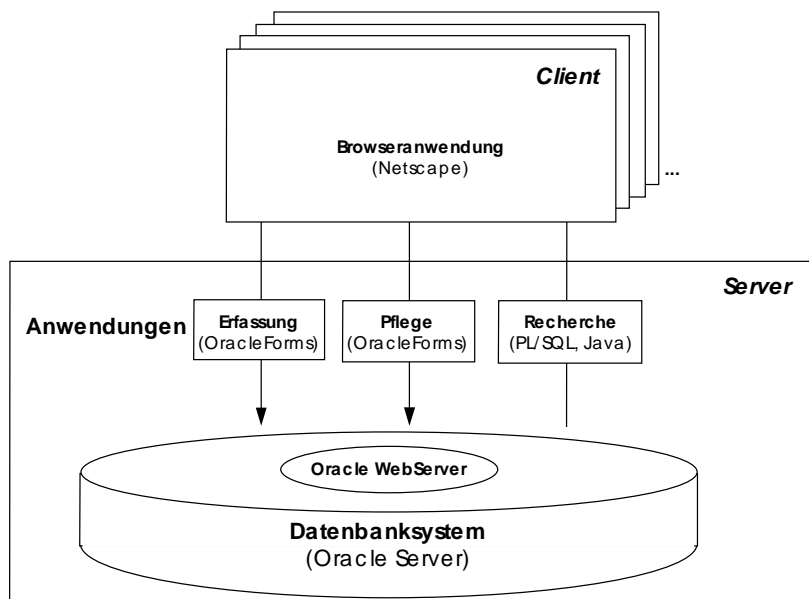


Abbildung 1: Die Systemarchitektur

4 Systemarchitektur

Die Datenstrukturen nehmen die zentrale Stellung in der Systemarchitektur ein. Auf ihnen beruht das gesamte System einschließlich der Anwendungsentwicklung und Ablaufsteuerung.

4.1 Objektrelationale Datenstrukturen

Für die Datenverwaltung ist ein objektrelationaler Ansatz gewählt worden. Die abgelegten Daten sind als Objekte definiert. Die Objekte und die Beziehungen sind in Relationen abgelegt. In einer Definitionsebene werden die Objekttypen und das allgemeine Datenmodell definiert. In der Datenebene werden die Sachdaten und deren Beziehungen verwaltet. Über Fremdschlüsselverknüpfungen und Verkettungsinformationen werden die komplexen Zusammenhänge dynamisch und über die Hierarchiestufen hinweg verarbeitet. Es kann in verschiedene Datenbankbereiche

unterschieden werden: strukturelles Datenmodell, struktureller Thesaurus und der Thesaurus (Abbildung 2).

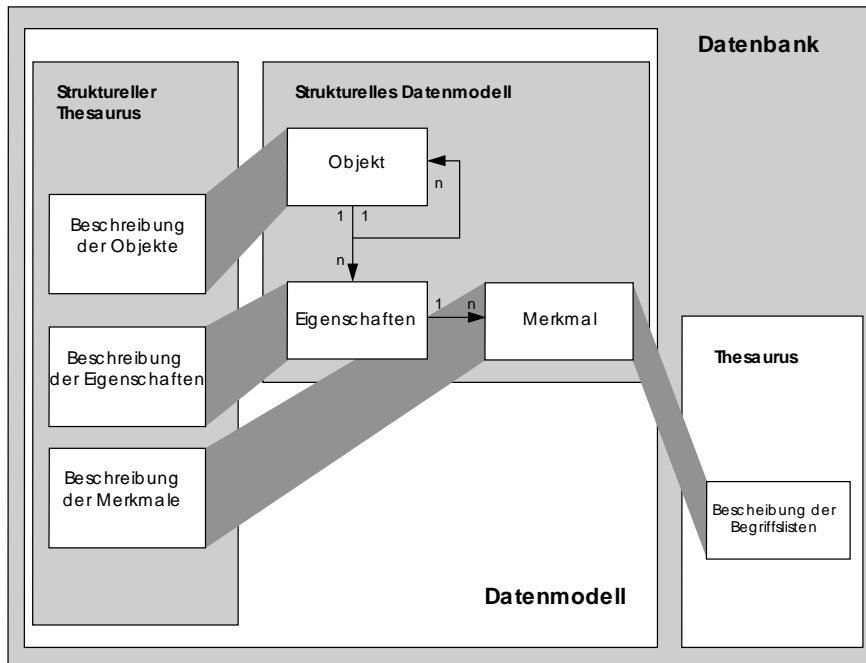


Abbildung 2: Das Datenbankmodell

4.2 Das strukturelle Datenmodell

Die grundsätzlichen Datenstrukturelemente sind allgemein und mit nur sehr wenigen Objekttypen definiert. Es wird unterschieden in Objekte, Eigenschaften und Merkmale. Das Beziehungsgefüge dieser Elemente ist klar definiert. Ein Objekt kann Vater eines oder mehrerer Objekte sein. Ein Objekt kann eine oder mehrere Eigenschaften haben. Eine Eigenschaft wird durch ein oder mehrere Merkmale beschrieben. Alle Beziehungen können in 1:n Ausprägung vorliegen. Durch die Reduzierung auf 3 Objekttypen ist es einerseits sehr allgemein gültig, andererseits erfordert es einen hohen Abstraktionsgrad bei der Datenmodellierung.

4.3 Der strukturelle Thesaurus

Im strukturellen Thesaurus werden alle abgebildeten geowissenschaftlichen Objekte beschrieben. In ihm ist somit das Abbild der realen Welt als Objekthierarchie beschrieben. Er repräsentiert die Datenebene des Systems.

4.4 Der Thesaurus

Der Thesaurus beschreibt die definierte, hierarchisch strukturierte Begriffswelt des Systems. Integriert in das Datenmodell liegen hier umfangreiche Begriffe in Begriffslisten als Objekte vor.

4.5 Implementierung

Das Informationssystem ist modular aufgebaut. Als zentrales Datenbanksystem dient ein OracleServer. Die Anwendungen zur Erstellung und Pflege eines Fachdatenmodelles sowie für die Erfassung der Geodaten sind zunächst als Oracle Forms Windowsanwendung realisiert worden. Die Recherche auf die Daten wurde mit Java und PL/SQL direkt als Serveranwendung umgesetzt. Über Browseranwendungen erfolgt der Zugriff von beliebigen Clientarbeitsplätzen. Die Recherche ist damit von Anfang an plattformunabhängig einer breiten Anwendergruppe verfügbar. Die Pflege und Erfassung der Daten unterliegt dagegen noch einigen clientseitigen Beschränkungen. Da diese Anwendungen von deutlich weniger Anwendern bedient werden, sind diese Einschränkungen akzeptabel.

Die fachlichen Konzepte und Strukturen sind in breit angelegten Diskussionen am HLFb entwickelt worden. Die Implementierung der Datenstrukturen und die Anwendung zur Erfassung der Daten ist ebenfalls vom HLFb geleistet worden. Die Pflege- und Rechercheanwendung wurde vom Institut für Geoinformatik der Universität Münster entwickelt und programmiert.

5 Fachinformationssysteme

5.1 Fachdatenmodelle

Das beschriebene Datenmodell ist die Grundlage für die Entwicklung unterschiedlicher Fachdatenmodelle, die jeweils einen abgeschlossenen Ausschnitt der realen Welt beschreiben. Verschiedene Fachdatenmodelle stehen nebeneinander oder auch sich gegenseitig ergänzend zueinander. Ein Fachdatenmodell wird durch die definierten Beziehungen der Objekte eindeutig charakterisiert. Ein Objekt liegt zunächst unabhängig und selbständig vor. Durch die Definition von Beziehungen zu

anderen Objekten wird es in ein Fachdatenmodell und damit in einen fachlichen Kontext integriert. Ein Objekt kann gleichzeitig mehreren Fachdatenmodelle angehören bzw. innerhalb eines Modells mehrfach in Erscheinung treten. Es liegt physikalisch aber immer nur einmal vor.

Diese objektorientierte Datenverarbeitung gewährleistet fachlich eine einheitliche Begrifflichkeit und Datenstruktur. Hieraus ergibt sich auch die Möglichkeit einer übergreifenden und vergleichbaren Mehrfachnutzung innerhalb und zwischen den Fachdatenmodellen. Geowissenschaftliche Objekte werden auf diese Art und Weise systemweit klar strukturiert und eindeutig definiert.

Durch die Definition von entsprechenden Objekten und deren Beziehungen lassen sich unter Zugriff auf bereits bestehende Fachdatenmodelle und ihrer Objekte neue Fachdatenmodelle schnell und in einheitlicher Form entwickeln. Ein entscheidender Vorteil wird bei der Entwicklung neuer Fachdatenmodelle und den Anwendungen deutlich. Neu- und Weiterentwicklungen haben direkten Zugriff auf die verschiedenen Entwicklungs- und Anwendungswerkzeuge, da sie auf die definierten Objektstrukturen zugreifen. Die Anwendungen können diese selbständig verfolgen, da sie im Datenbanksystem abgelegt sind. Spezielle Anwendungsentwicklungen und -programmierungen sind nicht erforderlich bzw. können auf ein Minimum reduziert werden.

5.2 Implementierung

Am Anfang der Entwicklung eines FIS steht die Erstellung eines logischen Datenmodells mit der Definition der Objekte und ihrer Beziehungen. Die Implementierung dieses Datenmodells in objektrelationale Strukturen des Datenbanksystems erfolgt direkt mit Hilfe der Pflegeanwendung. Die Objekte werden mit ihren Eigenschaften, Merkmalen und Beziehungen dem Datenmodell entsprechend definiert. Die Pflegeanwendung generiert hieraus die relationalen Datenstrukturen. Bereits existierende Objekte können problemlos integriert werden. Alle Objekteigenschaften, auch programmatische Eigenschaften, stehen den Anwendungen sofort zur Verfügung.

5.3 Fachinformationssystem Bodenkunde

Ein FIS hat ein in sich abgeschlossenes Datenmodell. Die Objektmenge eines FIS ist eine Teilmenge aller definierten Objekte des Gesamtsystems. Am HLfB befinden sich derzeit mehrere Fachdatenmodelle in der Entwicklung. Die Systeme für die Geologie, Bodenkunde und Geotoptypen sind als erste in ihrer Entwicklung bereits weit fortgeschritten. Als Beispiel werden einige Bereiche des FIS Bodenkunde nachfolgend näher erläutert.

Grundlegend ist anzumerken, daß gegenwärtig alle Daten als Punktdaten verwaltet werden, d.h. der Ursprung aller Daten im Informationssystem liegt immer in einem Punktobjekt. Jeder Punkt hat beschreibende Merkmale. Alle weiteren Objekte stammen direkt oder indirekt über mehrere Hierarchiestufen von einem Punktobjekt ab. Die Entwicklung eines FIS geht einher mit einer wachsenden Anzahl von Objekten. Die Objekte werden neu definiert oder es werden bereits bestehende Objekte verwendet.

5.3.1 Fachdatenmodell Boden/Bodenschutz

Das FIS Bodenkunde kann unterteilt werden in die Bereiche Bodenprofile, Bodenschätzung und Bodendauerbeobachtungsfläche. Abbildung 3 zeigt einen Ausschnitt aus dem Objektdiagramm des Datenmodells Boden/Bodenschutz.

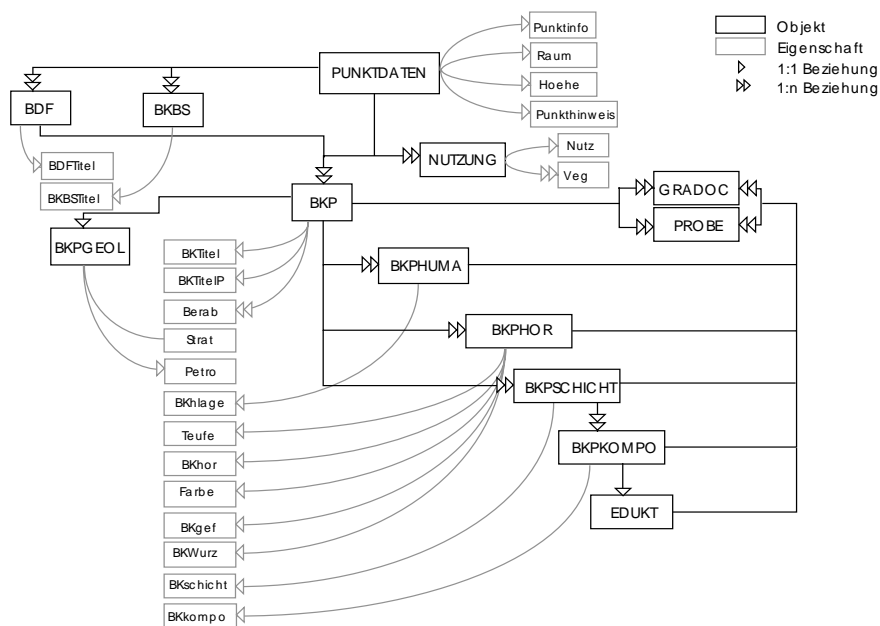


Abbildung 3: Das Fachdatenmodell Boden/Bodenschutz (Ausschnitt)

Als oberste Hierarchie ist der Raumbezug als Punkt definiert. Dieser besitzt 4 Eigenschaften, in denen beschreibende Merkmale gruppiert werden. Dieser Objekttyp mit seinen Eigenschaften ist allen Bereichen gemeinsam. Die erste Abgrenzung der FIS'e untereinander erfolgt in der 2. Hierarchieebene. Hier sind die Bereiche Bodendauerbeobachtungsfläche (Objekt BDF), Bodenschätzung (BKBS) und bodenkund-

Der beschriebene Auszug des FIS Bodenkunde umfaßt nur einen Teilbereich des Informationssystems HEISS. Verknüpfungen bestehen durch gemeinsam genutzte Objekte und Eigenschaften. Je größer die Anzahl gemeinsam genutzter Objekte ist, desto ähnlicher sind sich die FIS'e. Als gemeinsames Minimum steht immer ein Punktobjekt als Raumbezug aller Daten an oberster Stelle. Damit wird nochmals deutlich, daß ein FIS nur einen Ausschnitt der im System dargestellten 'Welt' wiedergibt (Abbildung 4).

5.5 Anwendung des FIS Bodenkunde

Mit der Implementierung des logischen Datenmodells in der Pflegeanwendung ist bereits ein Großteil der notwendigen Arbeiten für die Anwendung im HEISS geleistet. Nach der Erledigung einiger notwendiger Anpassungen in der Erfassungsanwendung kann mit der Dateneingabe begonnen werden. Diese Anpassungen betreffen die Benutzeroberflächenelemente, die nicht programmatisch anhand der definierten Datenstrukturen erzeugt werden können. Nach erfolgter Dateneingabe stehen die Daten über die Rechercheanwendung ohne weitere Anpassungen in vollem Umfang zur Verfügung.

Über die Recherche kann in strukturierter Form, den definierten Objekten entsprechend, auf die Daten zugegriffen werden. Sie erlaubt eine vollständige Suche über alle Hierarchiestufen des Systems. Je nach Fragestellung kann sowohl die Objekt- als auch die Merkmalsmenge beliebig definiert werden. Recherchen können als spezifische Anfragen abgespeichert werden. Sie lassen sich somit mehrfach auch auf unterschiedlichen Datengrundlagen ausführen. Recherchen können ebenfalls miteinander verknüpft werden. Standardisierte Recherchen können so spezifischen Anforderungen entsprechend aus verschiedenen Teilrecherchen zusammengestellt werden.

6 Bewertung

Die Datenverarbeitung am HLfB wird durch das System HEISS in wesentlichen Teilen optimiert. Mit diesem System steht eine ganze Anwendungsgruppe zur Verfügung, die geowissenschaftliche Daten in einheitlichen Datenstrukturen integriert und in angepaßten Anwendungen zugänglich macht. Dem Konzept der einfachen allgemeinen Objektstrukturen steht die Einpassung aller Fachdaten in diese Strukturen zunächst entgegen. Alle Objekte müssen zuerst entsprechend formalisiert bzw. allgemein definiert werden. Der Aufbau des logischen Datenmodells muß auf der Grundlage der fachlichen Abstimmung erfolgen. Die Definition und Abgrenzung der geowissenschaftlichen Objekte ist die Hauptaufgabe der Entwicklung eines Fachinformationssystems. Die einzelnen Anwendungen ermöglichen auf dieser Grundlage

eine rasche Implementierung und Nutzung, weitestgehend ohne spezifische interne Datenmodell- und Programmierkenntnisse.

Aus geowissenschaftlicher Sicht ist hervorzuheben, daß durch die Objektdefinitionen gleichzeitig geowissenschaftliche Themen übergreifend klar strukturiert und im System standardisiert definiert werden können. Daten können übergreifend genutzt werden, ohne daß Redundanzen auftreten.

7 Ausblick

Das System HEISS stellt in der gegenwärtigen Entwicklungsstufe noch nicht die Endausbaustufe dar. Mit dem in diesem Jahr (1998) anlaufenden produktiven Betrieb des Systems wird die Menge der definierten Objekte und gespeicherten geowissenschaftlichen Daten rasch ansteigen. Hieraus ergeben sich weitergehende Rechercheanforderungen, um dem gewachsenen Datenbestand mit effektiven Recherchewerkzeugen gerecht zu werden. Geplant sind insbesondere der Ausbau der Recherche zu einem stärker intuitiv fachlich und weniger objektgesteuerten Recherchesystem, in dem die Anfragen eher als Klartext als in Objektlisten zusammengestellt werden. Die Datenausgabe erfolgt derzeit in standardisierter, nach Objekten hierarchisierter Form. Hier ist eine Erweiterung hin zu standardisierten Fachanwendungen notwendig, um z.B. Berichte und Dokumentationen direkt in entsprechenden Formaten auszugeben. Neben diesen technischen Aspekten steht weiterhin die Einbeziehung von Flächendaten mit entsprechenden Grafikkomponenten und Geoinformationssystemen an. Neben der bisherigen inhaltlichen Recherche soll eine räumliche Recherche realisiert werden. Auf der Seite der Datenerfassung wird weiterhin über eine Schnittstelle und entsprechende Anwendung nachgedacht, die externen Stellen erlaubt, Daten direkt in den geforderten Strukturen zu erheben und den Datenfluß in das zentrale System zu vereinfachen.