

DK 528.9:681.3.04
519.256
528.063.9

Einführung in das Projekt „Entwurf einer Benutzerschnittstelle für eine graphisch-alphanumerische Datenbank“

(mit 3 Bildern)

Von Isa Ramm und Hans-Dieter Ehrich, Braunschweig

ZUSAMMENFASSUNG: Dieses Projekt ist ein Teil des DFG-Schwerpunkt-Projektes „Digitale geowissenschaftliche Kartenwerke“. Letzteres befaßt sich unter anderem mit der computergestützten Speicherung und Auswertung geowissenschaftlicher Daten.

Für diesen Zweck wurden bisher spezielle autonome Anwendungsprogramme und -programmsysteme erstellt, die jeweils auf einer individuell programmierten Dateiverwaltung aufbauen. Bei der Entwicklung dieser Systeme ist der Anwendungsprogrammierer neben seiner eigentlichen Aufgabe verantwortlich für die Strukturierung der Dateien, die Sicherung der Daten usw. Solche Verwaltungsaufgaben werden ihm beim Einsatz eines Datenbanksystems weitestgehend abgenommen.

Wir zeigen die Vorteile beim Einsatz eines Datenbanksystems gegenüber der individuellen Dateiverwaltung auf und skizzieren die Vorgehensweise für den Entwurf einer Datenbank, die die gewünschten Anwendungen auf den Daten ermöglichen soll.

SUMMARY: This project is part of the DFG-project "Digital geo-scientific maps". The latter deals with computer based storing and evaluating of geo-scientific data.

For this purpose, specifically tailored application programs and systems based on individual file management are used up to now. However, there are some disadvantages of application systems doing their own file management which are the reason why a database system should be used instead.

We show the superiority of a database system in comparison to the application dependent file management and give some hints how to design the database to cover the intended applications on the data.

RÉSUMÉ: Ce projet fait partie intégrante du projet de base «Digitale geowissenschaftliche Kartenwerke» (Séries de cartes Numériques Géoscientifiques) de la DFG, lequel s'est donné comme objectif, entre autres, la mémorisation et l'évaluation assistées par ordinateur de données géoscientifiques.

A cette fin, des programmes d'évaluation ainsi que des systèmes de programme spéciaux à conception autonome ont été mis au point jusqu'à présent, dont chacun est basé sur une gestion de fichiers programmée d'une manière individuelle. Néanmoins, il existe quelques désavantages en ce qui concerne les systèmes d'application effectuant leur propre gestion de fichiers; c'est pourquoi on devrait accorder la préférence à des systèmes de banque de données.

Les avantages de l'emploi d'un système de banque de données par rapport à une gestion de fichiers individuelle sont démontrés et les opérations à exécuter en élaborant le projet d'une banque de données, susceptible de répondre aux applications exigées, sont esquissées.

1 Einleitung

Im DFG-Schwerpunkt-Projekt „Digitale geowissenschaftliche Kartenwerke“ sind vielfältige Forschungsvorhaben zusammengefaßt, die unter anderem das Ziel haben, geowissenschaftliche Daten computergestützt zu speichern, zu manipulieren, auszuwerten und zur Ausgabe aufzubereiten.

Einige geowissenschaftliche Forschungsvorhaben existieren schon seit längerem. Hieraus resultieren viele verschiedene graphische und alphanumerische Programme und Programmsysteme, die für spezielle Anwendungen konzipiert wurden.

Nachdem im geowissenschaftlichen Bereich somit einige Erfahrungen mit der computergestützten Datenhaltung und -verarbeitung gesammelt wurden, ist es an der Zeit, die Dateiverwaltung, auf der die Programme und Systeme basieren, zu vereinheitlichen, um die Kommunikation zwischen verschiedenen Programmen zu ermöglichen. An dieser Stelle wurde der Einsatz eines Datenbanksystems erwogen.

In dem DFG-SPP sind unter dem Vorhaben „Graphische Datenbank“ mehrere Projekte zusammengefaßt, die das Ziel haben, eine speziell für geowissenschaftliche Belange konzipierte Datenbank aufzubauen. Eines dieser Projekte „Entwurf einer Benutzerschnittstelle für eine graphisch-alphanumerische Datenbank“, im folgenden kurz „Geo-Datenbank-Entwurf“ genannt, wird in unserer Gruppe bearbeitet.

In den nächsten Abschnitten wird erläutert, warum der Einsatz eines Datenbanksystems sinnvoll erscheint, wie das Vorhaben „Graphische Datenbank“ gegliedert ist, und es wird das Arbeitsprogramm zum „Geo-Datenbank-Entwurf“ skizziert.

2 Vorteile des Datenbankeinsatzes

Ein Datenbanksystem bietet entscheidende Vorteile gegenüber der Dateiverwaltung in Anwendungsprogrammen und -systemen. Um diese aufzeigen zu können, werden zuerst die Anforderungen des Endbenutzers an ein Anwendungssystem den Diensten, die ein Anwendungssystem auf Dateibasis tatsächlich bieten kann, einander gegenübergestellt. [HoWh 83]

Der Endbenutzer stellt die folgenden Anforderungen:

1. Daten können gespeichert, bearbeitet, wiedergewonnen werden, ohne daß sich der Endbenutzer um Datenstrukturen, Datenverwaltung und Programmierung zu kümmern braucht.
2. Das System soll automatisch für Datenintegrität (Widerspruchsfreiheit der eingegebenen Daten) und Datensicherheit (Schutz vor unberechtigtem Zugriff) sorgen.
3. Das System soll an den vorhandenen Rechner angepaßt sein.

4. Das System soll portabel und hardwareunabhängig sein (für Erweiterung der Hardware, Betrieb auf anderen Rechnern).
5. Die Leistungsfähigkeit oder Effizienz in bezug auf die Speicherung und Verarbeitung der Daten soll möglichst groß sein.
6. Das System muß in dem Sinn erweiterbar sein, daß neue, zusätzliche oder geänderte Anwendungen auf bestehenden Daten ausgeführt bzw. neue Datenfelder problemlos hinzugefügt werden können.

Der Endbenutzer wünscht, daß alle diese Anforderungen so kostengünstig wie möglich erfüllt werden.

Die Anwendungssysteme werden von Programmierern so erstellt, daß an erster Stelle die Entwicklung von funktionsfähigen Berechnungsprogrammen zur Manipulation und Aufbereitung der Daten steht. Dies ist das eigentliche Anliegen des Anwenders und Programmierers, durch das der Einsatz eines Computers erst gerechtfertigt wurde.

Hinzu kommen komfortable Ein-/Ausgaberoutinen, die dem Endbenutzer eine möglichst komfortable Kommunikation mit dem System bieten sollen.

Weiterhin muß sich jedoch der Programmierer, oder in einigen Fällen sogar der Endbenutzer, noch direkt um die Daten- und Dateiverwaltung kümmern, d. h. Speicherplatzberechnung, Mischen und Sortieren von Dateien, Speicherplatzfreigabe bei Datenlöschung, Platten-Ein-/Ausgabesteuerung, Verwaltung von Pointern und Indizes für Aufbau und Pflege von Zugriffspfaden, usw.

Der Aufbau eines Anwendungssystems auf Dateibasis ist in Bild 1 dargestellt.

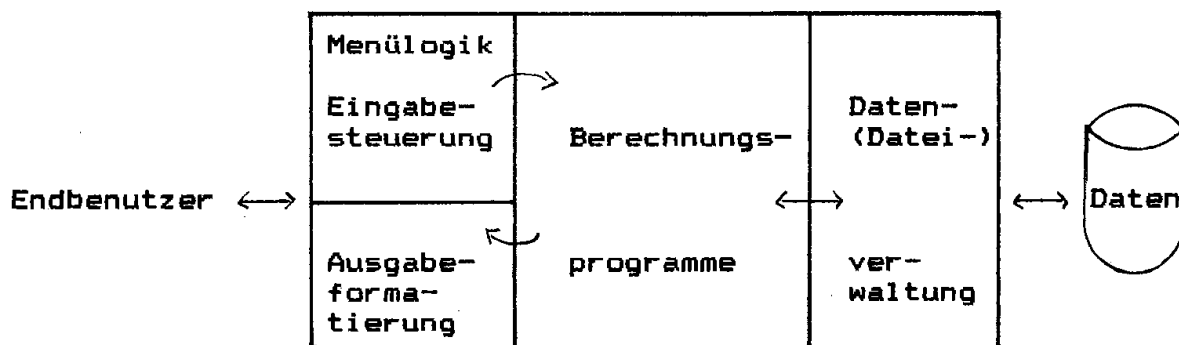


Bild 1 — Aufbau eines Anwendungssystems auf Dateibasis

Durch Anwendungssysteme auf Dateibasis werden die Benutzeranforderungen jedoch nur im folgenden Umfang erfüllt:

1. Es wird nicht immer gewährleistet, daß der Endbenutzer keine Kenntnisse über die physische Speicherung der Daten, Programmierkenntnisse bzw. Kenntnisse über den Aufbau des Anwendersystems benötigt.
2. Mechanismen für die Datensicherheits- und -integritätsüberwachung sind gar nicht oder kaum vorhanden. Wenn ja, werden diese in jedem Anwendungssystem erneut programmiert.

3. Die Anwendungssysteme selbst sind in starkem Maße von der physischen Speicherung der Daten abhängig. Daraus folgt zwar eine gute Anpassung an den Rechner, für den das System entwickelt wurde, eine Änderung der Hardware würde jedoch eine Anpassung aller Anwendungssysteme zur Folge haben müssen (z. B. Speichermedium Magnetband wird ersetzt durch Platte).
4. Aus Punkt 3 folgt in diesem Fall direkt, daß solch ein Anwendungssystem nicht unbedingt portabel, das heißt auf anderen Rechnern ohne Änderung einsetzbar, ist.
5. Wegen der Hardware-Nähe der Anwendungssysteme lassen sich hier die jeweiligen Rechner-spezifischen Gegebenheiten ausnutzen und spezielle, für eine Problemlösung optimale, Datenstrukturen verwenden. Daraus folgt unter Umständen eine hohe Verarbeitungsgeschwindigkeit. In vielen Programmen werden jedoch dieselben Daten in unterschiedlichem Dateiformat bzw. Teilmengen derselben Daten benötigt. Hierzu sind jeweils Umsetzungen nötig. Das bedeutet zusätzlichen Verarbeitungsaufwand, oder die Daten werden mehrfach gespeichert, was zu hoher Datenredundanz und somit Speicherplatzverschwendung führt.
6. Für die Entwicklung neuer, zusätzlicher oder geänderter Anwendungen auf bestehenden Dateien ist eine genaue Kenntnis über die Zugriffspfade, das jeweilige Dateiformat und den Dateiinhalt notwendig. Für diese Anwendungen sind die bestehenden Zugriffspfade nicht unbedingt effizient, sie können jedoch nur unter großem Aufwand geändert werden.

Jede Änderung der Dateiformate oder Zugriffspfade erzwingt die Änderung sämtlicher darauf operierender Programme. Die Dateiformate sind für bestimmte Anwendungen und nicht für leichte Änderbarkeit konzipiert, daher ist es oft nicht möglich, nachträglich neue Datenfelder hinzuzufügen.

Mit diesen Anwendungssystemen kann also nur ein Teil der Anforderungen des Benutzers erfüllt werden. Diese Dienste sind sogar recht kostenintensiv, da sie in jedem Anwendungssystem neu geplant und programmiert werden müssen.

Ein Datenbanksystem hingegen ist als Werkzeug für die Entwicklung von Anwendungssystemen zu sehen, das viele der Dienste abnimmt bzw. überhaupt erst ermöglicht. Das Datenbanksystem wird (evtl. für jeden Rechner-typ) nur einmal entwickelt und leistet gleichzeitig mehr. Die Dateiverwaltungs- und übergeordneten Betriebsaufgaben werden vom Datenbanksystem erledigt, der Anwendungsprogrammierer kann sich damit auf die für ihn wesentlichen Aufgaben der Erstellung von Datenaufbereitungs- und Auswertungsprogrammen konzentrieren und die Manipulation der Daten auf der logischen Ebene ausführen. Die Sicht des Endbenutzers bleibt erhalten oder wird dadurch vereinfacht, daß dieser sich nicht mehr um die physische Speicherung der Daten kümmern muß. Der Aufbau eines Anwendungssystems auf der Basis eines Datenbanksystems wird in Bild 2 dargestellt. Unter einem Datenbanksystem versteht man die Einheit aus Datenbankmanagementsystem (DBMS) und dem Datenbestand.

Durch den Einsatz eines Datenbanksystems wird eine physische Datenunabhängigkeit erreicht, die die Anwendungsprogramme immun gegenüber Änderungen der Datenstruktur, Speicheraufteilung und verschiedenen Zugriffsstrategien macht. Daraus folgt umgekehrt, daß Anwendungsprogramme leicht verändert, erweitert und neu hinzugefügt werden können.

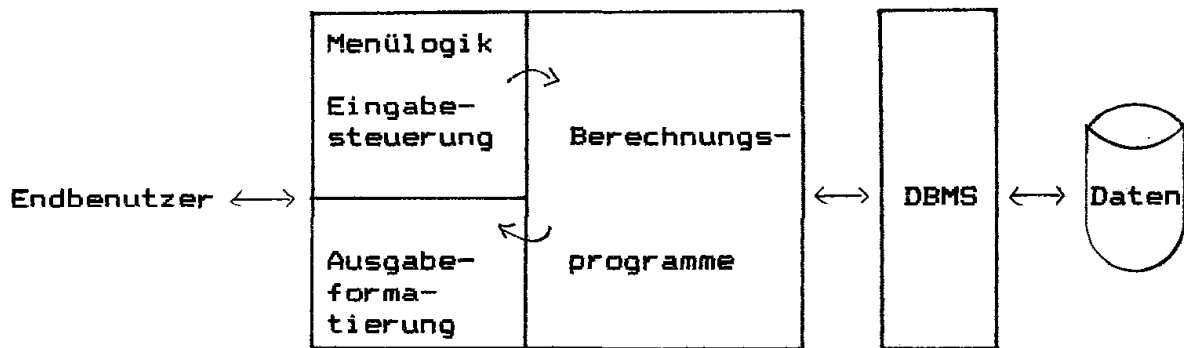


Bild 2 — Aufbau eines Anwendungssystems auf der Basis eines Datenbanksystems

Die Fähigkeiten eines Datenbanksystems werden im einzelnen hierunter näher beschrieben.

Mechanismen für die Gewährleistung der Integrität:

1. Spezifikation von Wertebereichen für Daten;
2. Spezifikation von Bedingungen für konsistente Datenbankzustände (solche Bedingungen werden bei der Ausführung von Aktionen überprüft);
3. Transaktionsverwaltung (zusammengehörige Aktionen werden in einem Block, erst nach Transaktionsende, auf einer Datenbank ausgeführt);
4. Schreibautorisation, sodaß nur benannte Personen Daten ändern dürfen.

Mechanismen für die Kontrolle der Datensicherheit sind:

1. Benutzerkennungsprüfung;
2. Autorisationskontrolle für den Zugriff auf ausgezeichnete Datenmengen;
3. Sicherung der zugrundeliegenden Dateien gegen unberechtigtes Lesen über das Betriebssystem mittels Verschlüsselung der Daten.

Portabilität:

- Das Datenbanksystem „paßt“ auf die Hardware, für die es konzipiert wurde. Es gibt bereits Datenbanksysteme, die auf mehreren Rechnertypen implementiert sind, die Anwendungsprogramme sind in diesem Rahmen portabel.

Flexibilität:

1. Die Datenschemata können im laufenden Betrieb geändert werden, d. h. es können neue Datentypen zugefügt werden.
2. Der Umfang der Daten braucht vorher nicht bekannt zu sein und es können sehr große Datenmengen verwaltet werden.
3. Ein Datenbanksystem läßt auch eine interaktive Verarbeitung der Daten zu.

Effizienz:

1. Die Zugriffspfade können im laufenden Betrieb geändert werden, damit eine höhere Verarbeitungsgeschwindigkeit für geänderte Anforderungen erreicht werden kann.
2. Es gibt einen Anfrage-Optimierungs-Mechanismus, mit dessen Hilfe der schnellste Zugriff auf die gewünschten Daten ermittelt wird.

3. Die Daten können gezielt nach einer gewünschten Ordnung im Speicher plaziert werden (Clustering).
4. Man hat die Kontrolle über den Grad der Redundanz von Daten; redundante Daten können gehalten oder gelöscht werden, wie es für die effiziente Verarbeitung am günstigsten ist.

Außerdem bietet ein Datenbanksystem noch die folgenden, bisher noch nicht erwähnten Dienste:

Mehrbenutzerbetrieb:

- Mehrfachbenutzbarkeit des Datenbanksystems, d. h. Sicherung des konkurrierenden Zugriffs auf Daten von verschiedenen Anwendungsprogrammen „zur gleichen Zeit“.

Recovery:

- Sicherung des Wiederanlaufes, d. h. Gewährleistung des „richtigen“ Datenbankzustandes nach einem Systemzusammenbruch.

3 Gliederung des Vorhabens „Graphische Datenbank“

Wegen der oben aufgezeigten Vorteile von Datenbanksystemen sollen diese nun auch im geowissenschaftlichen Bereich eingesetzt werden. Die kommerziell verfügbaren Datenbanksysteme sind für geowissenschaftliche Anwendungen nicht besonders geeignet. Sie sind so aufgebaut, daß nur einfache Objekte, wie z. B. Zahlen und Texte, direkt abgespeichert werden können, d. h. es werden keine Operationen auf komplexeren, z. B. graphischen Objekten unterstützt (vgl. auch [DKML 84] und [HäRe 83]). Deshalb erscheint ein speziell dafür zugeschnittenes System wünschenswert. (Erste Vorschläge zur Architektur von Nicht-Standard-Datenbanken werden in [Mit 84] vorgestellt.)

Die Einzelprojekte, die in dem Arbeitskreis „Graphische Datenbank“ des DFG-SPP koordiniert werden, sollen nach den Wünschen der Anwender ein solches geowissenschaftliches Datenbanksystem erstellen. Der Zusammenhang zwischen den Projekten, und damit der Aufbau der Graphischen Datenbank, ist in Bild 3 aufgelistet. Die einzelnen Schichten der Graphischen Datenbank werden hiernach in der Reihenfolge V-I näher erläutert.

In Ebene V ist das eigentliche Datenbanksystem angesiedelt. In dem Projekt von Prof. *Schek*, TH Darmstadt, wird eine für geowissenschaftliche Anwendungszwecke besser geeignete Kernsoftware (Datenbank-Kern), insbesondere ein Speicher- und Zugriffssystem, entwickelt. Bis zur Fertigstellung und für Vergleichszwecke wird ein verfügbares relationales Datenbanksystem eingesetzt. An der Schnittstelle zur nächsthöheren Ebene, der internen Datenbankschnittstelle, stehen Attribute und Relationen (das sind Merkmale und Beziehungen von Objekten) zur Verfügung.

In der Ebene IV wird in unserem Projekt eine komplexe Datenbankfunktionssoftware entwickelt, die auf die interne Datenbankschnittstelle der Ebene V aufsetzt und eine Schnittstelle bereitstellt, die die benutzerorientierten Aspekte der Datenbank-Kompo-

nente unter Berücksichtigung der kartographischen Weiterverarbeitung unterstützt. In der Schnittstelle zur Ebene III werden graphische Objekte wie z. B. Punkt, Linie, Polygon und Fläche, alphanumerische Objekte wie z. B. Zeichenketten und Bemaßungen zur Verfügung gestellt. Zu den allgemeinen Grundoperationen (Abspeichern, Verknüpfen, Wiederfinden) werden umfangreiche Such- und Änderungsoperationen zur Unterstützung eines graphischen Editors und anderer graphischer Weiterverarbeitungsprogramme angeboten: z. B. Bilden neuer Objekte aus vorhandenen, Suche von Schnittpunkten von Objekten, Suchen in der Umgebung von Objekten usw.

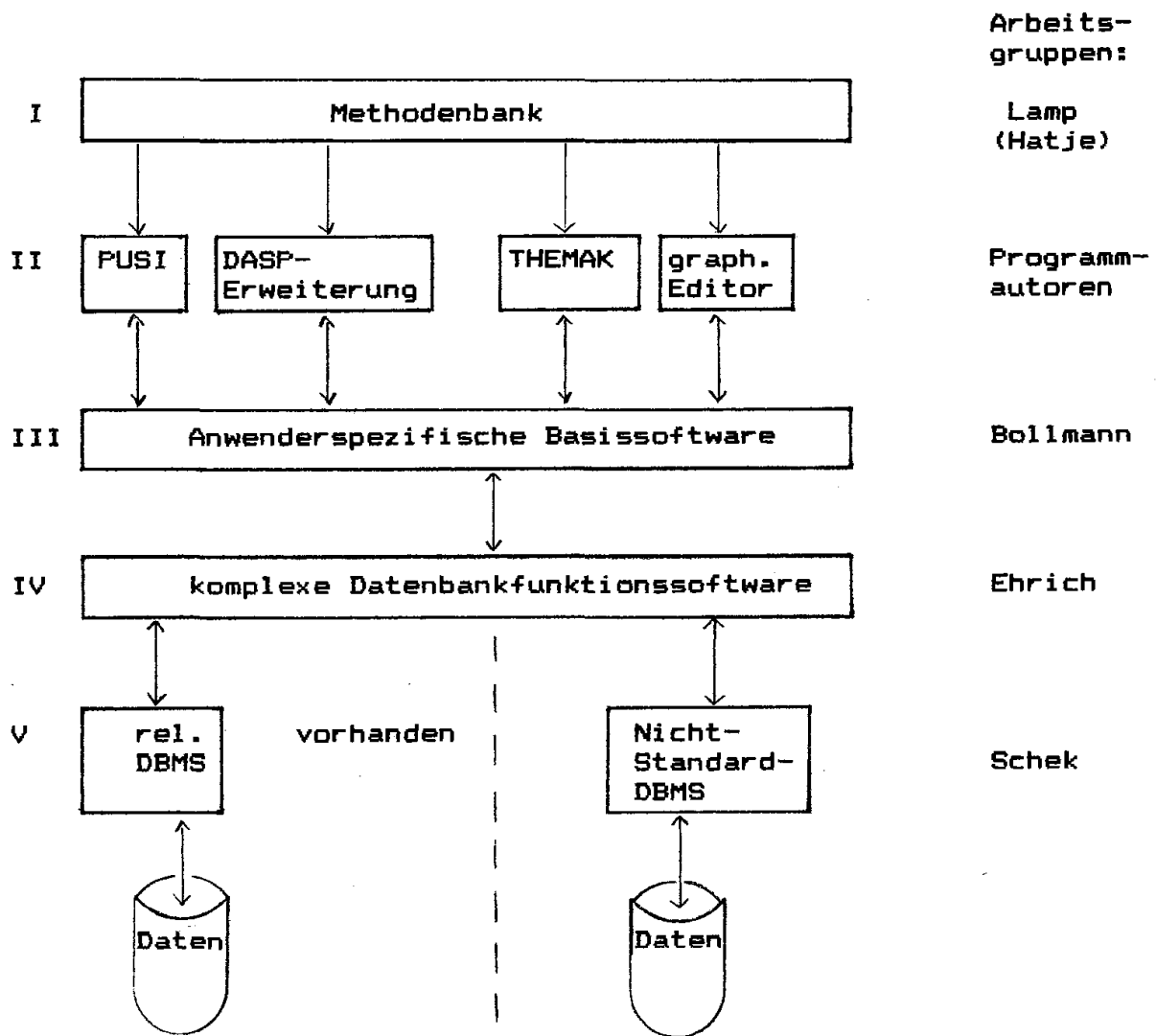


Bild 3 — Schichtenaufbau der Graphischen Datenbank

Die III. Schicht wird gebildet durch das Projekt von Dr. *Bollmann*, FU Berlin. Hierin wird die Anwenderspezifische Basissoftware erstellt. Die zugehörige Schnittstelle ist mit dem Ziel geplant, graphische Objekte mit Attributen (z. B. Linie A ist Straße) darzustellen und Graphikoperationen wie z. B. Linienglättung, Flächeninterpolation o. ä. und graphische Ein-/Ausgabe nach dem GKS-Standard (graphisches Kernsystem) zu ermöglichen.

Die Ebene II wird gebildet durch die zum Teil bereits bestehenden Weiterverarbeitungsprogramme, die von diversen Programmautoren stammen und auch weiterhin gepflegt werden. Es handelt sich hierbei um spezielle Anwendungsprogramme, die gespeicherte Daten für die graphische Ausgabe oder Weiterverarbeitung aufbereiten.

In der obersten Ebene ist das Projekt von Dr. *Lamp*, Universität Kiel, angesiedelt, in dem ein vollständiges Methodenbanksystem implementiert werden soll. Das System wird die Programme aus der Ebene II integrieren, leicht erweiterbar sein und ein Auskunftssystem enthalten. Hierüber sollen die Anwendungsprogramme angewählt und einheitlich aufgerufen werden können.

4 Arbeitsprogramm zum Geo-Datenbank-Entwurf

Im folgenden wird nach einer Erläuterung des Grobkonzeptes das Arbeitsprogramm für das Teilprojekt „Geo-Datenbank-Entwurf“ (Ebene IV) vorgestellt.

Es ist geplant, die graphischen und rein alphanumerischen Daten in getrennten Datenbanken zu halten, wobei eine Aufteilung in eine Karten-Datenbank und mehrere alphanumerische Datenbanken denkbar ist. In die alphanumerischen Datenbanken gehen unter anderem die bisher in Dateiform gespeicherten Daten ein; in der Karten-Datenbank sollen graphische Objekte gespeichert werden. Diese Datenbanken werden als sogenannte Server-Datenbanken konzipiert, die sich dadurch auszeichnen, daß sehr große Datenmengen gespeichert werden, auf denen nur begrenzte, weniger komplexe und komfortable Operationen möglich sind, die jedoch den Mehrbenutzerbetrieb mit konkurrierendem Zugriff zulassen.

Auf diese Server-Datenbanken sollen Arbeits-Datenbanken aufsetzen: Aus der Karten-Datenbank und den alphanumerischen Datenbanken werden Daten, wie z. B. einzelne Blattgebiete TK 25, extrahiert und in Arbeits-Datenbanken übertragen. Eine solche Arbeits-Datenbank kann jeweils nur von einem Anwender zur Zeit, ohne konkurrierenden Zugriff, benutzt werden; es können jedoch mehrere Arbeits-Datenbanken gleichzeitig auf die Server-Datenbanken zugreifen.

Somit können komplexe, relativ komfortable Operationen auf begrenzten Mengen von graphischen und alphanumerischen Daten ausgeführt werden (z. B. lokale und/oder thematische Kartenauszüge). In der Arbeits-Datenbank befindliche Daten können also beliebig manipuliert (gesucht, geändert, ...) und anschließend in die Server-Datenbanken zurückgeschrieben werden. Hierfür muß ein Konzept zur Zuordnung oder Verknüpfung graphischer und alphanumerischer Daten erarbeitet werden.

Die Arbeits-Datenbank stellt die Benutzerschnittstelle zur Verfügung, die entsprechend des Informationsbedarfs und der Anforderungen der verschiedenen Anwendergruppen entworfen werden muß. Denkbar wäre die Erstellung mehrerer Versionen von Arbeits-Datenbanken mit unterschiedlichen Schnittstellen für verschiedene Benutzergruppen; zunächst ist insbesondere eine kartographische Datenbank geplant.

Das Arbeitsprogramm, konzipiert und aufgeteilt für zwei wissenschaftliche Mitarbeiter, sieht folgendermaßen aus:

1. Praktischer Entwurf und Implementierung
 - a) Prototyp der Arbeits-Datenbank
 - konzeptioneller Entwurf der Benutzerschnittstelle der Datenbank-Komponente und deren Prototyp-Implementierung auf einem vorhandenen relationalen Datenbanksystem.
 - b) Entwurf der Karten-Datenbank
 - keine vollständige Datenbank-Implementierung, sondern Pilotimplementierung ausgewählter Teilaspekte auf einem vorhandenen relationalen Datenbanksystem zur Demonstration der Praktikabilität des Entwurfs.
 - c) Effizienter Neuentwurf und Implementierung
 - nach Fertigstellung des Datenbank-Kernes der Gruppe von Prof. *Schek*.
2. Grundlagenuntersuchungen zur Datenbank-Entwurfsmethodik mit komplexen Objekten
 - a) Abstrakte Datentypen
 - durchgängige formale Spezifikationstechnik für Nicht-Standard-Datenbankschemata mit solider semantischer und theoretischer Grundlage.
 - b) Entwurfsbewertung und -optimierung
 - Untersuchung über Methodik und Grundlagen des Datenbank-Entwurfs mit komplexen Objekten.
 - c) Werkzeuge.

Für den Zeitraum Oktober 1984 bis Oktober 1985 ist nur eine wissenschaftliche Mitarbeiterstelle bewilligt worden, daher wird Teil 2 zunächst nicht bearbeitet. Das Gesamtprojekt ist auf fünf Jahre angelegt.

5 Literatur

[DKML 84] *Dittrich, K. R.; Kotz, A. M.; Mülle, J. A.; Lockemann, P. C.*: Datenbankkonzepte für Ingenieur Anwendungen: eine Übersicht über den Stand der Entwicklungen. — Proc. 14. GI-Jahrestagung, *H.-D. Ehrich* (ed.), Informatik Fachbericht 88, Springer-Verlag, Berlin 1984, S. 175–192.

[HäRe 83] *Härder, T.; Reuter, A.*: Database Systems for Non-Standard Applications. — Proc. International Computing Symposium, *H. J. Schneider* (ed.), Teubner-Verlag, Stuttgart 1983, S. 452–466.

[HoWh 83] *Holsapple, C. W.; Whinston, A. B.*: Guidelines for DBMS Selection. — In: Data Base Management: Theory and Applications. — *C. W. Holsapple; A. B. Whinston* (ed.), D. Reidel Publishing Company, Dordrecht, Holland, 1983, S. 175–191.

[Mit 84] *Mitschang, B.*: Überlegungen zur Architektur von Datenbanksystemen für Ingenieur Anwendungen. — Proc. 14. GI-Jahrestagung, *H.-D. Ehrich* (ed.), Informatik Fachbericht 88, Springer-Verlag, Berlin 1984, S. 318–334.

Herrn Dr. Lipeck danken wir für seine kritischen Anmerkungen