

UNIGIS MSc (GIS) 2002

**Vom Coverage zur Geodatabase
der Umstieg von der filebasierten zur daten-
bankbasierten Haltung von Geodaten**

**eingereicht von
Jürgen Oberreißl (U964)**

Dornbirn, Oktober 2004

I. Eidesstattliche Erklärung

Ich versichere, dass ich die nachfolgende Arbeit selbständig und ohne fremde Hilfe angefertigt habe. Ich habe mich anderer als der im beigefügten Verzeichnis angegebenen Hilfsmittel nicht bedient. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht.

Dornbirn, im Oktober 2004

(Jürgen Oberreißl)

II. Zusammenfassung:

Die Welt der Geographischen Informationssysteme befindet sich derzeit in einem als fundamental zu beschreibenden Umbruch. Diese als Paradigmenwechsel zu bezeichnende Veränderung betrifft die Haltung von Geodaten. Die geographische Information, welche bislang ausschließlich in Dateien gespeichert wurde, wird in modernen Systemen in relationalen Datenbanken abgelegt. Der Umstieg vom Coverage zur Geodatabase ist Thema dieser Arbeit.

Ein wichtiger Aspekt ist dabei jener der Datenmodellierung. Ein Modell ist ein auf ihre Grundfunktionen reduziertes Abbild der realen Welt. Das Coverage Modell basiert auf den geometrischen Primitiven Punkt, Linie, Fläche. Diesen werden durch Attribute Eigenschaften zugeordnet. Die Suche nach einer, dem logischen Datenmodell näher stehenden Methode zur Beschreibung geographischer Entsprechungen führte zur Entwicklung des Geodatabase-Modells: GIS Datenbestände sollen dadurch intelligenter und mit natürlichem Verhalten ausgestattet werden können. Durch den erweiterten Umfang an Möglichkeiten ist das Geodatabase-Modell sehr komplex, weshalb Modellierungstechniken eine wichtige Unterstützung beim Umstieg darstellen.

Denn eine Datenbank, welche die in sie gesetzten Erwartungen nicht erfüllt, verursacht hohe Kosten und wird von den Nutzern nicht akzeptiert. Als Modellierungstechniken haben sich die ER-Modellierung (Entity-Relationship) und die Relationale Modellierung durchgesetzt. Durch die UML-Notation (Unified Modelling Language) wird die Modellierung von Objekten (wie sie in der objektorientierten Geodatabase zur Anwendung kommen) vereinfacht. Unabhängig von der benutzten Modellierungstechnik ist der entscheidende Vorteil der Modellierungsphase, dass ein tieferes Verständnis für die Datenwelt eines Unternehmens entwickelt wird. Ergänzend dazu werden Fragen der Organisation, Schulung, Ausstattung und der Zeitplanung erörtert. In dieser Arbeit wird auf Grundlage der Modellbildung ein generelles Ablaufschema für den Umstieg von Coverage auf Geodatabase erstellt, das anhand von Beispielen auf seine Gültigkeit untersucht wird.

Zunächst wird mit Hilfe des im Kapitel 4.2.4 erarbeiteten Ablaufschemas die Datenmodellierung durchgeführt. Ausgehend von einer mit UML Notation erstellten ER-Modellierung wird ein Relationales Modell (ebenfalls in UML Notation) erstellt. Dieses Relationale Modell dient als Vorlage für die Erstellung der einzelnen Tabellen, welche mit den entsprechenden Werkzeugen direkt in der Geodatabase angelegt werden können. Weiters werden die für die Erstellung topologischer Beziehungen notwendigen Regeln, Attributdomänen und Rangreihungen definiert.

Abschliessend werden die Erfahrungen aus der Anwendung des Ablaufschemas und somit der Überführung der Daten vom Coverage zur Geodatabase zusammengefasst.

Abstract:

The world of geographic information systems is going through a time of fundamental upheaval concerning the storage of geodata. Up to now data have been stored solely in files, so called coverages, whereas nowadays in modern systems data are stored in relational geodatabases. How to convert coverages into geodatabases is topic of this thesis. An appropriate approach are the different methods of data modeling. A model is an image of the real world reduced to its basic functions. The coverage model is based on the geometric primitives point, line and area to which characters are assigned by means of attributes. The geodatabase is a further development that is also based on geometric primitives but provides much more options with respect to the characteristic of an object. The geodatabase model is closer to a logical data model which increases its performance but makes it very complex to operate. So the right selection of the modeling technique is crucial for a successful conversion of coverages into geodatabases. Otherwise there is the risk of creating a database that does not fulfill the users' expectations causing high costs.

The modeling techniques of choice are ER-modeling (entity relationship) and relational modeling. Using the unified modeling language eases the modeling of objects (as applied in the object oriented geodatabase). A benefit of this modeling phase is a better understanding of a companies dataworld and furthermore additional questions of organisation, training, equipment and time shedule are discussed.

In this thesis, the modeling is taken as a starting point for creating a general procedure for the change from coverages to geodatabases. The applicability of this procedure is investigated by means of examples. Starting with the development of the change procedure in chapter 4.2.4, the datamodeling is done next. First a relational model is created from an ER-modeling, both made in UML. The relational model is the blueprint for making tables, which can be generated directly in the geodatabase with the appropriate tools. Next, the rules, attribute domains and rank are defined for creating topological relationships. Finally experiences in using the change procedure are described with the help of case studies.

Inhaltsverzeichnis

1	<i>Einleitung</i>	3
1.1	Problemstellung.....	4
1.2	Konkretisierung der Thematik, Abgrenzung zu anderen GIS-Systemen.....	5
1.3	Zusammenfassung.....	7
2	<i>Datenmodelle im Überblick</i>	8
2.1	Filebasierte Datenhaltung versus datenbankbasierte Datenhaltung.....	10
2.2	Coverage – der Erfolg eines Datenmodells.....	11
2.2.1	Die Entwicklung.....	11
2.2.2	Das Datenmodell.....	12
2.2.3	Das Gesamtsystem.....	14
2.2.4	Vorteile und Nachteile.....	14
2.2.5	Die Zukunft der Coverage.....	15
2.3	Exkurs: Topologie – d a s Argument für GIS.....	15
2.3.1	Was ist Topologie?.....	16
2.3.2	Der Nutzen topologischer Strukturen.....	17
2.3.3	Der Anlass für die Entwicklung neuer Technologien.....	18
2.4	Das Geodatabase Konzept.....	18
2.4.1	Die Abbildung bisheriger GIS Funktionalitäten in der Geodatabase.....	19
2.4.2	Neue Funktionalitäten einer Geodatabase.....	19
2.4.3	Die Zukunft der Geodatabase.....	23
2.5	Zusammenfassung.....	24
3	<i>Der Unterschied zwischen Coverage und Geodatabase</i>	25
3.1	Hardware.....	26
3.2	Software.....	26
3.3	Daten.....	26
3.4	Personal.....	27
4	<i>Der Umstieg in der Praxis – Methodik</i>	28
4.1	Die Sicht des Herstellers.....	28
4.1.1	Dokumentation und Whitepapers.....	28
4.1.2	Case studies.....	31
4.2	Planung: technische Aspekte des Umstiegs.....	32
4.2.1	Theorie der Datenmodellierung.....	32
4.2.2	Vorgangsweise zur Erstellung logisch zweckmässiger Datenmodelle... ..	36
4.2.3	Vorschlag für ein Ablaufschema auf Basis der UML-Notation.....	38
4.2.4	Ablaufschema:.....	46

4.3	Planung: organisatorische Aspekte des Umstiegs	48
4.3.1	Exkurs – der Wechsel als Prinzip (Change Management)	48
4.3.2	Organisation und Schulung	49
4.3.3	Ausstattung	50
4.3.4	Zeitplan.....	51
4.4	Die Umsetzung	52
4.4.1	Exkurs – die Durchführung komplexer Vorhaben als Aufgabe (Systems Engineering)	52
4.4.2	Anwendung des Ablaufschemas.....	54
4.4.3	Dauer der Umsetzung	55
4.4.4	Betrieb während der Umsetzung	55
4.5	Die Nachbereitung.....	55
4.5.1	Abschliessende Arbeiten (final Clean Up)	55
4.5.2	Kritik der Anwendung des Ablaufschemas	56
4.5.3	Dokumentation des Umstieges	56
4.6	Erfahrungsberichte	57
4.7	Zusammenfassung	58
5	<i>Fallstudien</i>	60
5.1	Grundlagenthema: DKM	60
5.1.1	Datenmodellierung	60
5.1.2	Schlussfolgerungen.....	67
5.2	Fachthema: Fliessgewässer	68
5.2.1	Datenmodellierung	68
5.2.2	Schlussfolgerungen.....	71
5.3	Erfahrungen aus der Anwendung des Ablaufschemas	73
5.4	Zusammenfassung	74
6	<i>Status in den neun Bundesländern</i>	75
6.1	Derzeitiger Stand der Entwicklung.....	75
7	<i>Schlußfolgerungen</i>	76
8	<i>Literaturverzeichnis</i>	77
9	<i>Glossar</i>	79

1 Einleitung

Die GIS Welt steht derzeit an der Grenze zu einer Neuorientierung in der Haltung von Daten. Ein jahrzehntelang bewährtes Modell zur Erfassung, Speicherung und Bearbeitung von Daten ist dabei, von einem neuen abgelöst zu werden. Die Unsicherheit beim Umstieg von Bewährtem zu Neuem und Unbekanntem ist wie in jedem anderen Bereich auch im Gebiet der GIS-Technologie deutlich spürbar. Es ist die Rede vom Wechsel von der filebasierten Datenhaltung zur datenbankbasierten Datenhaltung; vereinfacht gesagt vom Coverage zur Geodatabase.

Dabei reden wir nicht von einem Versionswechsel, auch nicht von einer Ergänzung bisheriger Technologien; es handelt sich vielmehr um einen Paradigmenwechsel, eine grundlegende Änderung. Zwar wurden auch bislang schon Datenbanktechnologien verwendet, allerdings nicht in dieser alles durchdringenden und umfassenden Form wie sie in Zukunft zu denken und zu leben sein wird.

Das wird eine Änderung in der Gewinnung und Anwendung geographischer Daten zur Folge haben. Diese Umstellung (und auch deren Begleiterscheinungen) ist das Thema dieser Arbeit. Es wird insbesondere zu untersuchen sein, welche Auswirkungen ein Wechsel von der gegenständlichen Technologie auf die neue auf bestehende Systeme haben wird und wie ein solcher Umstieg zweckmässigerweise zu gestalten ist.

Jedenfalls muss bereits an dieser Stelle der Hinweis darauf gemacht werden, dass nicht im geringsten die Illusion einer friktionsfreien Ablösung des bisherigen Systems aufkommen darf. Vielmehr sind die Befürchtungen und Ängste der Anwender einer entsprechenden Würdigung zu unterziehen.

Einige erläuternde Worte und eine damit verbundene Einschränkung des Themas seien an dieser Stelle ebenfalls erlaubt:

Es handelt sich bei der Darstellung der Beispiele um ein System eines bestimmten Anbieters (nämlich ESRI), weil hier der Zugang für mich am leichtesten zu bewerkstelligen ist. Das Thema bezieht sich jedoch zumindest im allgemeinen Teil nicht auf ein bestimmtes System oder einen bestimmten Anbieter – der Paradigmenwechsel von dem die Rede ist, hat die gesamte GIS-Welt erfasst und ist damit prinzipiell für jedes System gültig.

Am Ende dieser Arbeit ist klar, welche Änderungen zweckmässigerweise durchzuführen sind und wie diese am besten durchgeführt werden können. Damit die ganze Thematik nicht nur im grauen Bereich der Theorie abgehandelt werden soll, sind Beispiele angeführt, welche sich auf Datenbestände des VoGIS (Vorarlberger Geographisches

Informationssystem) beziehen. Ein Blick auf die Entwicklung in anderen österreichischen Bundesländern soll diese Untersuchung vervollständigen.

Es ist aufzuzeigen, dass es nicht ausreicht, Datenbestände und Applikationen in eine neue Form zu bringen, sondern die ganzen Abläufe, welche dahinter stehen, die Organisation, die Mitarbeiter, die Anwendungen und ähnliches, zu beleuchten.

Arbeitsabläufe, logische Datenbeziehungen und Bearbeitungsmethoden sind gegebenenfalls neu zu definieren oder anzupassen.

Auf eine wichtige Einschränkung dieser Arbeit muss an dieser Stelle noch hingewiesen werden: Ich beschränke mich bei der gegenständlichen Untersuchung (insbesondere bei der Darstellung der Beispiele) auf Daten welche auf Vektoren basieren. Die Besonderheiten von Rasterdaten werden dabei nicht behandelt.

1.1 Problemstellung

In der Einleitung ist von Befürchtungen und Ängsten die Rede, welche zu diesem Zeitpunkt nicht konkret angesprochen wurden. Die vielfach (meist von GIS-Laien) geäußerte Meinung, dies könne doch nicht so schwer sein und man solle einfach die Daten durch geeignete Werkzeuge in eine räumliche Datenbank schieben, kann sehr leicht zum einen GIGO-Effekt (Akronym für „garbage in, garbage out“) führen. Weisen nämlich die importierten Daten nicht jene Struktur auf, welche vom System für eine zweckmässige (oder vielleicht sogar verbesserte) Verarbeitung benötigt werden, hat man nichts gewonnen, vielleicht sogar wesentliches verloren. Zur Verdeutlichung der Problemstellung sei das Ergebnis eines Migrationsprojektes vorangestellt, welches die Überführung von filebasierter zu datenbankbasierter Datenhaltung zum Gegenstand hatte. Seine Erkenntnisse über Vor- und Nachteile können so zusammengefasst werden [siehe SCHI03b]:

„...“

- *Konvertierung eines großen, strukturierten Geo-Datenbestandes in integriertes Datenhaltungskonzept ist zeitaufwendig*
- *Vorhandene Datenstrukturen (z.B. Thematische Klassifizierung) gehen verloren*
- *Abhängigkeit von GIS-Herstellern (zugrundeliegendes Datenmodell)*
- *Black-Box-ähnliche Datenstrukturen, die nur mit der spezifischen GIS-Software erkannt und genutzt werden können*
- *Kartographische Gestaltung von Geodaten geht verloren (vollständige Neukonzeption von Kartendarstellungen)*
- *Mitunter Performanceeinbußen (Räumliche Indizierung essentiell)*

„...“

dem stehen folgende Vorteile gegenüber

„...“

- *Sicherungsmechanismen und Transaktionskonzept der Datenbank auch für Geometriedaten (lange Transaktionen)*
- *Möglichkeit zum gesicherten Mehrbenutzerbetrieb (Versionierung und Sperrverhalten)*
- *Internetfähigkeit: Webbasierter Zugriff auf Geometrie- und Sachdaten*
- *Gezielte Zuweisung von Zugriffsrechten (read/write) an Datenbanknutzer*
- *Möglichkeit zur integrierten Verwaltung von Geodaten und Meta-Daten*
- *Möglichkeit zum Mosaikieren von Rasterbildern in der Datenbank*

„...“

Die Problemstellung kann also folgendermassen zusammengefasst und konkretisiert werden: Der technische Fortschritt hat zu einer Weiterentwicklung der Verfahren in der Haltung und Verarbeitung von GIS Daten geführt. Diese Möglichkeiten umfassen in erster Linie die Abbildung nicht nur alphanumerischer, sondern auch graphischer Daten in relationalen Datenbanken. Diese neue Methodik eröffnet neue Perspektiven in der Datenmodellierung, was sich auf alle Bereiche der Verarbeitung und Haltung von GIS relevanten Informationen erstreckt. Für diese Neuerungen muß ein Preis bezahlt werden; Vor- und Nachteile sind die Folge.

Ein etwaiger Technologiewechsel wie oben angedacht, muß geplant werden; ausserdem sollen mögliche Fallstricke frühzeitig erkannt werden können.

Die gegenständliche Arbeit hat sich zum Ziel gesetzt, die genannten Aspekte zu diskutieren und anhand von Beispielen die gefundenen Gesetzmässigkeiten zu verifizieren.

1.2 Konkretisierung der Thematik, Abgrenzung zu anderen GIS-Systemen.

Eine große Gefahr, welche in der Bearbeitung eines so weiten Themenbereiches liegt, ist jene der zu weiten Fassung des Themas und damit der großen Fülle an zu behandelnden Fragen. Um diesem Problem aus dem Weg zu gehen wird nachfolgend versucht, anhand einer konkreten Fragestellung die allgemein gültigen (bzw. für alle Systeme zutreffenden) Aussagen herauszuarbeiten.

Dies ist allein deshalb notwendig, weil durch die Vielzahl auf dem Markt im Einsatz befindlicher GIS Systeme die unterschiedlichsten Datenformate zum Einsatz kommen und die Behandlung all dieser Varianten im Rahmen dieser Arbeit nicht möglich ist. Dennoch ist beabsichtigt, Allgemeingültigkeit in den Aussagen und Erkenntnissen in den Vordergrund zu stellen.

Die Zusammenführung dieser beiden Absichten, nämlich der Darstellung der allgemeingültigen Aussagen und die Behandlung eines konkreten Umstiegsszenarios, soll bewerkstelligt werden, indem auf Basis eines konkreten Systems (und damit eines Datenmodells) eines Herstellers der Umstieg aufgezeigt wird. Die daraus gewonnenen Erkenntnisse werden verallgemeinert dargestellt.

Dies ist der Grund, weshalb in dieser Arbeit die Umstellung ausgehend von dem im Titel angeführten Datenformat, nämlich dem Coverage-Format, in eine Geo-Datenbank geht.

Dort, wo es der Erläuterung konkreter Vorgänge dient, wird die Grundlage das spezifische Datenmodell sein, dort, wo eine allgemein darstellbare Erkenntnis im Zusammenhang mit dem Umstieg gefunden wird, soll der allgemeingültige Merksatz als Richtschnur dienen.

1.3 Zusammenfassung

Die Welt der Geographischen Informationssysteme befindet sich derzeit in einem als fundamental zu beschreibenden Umbruch. Diese als Paradigmenwechsel zu bezeichnende Veränderung betrifft die Haltung von Geodaten. Die geographische Information, welche bislang ausschliesslich in Dateien gespeichert wurde, wird in modernen Systemen in relationalen Datenbanken abgelegt. Damit werden Attributdaten und Geographische Daten in einem Speicher zusammengefasst. Damit verbunden ist eine Erweiterung der Funktionalitäten von GIS-Systemen und neue Möglichkeiten der Datenmodellierung.

Die Vor- und Nachteile dieser neuen Möglichkeiten werden untersucht und aufgezeigt. Dies erfolgt anhand eines konkreten GIS-Systems. Die gefundenen Ergebnisse sollen in verallgemeinert als Unterstützung für eine Anpassung und Umarbeitung von Geographischen Datenbeständen dienen.

2 Datenmodelle im Überblick

Bevor die reale Welt in einem GIS-System abgebildet werden kann, muss sie vereinfacht und auf ihre Grundfunktionen reduziert werden. Dabei entsteht ein Modell der realen Welt.

„Das Datenmodell beschreibt die Leistungsfähigkeit eines GIS!“ So überschreibt SCHILCHER [SCHI03b] seine Vorlesungsunterlagen. Er will damit zum Ausdruck bringen, dass ein GIS-System nur so gut sein kann, wie die Daten, welche in ihm abgelegt sind. Und diese wiederum können ihren Zweck nur so gut erfüllen, wie sie in der Lage sind, die Realität abzubilden. Genauer gesagt geht es hierbei um räumliche Datenmodelle. Über die Theorie der Datenmodelle und ihre Erstellung wird später noch zu diskutieren sein. Zunächst soll, mit einem stärkeren Fokus auf die Praxis, die Systematik der Datenmodelle im Vordergrund stehen.

Die Fachliteratur [ESR94, ZEI99] kennt drei Kategorien, nämlich das Vektor-Datenmodell, das Raster-Datenmodell und das TIN-Datenmodell (TIN steht dabei für Triangular Irregular Network). Im Zusammenhang mit der in dieser Arbeit behandelten Frage nach der Portierung von Daten vom Coverage zur Geodatabase ist in erster Linie das Vektor-Modell von Interesse. Es setzt sich aus Punkten, Linien und Flächen zusammen, welche jeweils Objekte der realen Welt im GIS repräsentieren und weist unterschiedliche logische Verbindungen zwischen diesen Elementen auf.

Im GIS Bereich im weiteren Sinne kommen diese Datenmodelle in zumindest drei „Modell-Welten“* zur Anwendung. Diese weisen unterschiedliche Strukturen auf, welche zunehmend komplexere Möglichkeiten der Modellbildung umfassen.

- CAD Datenmodell
- Coverage Datenmodell
- Geodatabase Datenmodell

Während das CAD Datenmodell lediglich die Grundobjekte darzustellen vermag, wobei jedes Objekt unabhängig von den anderen ist (dafür hat sich der Begriff „Spaghetti Modell“ eingebürgert) wurde im Jahr 1981 (gemeinsam mit der Markteinführung von ArcInfo) ein Modell eingeführt, welches logische Beziehungen zwischen den Elementen ermöglicht; das Coverage-Modell.

* Der Begriff „Welt“ wird hier verwendet, um eine Abgrenzung zum zuvor verwendeten Modell Begriff herzustellen.

Als Weiterentwicklung des Coverages wurde, einem Trend in der Haltung geographischer Daten folgend, Ende der 90er Jahre des 20. Jahrhunderts das Geodatabase-Modell eingeführt. Zur Charakterisierung der beiden letztgenannten Modelle (das CAD-Modell wird aufgrund seiner geringen Bedeutung in der GIS Welt nicht weiter behandelt) sehen wir uns die Beschreibung von ZEILER [ZEI99] an:

Das Coverage-Modell besteht aus zwei wichtigen Teilbereichen:

- 1.) Räumliche Daten sind mit Attributdaten verknüpft. Die geographischen Daten werden in binären Dateien gespeichert, die Attributdaten in Tabellen. Über einen gemeinsamen Schlüssel werden einander entsprechenden Werte zugeordnet.
- 2.) Topologische Beziehungen zwischen Vektorelementen können gespeichert werden.

Mit diesem Konzept können komplexe GIS-Anwendungen mit geringem Ressourcenverbrauch realisiert werden. Der Nachteil dieses Ansatzes liegt darin, dass räumliche Objekte jeweils in Sammlungen von Punkten, Linien oder Flächen (d.h. geometrischen Primitiven) abgelegt werden, welche ohne die entsprechenden (und für die Unterscheidung wichtigen) Attribute gleiche Eigenschaften aufweisen.

Es wurde nach Wegen gesucht, wie die Eigenschaften von Elementen eng mit der geometrischen Ausprägung verknüpft werden konnte. Das soll mit dem Geodatabase-Modell erreicht werden. Das Physikalische Datenmodell soll dem logischen Datenmodell näher gebracht werden; das bedeutet mit anderen Worten, der objektorientierte Ansatz, welcher schon seit einiger Zeit die IT Welt in zunehmendem Maße durchdringt, soll auch im GIS eine neue Ära einleiten.

Die Vorteile des Geodatabase Modells werden dabei wie folgt dargestellt:

- Einheitlicher Verwahrungsort für geographische (geometrische und alphanumerische) Daten
- Verlässlichere Eingabe und Bearbeitung von Daten
- Benutzer arbeiten mit der realen Welt ähnlichen Objekten
- Elemente bekommen ein reichhaltigeres Umfeld
- Viele Nutzer können die Daten gleichzeitig bearbeiten
- ...und einige mehr.

Dieser Überblick soll einen ersten Hinweis darauf geben, in welche Richtung sich derzeit die Diskussion entwickelt. Obgleich die Vorteile des Geodatabase-Modells un-

bestritten sind (vor allem und in erster Linie in der Technologie der verteilten Datenhaltung), stellen sich Fragen, die nicht ohne weiteres befriedigend beantwortet werden können. Nachfolgend wird versucht, sich dem Problem Schritt für Schritt anzunähern und entsprechende Antworten aufzuzeigen.

2.1 Filebasierte Datenhaltung versus datenbankbasierte Datenhaltung

Im Grunde müsste der Begriff filebasierte Datenhaltung in Anführungszeichen gesetzt sein, denn letztendlich werden alle Datenbestände, welche in EDV-Systemen verarbeitet werden in einer Datei abgelegt und gespeichert; auch solche in einer Datenbank.

In diesem Zusammenhang bezeichnet der Begriff „filebasiert“ die duale Konzeption in der Verspeicherung von Daten in einem Coverage; die geometrischen Inhalte werden im binären Format in einer Datei persistent vorgehalten, während die alphanumerisch vorhandenen Attribut (- und Systemdateien) in den Tabellen eines relationalen DBMS (INFO oder anderen RDBMS) abgelegt werden.

Die filebasierte Datenhaltung kennzeichnet somit ein duales System aus binären Dateien, welche die Geometrie enthalten und den die Relationen enthaltenden Datenbankdateien – Arc und Info. Diese Verbindung wurde in der Entstehungszeit des Systemes darum gewählt, weil es insbesondere in einer Zeit beschränkter Rechenleistung den besten Kompromiss zwischen Funktionalität und Antwortverhalten darstellte.

Die datenbankbasierte Datenhaltung geht mit den revolutionär einzustufenden Entwicklungen in der Hardwaretechnologie und der damit verbundenen Vervielfachung der Rechenleistung einher. Bis vor wenigen Jahren waren Relationale Datenbanksysteme darauf ausgelegt, einfach strukturierte Massendatensätze zu verwalten. Die Verarbeitung von Daten mit einem Raumbezug überforderte die Systeme. Die Softwareanbieter einschlägiger Produkte haben in den vergangenen Jahren große Anstrengungen unternommen, um ihre Produkte um die räumliche Komponente zu erweitern, um neue Märkte zu erschliessen und damit das Einsatzgebiet zu erweitern. Die Systeme haben, zumindest vom Gesichtspunkt der Anbieter, die Marktreife erlangt. Damit können Systeme angeboten werden, welche die getrennte Verspeicherung von geometrischen und alphanumerischen Daten überflüssig machen und ein neues Zeitalter der Datenhaltung in Geographischen Informationssystemen einläuten. Einschränkend muss bemerkt werden, dass nicht alle RDBMS geometrische Inhalte verwalten können. Derzeit werden zwei Ansätze von der einschlägigen Industrie verfolgt. Der Ansatz der Middleware einerseits (hierher gehört die ESRI Software SDE) und der Ansatz der um räumliche Funktionen erweiterten relationalen Datenbank wie z.B. Oracle Spatial.

2.2 Coverage – der Erfolg eines Datenmodells

Mittlerweile ist das Coverage-Modell mehr als 20 Jahre auf dem Markt. Viele Systeme mit großen Datenmengen sind sehr erfolgreich im Einsatz. Die ständige Weiterentwicklung der Funktionalitäten und die beständige Pflege des Systems haben dazu geführt, dass es zu einem verlässlichen, wenn auch nicht einfach aufgebauten und damit nicht leicht zu handhabendem Modell gereift ist. Der Grund, weshalb dem Coverage soviel Erfolg beschieden war (und immer noch ist), soll hier kurz dargelegt werden.

2.2.1 Die Entwicklung

Bereits in den 60er Jahren des vorigen Jahrhunderts wurden verschiedentlich Anstrengungen unternommen, die Kartenproduktion durch elektronische Systeme zu unterstützen. Die ersten Systeme bauten auf einem Datenmodell auf, welches heute noch als „Spaghetti Modell“ bezeichnet wird. Es zeichnet sich dadurch aus, dass die räumlichen Objekte vollkommen unabhängig voneinander beschrieben werden. Der Hauptvorteil dieses Modells ist seine Einfachheit, der Hauptnachteil ist die dem Modell innewohnende Redundanz. Dies war, insbesondere im Hinblick auf die beschränkte Verarbeitungskapazität ein wichtiger Grund für die Suche nach anderen Möglichkeiten, wenngleich dieses Prinzip auch heute noch für bestimmte Anwendungen (vgl. CAD) vollkommen ausreichend ist.

Die Entwicklung von ArcInfo wurde 1969 von ESRI in Angriff genommen. Seit 1982 ist das System auf dem Markt. Von Anfang an war das Coverage-Modell für die Ver- und Bearbeitung von Vektoren Bestandteil des GIS Systems. ArcInfo gehört zu den sogenannten „loosely coupled systems“. Diese werden durch die integrierten Systeme („integrated architecture“) sprich Geodatenbanken sehr wahrscheinlich nach und nach abgelöst.

Erstere zeichnen sich dadurch aus, dass das räumliche Datenmanagement vom thematischen Datenmanagement voneinander getrennt ist. In ArcInfo beispielsweise ist das Arc-Modul für die Verwaltung der räumlichen Daten zuständig, welche mit nicht-räumlichen Daten (das sind in der Regel alphanumerische Daten) aus dem Info-Modul oder externen relationalen DBMS verknüpft sind. Über Zeiger wird die Verbindung zwischen Attributdaten und Geometriedaten hergestellt.

Zweitere bieten einen datenbankbasierenden Zugriff auf geographische Information.

Das geographische Objekt hat alphanumerische und räumliche Attribute – beide werden in der Datenbank verwaltet. Folgende Darstellungen sollen dies verdeutlichen:

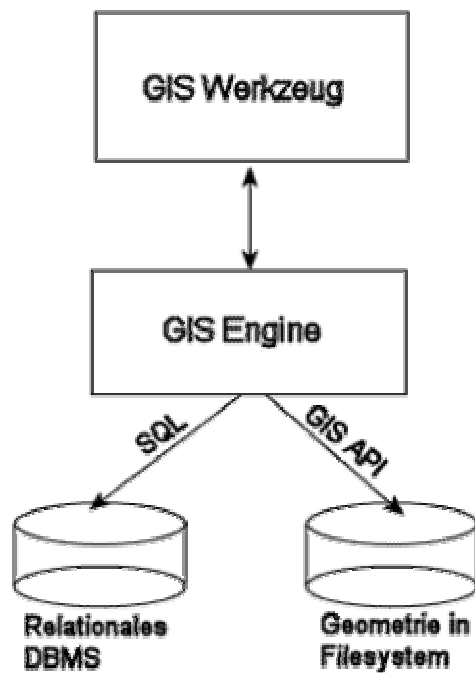


Abb.2.1: Schema eines GIS Systems der zweiten Generation – loosely coupled system (Geometrie in Dateien, Attributdaten in DBMS) (verändert nach [RSV02])



Abb.2.2: Schema eines GIS Systems der dritten Generation – integrated architecture („Geodatabase“)

2.2.2 Das Datenmodell

Es folgt eine kurze Beschreibung des Coverage-Datenmodells, beschränkt auf die wichtigsten Elemente (nach [ESR94]). Dies dient dem Verständnis der Problematik beim Umstieg auf das Geodatabase Modell.

Das Coverage-Datenmodell ist in der Lage drei grundlegende geometrische Elemente abzubilden, nämlich Punkt (point), Linie (arc), Fläche (polygon). Es gibt darauf aufbauend komplexe, d.h. aus diesen Grundformen zusammengesetzte Elemente; dies sind Routen (bestehend aus Linien) und Regionen (bestehend aus Flächen).

Punkt (point):

Diese werden als x,y Koordinatenpaare in einem .LAB file gespeichert. Die zugehörige Attributtabelle steht im .PAT (point attribute table) file. Als Punkte werden all jene Elemente gespeichert, deren Ausdehnung im logischen Modell nicht vorhanden oder zu klein ist, um als Fläche dargestellt zu werden (Beispiel: Bergspitze oder Trinkwasserquelle). Knoten sind eine spezielle Ausprägung von Punkten – sie markieren den Anfang bzw. das Ende einer Linie.

Linie (arc):

Ein Arc ist eine kontinuierliche Abfolge von x,y Koordinatenpaaren, welche sich von einem Anfangspunkt zu einem Endpunkt erstrecken. Als Arcs werden Inhalte dargestellt, die im logischen Modell als lineare Elemente definiert sind, wie Straßen oder Höhengichtenlinien. Gespeichert werden die arcs in .ARC Dateien. Die zugehörigen Attributtabellen stehen in .AAT (arc attribute table) Dateien.

Routen sind eine geordnete Sammlung von arcs oder Teilen von arcs welche lineare Elemente (wie z.B. U-Bahnlinien) repräsentieren. Informationen zu Routen werden in .RAT und .SEC files gespeichert, wobei diese lediglich Referenzen auf arcs enthalten und nicht die eigentliche Geometrie speichern.

Fläche (polygon):

Ein Polygon bezeichnet eine von mehreren Seiten umschlossene Figur. Sie wird durch mehrere Arcs definiert, welche die Ränder darstellen, und einem Bezeichnungspunkt (label point), welcher innerhalb der Figur liegt. Die Informationen zu jedem Polygon liegen in .ARC (Koordinaten und topologische Information der arcs), .PAT (Liste der das Polygon zusammensetzenden arcs), .CNT (Bezeichnungspunktnummern) und .LAB (Koordinaten der Bezeichnungspunkte) Dateien.

Regionen sind beschreiben zusammengesetzte Flächen, wie z.B. Staatsgebiete mit Exklaven (die Exklave ist eine räumlich getrennte, jedoch zum Staatsgebiet gehörige Fläche). Die zugehörigen Informationen werden in .RXP files gespeichert.

Diese Darstellung ist nicht vollständig, enthält jedoch die wichtigsten und für das Verständnis des Coverage-Modells notwendigen Erläuterungen. Anhand dieser Aufstellung

erkennbar, dass ein Coverage aus einer Vielzahl von Dateien besteht und zur Abbildung des Modells ein komplexes Gebilde vonnöten ist. Ein Coverage ist im Grunde ein Verzeichnis, welches eine Sammlung von Dateien enthält mit einer zugehörigen INFO Datenbank, die ebenfalls als Verzeichnis in einem gemeinsamen Arbeitsbereich (Workspace) liegt.

Der Arbeitsbereich enthält also das INFO Verzeichnis, welches (neben anderen) die Attributdaten eines Coverages enthält und das Coverage Verzeichnis, welches die Geometriedaten der einzelnen Elementtypen(Punkt, Linie, Fläche) enthält.

Von Bedeutung ist bei der Beschreibung des Coverage-Datenmodells anzumerken, dass grundsätzlich jeweils nur ein Elementtyp in einem Coverage gespeichert wird. D.h. dass zum Beispiel Fließgewässer und die zugehörigen Pegelmessstellen in zwei verschiedenen Coverages (Fließgewässer – Liniencoverage und Pegelmessstellen – Punktcoverage) gespeichert werden.

2.2.3 Das Gesamtsystem

Zusammengefasst kann das System wie folgt dargestellt werden [ZEI99].

Die primären Typen von Coverage Elementen sind Punkte, Linien (arcs), Polygone und Knoten. Diese Elemente können topologische Beziehungen aufweisen. Linien definieren den Umriss von Polygonen, Knoten markieren die Endpunkte von Linien, Punkte markieren die Innenseiten von Polygonen oder geographische Objekte mit kleiner räumlicher Ausdehnung.

Sekundäre Typen sind Tics (Registrierungspunkte), Links (Verschiebevektoren) und Annotations (Bezeichnungstexte).

Zusammengesetzte (komplexe) Typen sind Routen (Sammlungen von logisch zusammengehörigen Linien mit zugeordnetem Maßsystem) und Regionen (Sammlungen von logisch zusammengehörigen Flächen, die überlappend oder räumlich getrennt vorliegen können).

2.2.4 Vorteile und Nachteile

Die Vorteile des Coverages können kurz wie folgt dargestellt werden:

- Redundanzfreiheit durch topologische Strukturen. Als Beispiel können Verwaltungseinheiten dienen, wie Grundstücke, Katastralgemeinden, politische Gemeinden, Gerichtsbezirke, politische Bezirke, Bundesländer. Diese sechs Verwaltungseinheiten bauen auf ein- und demselben Satz von Linien (arcs) auf, nämlich den Grundstücksgrenzen. Aus diesen werden Regionen definiert, wel-

che jeweils nur Referenzen auf die zusammensetzenden Linienelemente enthalten. Mit anderen Worten: der Liniendatenbestand wird für die auf ihm basierenden Grenzen nur einmal physisch vorhanden. Die verschiedenen darauf aufbauenden Verwaltungsebenen sind jeweils Referenzen auf diesen Datenbestand.

- Systemimmanente Prüfung auf topologische Integrität. Die dem System eigenen Schritte zur Herstellung topologischer Strukturen führt zu einer stabilen Datenintegrität, weil Fehlermeldungen auf nicht bereinigte Datenbestände hinweisen.
- Einfacher und damit leicht zu verstehender Aufbau aus den geometrischen Grundelementen Punkt, Linie, Fläche aller Coverages.

Die Nachteile werden wie folgt beschrieben:

- Keine von vorne herein zugeordnete logische Zugehörigkeit. Datenelemente sind jeweils zu Datenbeständen mit demselben Elementtyp (Linie, Punkt, Fläche) zusammengefasst. Die Eigenschaften beispielsweise einer Linie, welche einen Fluß darstellt, ist identisch mit den Eigenschaften einer Linie, welche eine Straße darstellt. Die Unterscheidung zwischen Fluß und Straße ist nur durch ergänzende Attribute möglich.
- Komplexe Handhabung der Vielzahl von angelegten files, welche ein Coverage ausmachen.

2.2.5 Die Zukunft der Coverage

Seit der Version von ArcGIS 8.3 wird die Bearbeitung von Coverages in ArcMap nicht mehr unterstützt; die Darstellung und die Konversion in andere Datenbestände ist möglich. Gleichwohl wird von ESRI darauf hingewiesen, dass Geodatenbasen die Coverage nicht ersetzen, sondern ergänzen sollen [ZEI99].

Der Umstieg vom Coverage zur Geodatenbank solle demnach dann erfolgen, wenn die Vorteile der integrierten Datenverwaltung und die erweiterten Modellierungsmöglichkeiten den Aufwand der Migration übersteigen.

Dies deutet darauf hin, dass in den nächsten Jahren das Gewicht des Coverage-Modelles zugunsten des Geodatabase-Modelles verschoben werden soll.

2.3 **Exkurs: Topologie – das Argument für GIS**

Topologie wird in der Regel als Charakteristikum von GIS Systemen gesehen. Eine der ersten Ausführungen dieses mathematischen Ansatzes in der GIS-Welt wurde durch

das Coverage-Format realisiert. Es hat sich bis zum heutigen Tag als sehr leistungsfähiges Datenmodell bestätigt.

2.3.1 Was ist Topologie?

Topologie beschreibt – vereinfacht gesagt – „Nachbarschaftsbeziehungen“ von und zwischen räumlichen Objekten. Dabei bleiben geometrische Eigenschaften von untergeordneter Bedeutung. Als Beispiel mögen zwei Bushaltestellen (A, B) dienen, welche 500m voneinander entfernt liegen. Geometrisch gesehen liegt zwischen den Punkten A und B eine Strecke von 500m. Topologisch gesehen sind die Punkte A und B Nachbarn; es spielt keine Rolle, ob sie 500m oder 5000m voneinander entfernt liegen – sie bleiben Nachbarn.

Ausgangspunkt für die Topologie ist die Kanten-Knoten-Struktur. Diese bildet die Grundlage für drei bedeutende topologische Konzepte:

Verknüpfung (connectivity) – die arcs werden durch eine Liste von „von Knoten“ – „nach Knoten“ identifiziert.

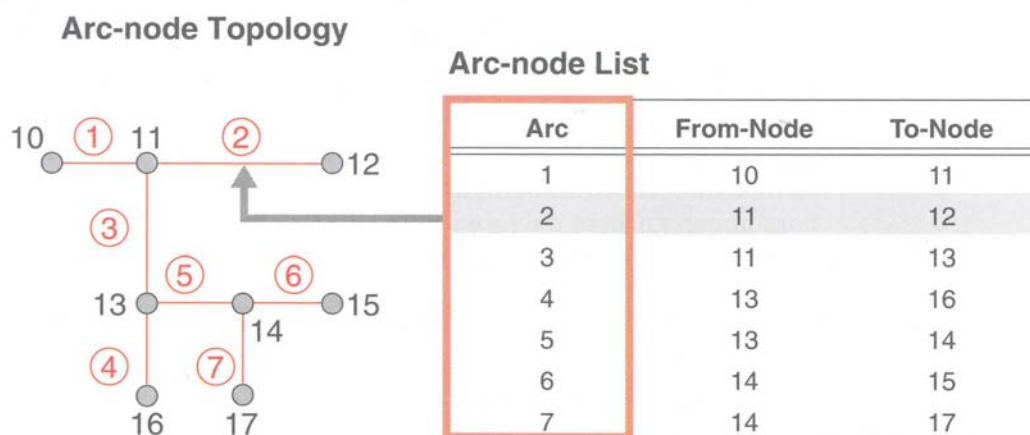


Abb. 2.3: Kanten-Knoten Topologie und der Bezug zur Tabellendarstellung.

Flächendefinition (area definition) – die polygone werden durch eine Liste von arcs definiert

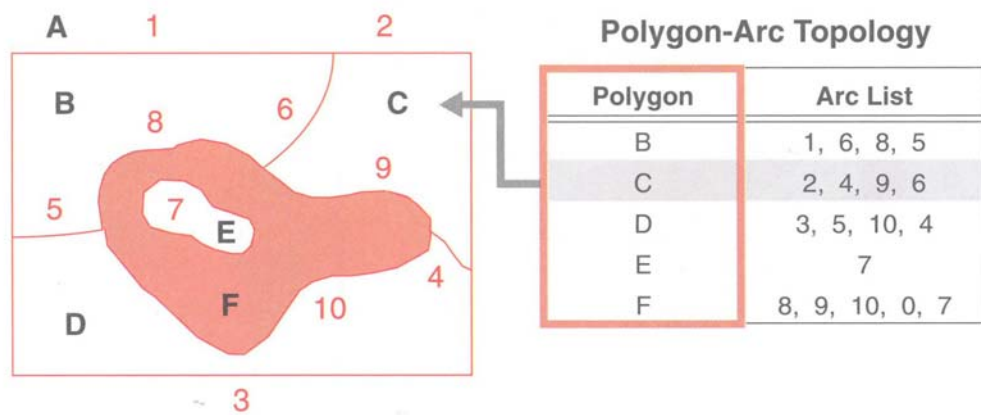


Abb. 2.4: Polygon-arc Topologie und der Bezug zur Tabellendarstellung.

Nachbarschaft (contiguity) – die Nachbarschaft (rechts oder links eines arcs) wird durch eine Liste von rechts und links liegenden Polygonen beschrieben.

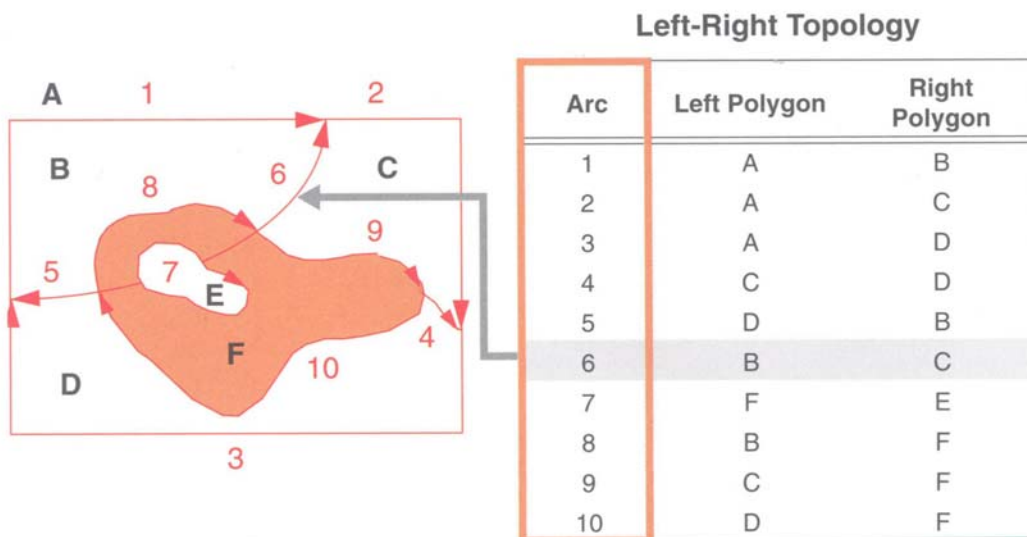


Abb. 2.5: Left-Right Topologie und der Bezug zur Tabellendarstellung.

2.3.2 Der Nutzen topologischer Strukturen

Folgende grundlegende Eigenschaften sind jedoch allen topologischen Strukturen gemeinsam [UGIS02a]:

- Redundanz-Minimierung durch „Mehrfachverwendung“ von (geometrischen) Primitiven.
- Hierarchischer Aufbau mittels Zeigern von komplexen Objekten auf (geometrische) Primitive.

- Ableitung von topologischen Eigenschaften (Kontiguität – Nachbarschaft, Konnektivität, Einschluss) aus absoluter geometrischer Lageinformation (Koordinaten).

ESRI baute daher bei der Entwicklung seines Coverage-Modells auf die „Kanten-Knoten-Struktur“. Das Prinzip dabei ist, dass Linien durch Knoten definiert werden und diese wiederum Polygone aufbauen. Dabei werden Ausgangs- und Endpunkt einer Linie und damit eine Richtung definiert. Aus der Festlegung der Richtung ergibt sich auch eine „Innen“ und „Aussen“ – Definition bei Flächen. Dieses Konstrukt führt dazu, dass es möglich ist, komplexe geometrische Elemente zu definieren, ohne geometrische Primitive redundant verwenden zu müssen.

Der Begriff der Topologie begegnet uns auch in der neuen Welt der „Geodatabase“, allerdings in anderem Zusammenhang.

Hier ist die Topologie nicht an ein physikalisches Datenmodell wie bei der Coverage gebunden. Sie wurde „verallgemeinert“, um mit großen, zusammenhängenden Datenbanken zu funktionieren und ist nicht mehr an das bekannte Coverage-Modell gebunden [ESR03]

2.3.3 Der Anlass für die Entwicklung neuer Technologien

Wie bereits oben erwähnt wurde durch den technischen Fortschritt bei der Entwicklung leistungsfähiger Hardware Bahn für die konzeptionelle Neugestaltung bei der Datenhaltung von GIS-Systemen gebrochen. Letzlich waren die technischen Möglichkeiten und die Suche nach einer besseren, dem logischen Datenmodell näher stehenden Methode zur Abbildung geographischer Entsprechungen, die Gründe für die Entwicklung des Geodatabase Modells. Die besonderen Aspekte geographischer Elemente sollte nicht allein durch attributive Eigenschaften, sondern durch dem Objekt inhärente Ausprägungen Berücksichtigung finden.

Ausserdem sollte der Dualismus zwischen dem file, welches die geometrischen Elemente in binärer Form enthält, und der Datenbank, die Attributinformationen speichert, ein für alle mal der Vergangenheit angehören.

2.4 Das Geodatabase Konzept

Diese Konzept ermöglicht aber auch eine weitere grundlegende Weiterentwicklung. Sie läuft darauf hinaus, dass GIS Datenbestände dadurch „intelligenter“ gemacht werden sollen, indem man sie mit „natürlichem Verhalten“ ausstattet und beliebige Beziehungen zwischen ihnen definierbar sind [ZEI99].

2.4.1 Die Abbildung bisheriger GIS Funktionalitäten in der Geodatabase

Im Grunde bleiben die Funktionalitäten, wie sie die Coverage bereitgestellt hat, erhalten, weil es sich lediglich um eine Veränderung des Datenmanagements handelt.

Einer der vielen Kommentare welcher im Internet nachzulesen ist hält fest, dass es sich konzeptionell bei der Geodatabase um nichts anderes als bei der Coverage oder dem Shape handelt. Die Datenmodelle würden lediglich um Funktionalitäten erweitert [BOZ04].

Dieser Einschätzung ist nur eingeschränkt zu folgen; es handelt sich bei der Geodatabase um ein Konzept, welches Funktionalitäten aus der Welt der filebasierten Datenhaltung fortführt, aber eine gänzlich andere Technologie verwendet und hinter welcher eine neue Philosophie steckt.

Natürlich bestehen weiterhin Möglichkeiten der Bearbeitung von geographischen Datenbeständen, so zum Beispiel die topologische Bereinigung von Datenbeständen und weitere Funktionen; diese werden jedoch auf andere Art und Weise realisiert und laufen vor einem Hintergrund der Datenhaltung ab, der sich grundlegend verändert hat.

2.4.2 Neue Funktionalitäten einer Geodatabase

Hierzu gehören u.a. zweifellos die verbesserten Editiermöglichkeiten, welche es erlauben, die Eingabe von Daten nur innerhalb bestimmter, vordefinierter Wertebereiche stattfinden zu lassen. Wenngleich diese Funktionalität bei weitem einen nicht so großen Stellenwert einnimmt, wie uns die Aussagen der Geodatabase-Macher glauben machen wollen, so unterstützt diese Funktionalität doch die Definition von Datenbeständen, welche sich mehr an thematischen Qualitäten, als an graphischen Elementen orientieren: Wertebereiche, Wertetabellen, Standartwerte.

GIS Datenbestände werden „schlauer“. Indem die geographischen Daten mit im Datenmodell abgebildeten Eigenschaften der realen Welt ausgestattet werden, können GIS Nutzer intuitiver mit den Objekten arbeiten. Zudem ermöglicht das „intelligente“ Verhalten genauere Dateneingabe und –änderung.

Genauere Abbildung der Form eines Objektes. Die der Geodatabase eigene Geometrie unterstützt komplexe Formen (Bezier-Kurven, echte Kreisbögen), bietet erweiterte Möglichkeiten der Koordinateneingabe und verbesserte Schnappfunktionalität.

Vor allem aber findet der Anschluß an eine einheitliche und IT konforme Speicherung geographischer Daten statt. Sowohl Vektoren, als auch Raster und TIN-Daten können im selben Verwahrungsort erstellt, gespeichert und verwaltet werden.

Viele Nutzer können gleichzeitig geographische Daten bearbeiten. Das Datenmodell erlaubt Arbeitsabläufe, bei welchen viele Bearbeiter auf denselben Bestand zugreifen,

diesen verändern und möglicherweise auftretende Konflikte benennen und bereinigen können. Dieser Prozess der Versionierung unterstützt somit langandauernde Transaktionen, welche in GIS-Projekten regelmässig auftreten [BOZ04].

Einer genaueren Betrachtung bedarf das Kapitel Topologie in der Geodatabase (nachdem wir ihm schon vorher beim Coverage breiteren Raum zugestanden haben).

Die Topologie begegnet uns auch im neuen Datenmodell, allerdings in einer ganz anderen Art und Weise.

Der Hauptunterschied zwischen Topologie hier (Geodatabase) und da (Coverage) liegt darin, dass Topologie insoweit eine Verallgemeinerung erfahren hat, als sie großen zusammenhängenden Datenbanken eingesetzt werden kann und nicht mehr allein an ein physikalisches Datenmodell gebunden ist. Sie ist daher nicht mehr obligatorisch, sondern fakultativ.

Das Geodatabase-Modell definiert ein allgemeines Modell für geographische Information. Es ermöglicht die Definition von Beziehungen zwischen Objekten, ebenso wie die Definition von Regeln für die Sicherstellung der referentiellen Integrität zwischen den Objekten.

Folgende Charakteristika lassen sich für Topologien in einer Geodatabase festhalten [nach ESR03]

- Sie existieren innerhalb von Element-Datensätzen (feature data sets), weil alle teilhabenden Elementklassen (feature classes) denselben räumlichen Bezug aufweisen müssen.
- Es können mehrere Topologien innerhalb eines Element-Datensatzes vorkommen
- Elementklassen (feature classes) können nur an einer Topologie teilhaben
- Elementklassen (feature classes) können nicht gleichzeitig an einer Topologie und einem geometrischen Netzwerk beteiligt sein.
- Die Topologie kann mehrere Punkt-, Linien-, und Flächenelementklassen enthalten.

Ausserdem kennzeichnen weitere Parameter die Definition der Topologie; das sind

- Cluster-Toleranz
- Rangreihung von Elementklassen
- Regeln (für die Erstellung von Topologien)

Die Cluster-Toleranz:

Zur Umsetzung topologischer Strukturen wird die Definition einer Toleranz vorausgesetzt. Diese ist jene Distanz innerhalb welcher benachbarte Scheitelpunkte (vertices) oder Grenzen als identisch oder überdeckend betrachtet werden. Scheitelpunkte bzw. Endpunkte, welche sich innerhalb der Cluster-Toleranz voneinander entfernt befinden, werden während eines Topologie-Prüfprozesses (Validate Topology) zu einem Punkt zusammengezogen. Die Cluster-Toleranz kann zur Verschiebung von Elementen und daher zur unerwünschten Veränderung der Geometrie führen, wenn sie falsch (daher zu groß) gewählt wird.

Rangreihung:

Die Rangreihung wird für eine Elementklasse (feature class) festgelegt und bestimmt, welche Elemente im Fall eines Topologie-Prüfprozesses ihre Lage relativ zu anderen Elementen verändern und welche nicht. Je besser der Wert der Rangreihung, desto weniger werden die Elemente in ihrer Lage verändert. Elementklassen mit der Rangreihung 1 werden bei einer Topologie-Prüfung als unveränderlich erkannt. Elementklassen mit den Rangreihungswerten 2 oder 3 werden zu den Elementen mit der Rangreihung 1 hingezogen oder allgemeiner ausgedrückt: wenn Elemente innerhalb der Cluster-Toleranz liegen, werden jene Elemente mit dem geringeren Rangreihungswert zu den Elementen mit dem höheren Rangreihungswert verschoben.

Wenn zwei Elementklassen denselben Wert für die Rangreihung aufweisen, wird ein Mittelwert für die Lageveränderung berechnet (die beiden Elemente treffen sich sozusagen in der Mitte).

Topologie-Regeln:

Jede Topologie in der Geodatenbank ist mit einer oder mehreren Topologie-Regeln verknüpft. Der Benutzer sorgt für die Integrität der Topologie durch die Definition von Regeln. Die Topologie-Regel beschreibt einen Zustand in der Topologie, bei welchem die Integrität verletzt werden könnte. Bei der Topologie-Prüfung (Validate Topology) wird der Datenbestand anhand der Regeln geprüft.

Beispiel:

- Katastergrundstücke dürfen sich nicht überlappen
- Hausgrenzen müssen innerhalb von Grundstücksgrenzen liegen

Jede Verletzung der Topologie-Regeln wird markiert und somit eine Bereinigung durch den Nutzer ermöglicht. Eine automatische Bereinigung ist nicht vorgesehen.

Damit ergeben sich drei Möglichkeiten für die Behandlung von aufgefundenen Regelverstößen:

- Ignorieren des Regelverstoßes
- Fehlerbereinigung durch Editieren des Datenbestandes
- Regelverstoß als Ausnahme markieren

Weitere bisher nicht enthaltene Funktionalitäten fallen ebenso in den (ausgesprochen wichtigen) Bereich der Pflege von Datenbeständen [ESR04SDE].

- Lange Transaktionen (long transactions): Im Gegensatz zu rein alphanumerischen Änderungen in DBMS (wie beispielsweise im Handel oder in der Finanzwirtschaft) betreffen Manipulationen in GIS Datenbeständen eine Vielzahl von Datensätzen über längere Zeiträume. Hier müssen besondere Maßnahmen gesetzt werden, um die Integrität der Daten nicht zu gefährden
- Mehrfache Editiersitzungen (multiple edit sessions): Die Aktualisierung von GIS Datenbeständen kann längere Zeiträume (Tage oder Wochen) in Anspruch nehmen.
- Gleichzeitiges Editieren durch mehrere Benutzer (multiuser editing): Viele Mitarbeiter ändern gleichzeitig jeweils ihren Datenbestand, welcher unmittelbar benachbart zu anderen Änderungsgebiete liegt, oder sich sogar mit diesen überschneidet.
- Abgemeldete und angemeldete Abläufe (check out – check in transactions): Dies ist vor allem dann der Fall, wenn Benutzer „ihren“ Teilbereich von Daten aus dem Gesamtbestand herauslösen und getrennt von diesem Bearbeiten (beispielsweise bei Feldarbeiten).
- Historisierung und Versionierung: Aus Dokumentationsgründen werden die verschiedenen Stände einer Änderung abrufbar gehalten. Somit kann jederzeit der Stand zu einem bestimmten Zeitpunkt rekonstruiert werden.
- Aktualisieren von Datenbeständen (change-only updates): Über XML-Schnittstellen können Datenbestände ausgetauscht werden.

- Replikation von Datenbeständen: Austausch von Datenbeständen in replizierten (d.h. kopierten) Datenbanksystemen.

2.4.3 Die Zukunft der Geodatabase

Nach dem heutigen Entwicklungsstand ist davon auszugehen, dass sich das Geodatabase-Modell als Standard etablieren wird. Dies nicht nur wegen der bereits genannten Vorteile der vereinheitlichten Datenspeicherung, sondern vor allem wegen der Weiterentwicklung von Datenbankanwendungen und dem weiter steigenden Funktionsumfang derselben (Stichwort „räumliche Erweiterung“).

Präziser ausgedrückt sind es folgende Faktoren, welche der Geodatenbank zu rascher und weiter Verbreitung verhelfen werden:

- Zunehmende Integration von GIS in die Standard-IT-Landschaft. Die ursprüngliche Sonderstellung von GIS in einer IT Nische ändert sich. Standard-IT Produkte, Methoden, Standards halten Einzug und führen zu einer neuen Dynamik im GIS Bereich.
- Entwicklung von räumlich erweiterten DBMS (spatial enabled DBMS). Die Funktionalitäten entwickeln sich sowohl hinsichtlich der Datenbanktechnologie, als auch bezüglich der Zugriffs- und Abfragemöglichkeiten weiter (Stichwort: SQL Standard für die Abfrage räumlicher Elemente).
- Standardisierte Geodatenmodelle (content standards)
- Ausbau der Client/Server Architekturen, besonders im Hinblick auf webbasierte Technologien und die damit zweckmässigerweise verbundene Verwendung von Geodatenbanken.

Nichtsdestotrotz soll an dieser Stelle festgehalten werden, dass vor allem eingefleischte Coverage-Nutzer insbesondere mit den geänderten Topologie-Eigenschaften Mühe haben. Die Verantwortung des Bearbeiters nimmt nämlich mit der Anzahl an Bearbeitungs- bzw. Manipulationsmöglichkeiten des Datenbestandes zu. Immer mehr Kenntnisse und Fähigkeiten werden verlangt oder den Zukauf ergänzender Expertise durch zusätzliches Personal oder vermehrte Beratungsleistungen.

2.5 Zusammenfassung

Ein Modell ist ein auf ihre Grundfunktionen reduziertes Abbild der realen Welt. Ein Datenmodell beschreibt die Elemente des Systems, welches das Abbild (meist eines Ausschnittes) der realen Welt beschreibt.

Im GIS Bereich treten Datenmodelle in drei unterschiedlichen Ausprägungen in Erscheinung: im CAD-, im Coverage- und im Geodatabase Modell.

Das CAD-Modell ist weitgehend bedeutungslos, während das Geodatabase Modell das Coverage abzulösen beginnt.

Im Coverage werden die geometrischen Inhalte in einer Datei binär, die attributiven Inhalte in einer Tabelle alphanumerisch abgelegt. In der Geodatabase werden beide Inhalte (geometrische und alphanumerische) in Datenbanktabellen gespeichert. Erstere werden als „loosely coupled systems“, letztere als „integrated architecture“ bezeichnet. Beide Technologien weisen zur Gewährleistung der Integrität von Daten topologische Strukturen auf. Die Topologie beschreibt Nachbarschaftsbeziehungen zwischen räumlichen Objekten. Ausgangspunkt für die Topologie ist die sogenannte Kanten-Knoten-Struktur. Dabei werden geographische Elemente als Punkt, Linie oder Fläche abgebildet. Die Suche nach einer, dem logischen Datenmodell näher stehenden Methode zur Beschreibung geographischer Entsprechungen führte zur Entwicklung des Geodatabase-Modells: GIS Datenbestände sollen dadurch intelligenter und mit natürlichem Verhalten ausgestattet werden können. Ausgangspunkte für die Topologie in einer Geodatabases sind die Parameter Cluster-Toleranz, Rangreihung und Topologie-Regeln. Nach dem heutigen Entwicklungsstand ist davon auszugehen, dass sich das Geodatabase-Modell als Standard etablieren wird. Durch den erweiterten Umfang an Funktionalitäten und das mächtigere (d.h. komplexere) Methodengebilde werden vom Nutzer immer mehr Kenntnisse und Fähigkeiten verlangt.

3 Der Unterschied zwischen Coverage und Geodatabase

Der Unterschied zwischen Coverage und Geodatabase, insbesondere das Umfeld betreffend, sollen an dieser Stelle noch einmal zusammengefasst werden.

Auf den ersten Blick scheint die in einer Frage formulierten Überschrift leicht zu beantworten.

„Worin besteht der Unterschied zwischen Coverage und Geodatabase?“

Das Coverage besteht aus zwei Verzeichnissen, wovon eines die geographischen Daten in binär codierter Form in mehreren Dateien enthält und das andere die Attributdaten in verschiedenen Datenbankfiles.

Die Geodatabase basiert auf einer relationalen Datenbank. Die Inhalte, welche die geographischen und die attributiven Daten beschreiben, werden in Datenbanktabellen abgelegt. Dadurch ist sie leistungsfähiger, aber auch komplexer.

Mit diesen beiden kurzen Formulierungen sind jedoch eine ganze Reihe von Folgen verbunden, die den großen Unterschied zwischen Coverage und Geodatabase bedingen. Die GIS-Welt bestand bis vor wenigen Jahren aus Anwendern, welche über ihr Fachgebiet zur Arbeit mit Geographischen Informationssystemen gekommen sind. Heute bewegt sich die Entwicklung sehr stark in den Bereich der reinen IT-Anwendungen.

Oder noch konkreter formuliert:

Während vor wenigen Jahren die Kenntniss der jeweiligen GIS-Software genügte, um ein System bedienen zu können, sind neuerdings Netzwerkkennnisse, Datenbankwissen, das Erlernen einer oder mehrerer Programmiersprachen und mehr undabdingbar. Diese Entwicklung bewirkt weiterhin, dass die IT Infrastruktur sowohl im Hard- und Softwarebereich, als auch bei der personellen Ausstattung immer anspruchsvoller wird. Damit wird es kleineren Unternehmen oder Organisationen immer schwerer gemacht, der Entwicklung zu folgen.

Nachfolgend werden jene Unterschiede genauer beschrieben, welche in der IT-Infrastruktur in Erscheinung treten und die Nutzung wesentlich beeinflussen. Dabei wird eine Konfiguration vorausgesetzt, wie sie in vielen Systemen als Standard bezeichnet werden kann.

Als von grundlegender Bedeutung ist dabei die Einführung des Objektorientierten Ansatzes der Geodatabase in die GIS-Welt.

Die Vorteile der Geodatabase werden wie folgt beschrieben [ZEI99]:

- Einheitlicher Verwahrungsort geographischer Daten
- Ermöglichung verbesserter Dateneingabe und -pflege
- Benutzer arbeiten mit der Realität näher kommenden Objekten
- Elemente weisen einen reicheren inhaltlichen Zusammenhang auf
- Es können bessere Karten erstellt werden
- Elementformen können realitätsnäher beschrieben werden
- Kontinuierliche Datenbestände großer Ausdehnungen
- Gleichzeitige Bearbeitung von Datenbeständen durch mehrere Nutzer

Die Unterschiede in den Anforderungen lassen sich demnach wie folgt formulieren:

3.1 Hardware

Coverage: Arbeitsstation für die Installation von ArcGIS/ArcInfo und ein Fileserver, welcher die zentrale Verwaltung der Datenbestände ermöglicht.

Geodatabase: Arbeitsstation für die Installation von ArcGIS/ArcInfo und ein SDE Server mit einem Datenbankserver, wobei SDE und DBMS auf demselben Server oder auf unterschiedlicher Hardware installiert sein können. Da Datenbank Anwendungen sehr komplex sind, ist die benötigte Rechenkapazität im oberen Leistungsbereich anzusiedeln.

3.2 Software

Coverage: ArcGIS/ArcInfo Lizenz auf der Arbeitsstation.

Geodatabase: ArcGIS/ArcInfo Lizenz für die Arbeitsstation, Lizenz für ArcSDE, Lizenz für DBMS.

3.3 Daten

Coverage: Die Coverage besteht aus zwei Verzeichnissen, wovon eines die geographischen Daten in binär codierter Form in verschiedenen Dateien enthält und das andere die Attributdaten in diversen Datenbankfiles.

Datenmodellierung: Zuordnung der Inhalte zu geometrischen Primitiven wie Punkt, Linie, Fläche. Zuordnung der im Datenmodell vorgesehenen Attribute.

Geodatabase: Die Geodatabase basiert auf einer relationalen Datenbank. Die Inhalte, welche die geographischen und die attributiven Daten beschreiben, werden in Datenbanktabellen abgelegt. Datenverwaltung bietet und erfordert mehr Möglichkeiten;

Datenmodellierung: Zuordnung der Inhalte zu geometrischen Primitiven wie Punkt, Linie, Fläche. Zuordnung der im Datenmodell vorgesehenen Attribute. Festlegung von Regeln für die Manipulation der Daten, festlegen von Wertebereichen (Attributdomänen).

3.4 Personal

Abhängig von der Tätigkeit – ausgehend von GIS Sachbearbeiter, welcher Datenerstellungs- und verwaltungsaufgaben wahrnimmt.

Coverage: Kenntnisse der Anwendung von ArcGIS/ArcInfo.

Geodatabase: Kenntnisse der Anwendung von ArcGIS/ArcInfo, Datenbank-, ArcSDE, Programmier-, Metasprachenkenntnisse.

Das Coverage ist prozessorientiert und besteht aus zwei Datenverzeichnissen, die einfach in der Anwendung sind. Die Weiterentwicklung stellt die objektorientierte Geodatabase dar, die auf einer relationalen Datenbank basierend. Das System ist komplexer, leistungsfähiger und stellt höhere Anforderungen an den Nutzer.

Die Unterschiede der beiden Datenmodelle im Überblick:

	Coverage	Geodatabase
Hardware	Arbeitsstation mit ArcGIS/Arc Info Fileserver mittlere Rechenkapazität	Arbeitsstation ArcGIS/Arc Info SDE Server und DBMS hohe Rechenkapazität
Softwarelizenzen	ArcGIS/ArcInfo	ArcGIS/ArcInfo ArcSDE Relationales DBMS
Datenverwaltung	geographische und attributive Daten an getrennten Speicherorten	geographische und attributive Daten in einem Datenspeicher
Datenmodellierung	Zuordnung der Elemente zu geo-metrischen Primitiven Zuordnung der Attribute	Zuordnung der Elemente zu Objekten, Objekte weisen Eigenschaften auf (Attribute) und Funktionen, welche auf sie angewendet werden können (Methoden)
Personelle Anforderungen	Anwenderkenntnisse ArcGIS/ArcInfo	Anwenderkenntnisse ArcGIS/ArcInfo Kenntnisse in Datenbanken, ArcSDE, Programmier- und Metasprachen

4 Der Umstieg in der Praxis – Methodik

4.1 Die Sicht des Herstellers

4.1.1 Dokumentation und Whitepapers

Kochrezepte – das ist, was sich jeder in den täglichen Wirren und Verstrickungen der IT gefangene Techniker wünscht. Die Software-Industrie muss sich darauf einstellen und bedient die Kundschaft entsprechend. Naturgemäß kann eine Anleitung nur grundsätzliche Vorgangsweisen enthalten und diese beispielhaft darstellen, weil die zu bearbeitenden oder zu konvertierenden Datenbestände zu unterschiedlich sind.

Dementsprechend umfangreich sind die Dokumentationen zum Thema Erstellung einer oder Umstellung auf eine Geodatenbank.

Grundsätzlich werden in der Literatur drei Möglichkeiten aufgezeigt, eine Geodatabase anzulegen:

1. Ein neues Schema in ArcCatalog erstellen
2. Bestehende Daten importieren
3. Mit Hilfe von CASE (Computer Aided Software Engineering) Tools

Es ist festzustellen, dass bei allen drei Varianten nach der Erstellung der Geodatabase im ArcCatalog Nachbearbeitungen notwendig werden. Wie schon zuvor erwähnt, müssen Topologien erstellt, Regeln definiert, Subtypen und Attributdomänen eingetragen werden. Die konkrete Vorgangsweise wird nachfolgend bei der Behandlung der Musterbeispiele anhand der oben genannten drei Möglichkeiten aufgezeigt.

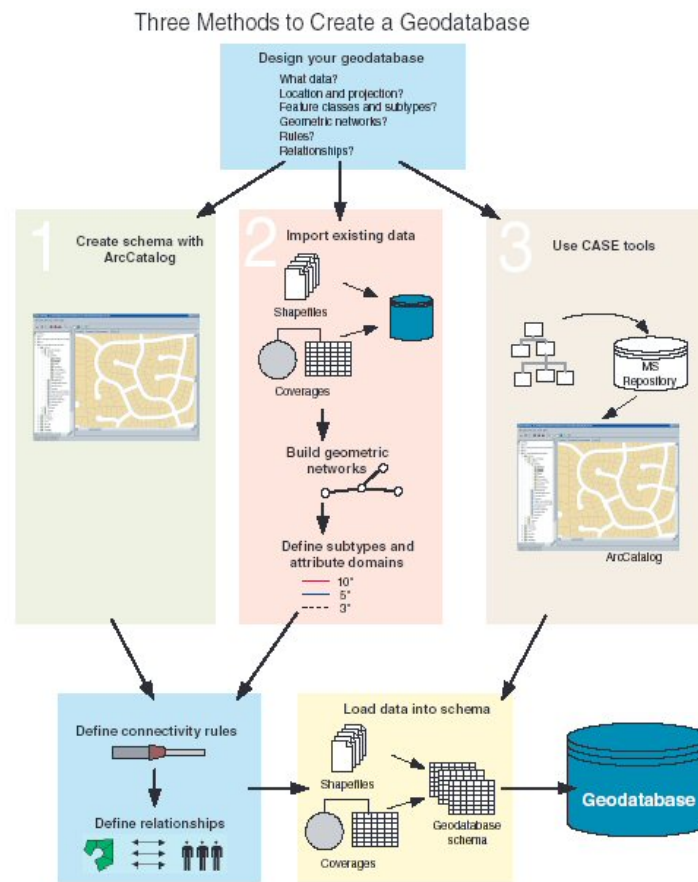


Abb. 4.1: Drei Methoden zur Erstellung einer Geodatabase (aus[ZEI99]).

Zu 1:

Bei der ersten Variante wird eine Geodatenbank von Grund auf in ArcCatalog neu aufgebaut. Es werden die Feature Datasets, die Feature Classes (sie enthalten die einzelnen thematischen Ebenen), die ergänzenden Relationen, die Topologieregeln und andere Vorgaben erstellt und die Datenbank dann mit entsprechenden Informationen befüllt.

Zu 2:

Vorhandene Daten (shapes, Coverages, ...) werden über Konvertierungswerkzeuge in das Geodatabase-Format überführt. Nach der Übernahme der Daten in die entsprechenden Feature classes (bzw. Layer) müssen noch die für die Erstellung der Topologie wichtigen Regeln definiert und formuliert werden, ebenso wie die Festlegung von Subtypen und Attributdomänen.

Zu 3:

Der interessanteste Ansatz zur Erstellung einer Geodatenbank ist jener mit Hilfe von CASE Tools. Hier werden Maßgeschneiderte Objekte definiert und in ein Datenbank-

schema integriert. Dabei wird zunächst mit Hilfe von MS VISIO oder Rational Rose ein UML Diagramm zu den einzelnen Objekten erstellt, dieses dann in ein XML Austauschformat (XMI ... XML Metadata Interchange) gebracht, in einer COM fähigen Programmiersprache in Code umgewandelt und dann im Datenbankschema implementiert.

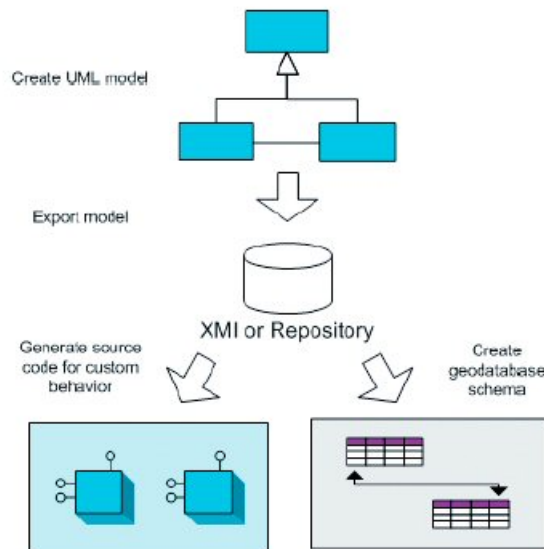


Abb. 4.2: Schema für die Erstellung einer GeoDB mit CASE tools (aus[ESR02bg]).

Die oben angeführten Anleitungen enthalten vor allem technische Vorgangsweisen (also hands-on Anleitungen) für die Umsetzung der Anlage einer Geodatabase. Wir werden diese nachfolgend noch genauer, nämlich bei der Umsetzung des Konzeptes in eine Geodatabase betrachten.

Die konzeptionelle Vorgangsweise wird an anderer Stelle und dabei weit weniger detailliert beschrieben, weil sie sehr datenspezifisch und daher nur schwer zu verallgemeinern ist.

Doch zunächst zu einem Schritt, der noch vor der „technischen“ Umsetzung der Geodatenbank-Erstellung durchzuführen ist, nämlich dem Design der Geodatenbank oder genauer gesagt die Modellierung der Datenbank.

Dieses Kapitel soll mit einer Standard-Vorgangsweise begonnen werden, nämlich dem „Geodatabase design guide“ [ZEI99]. Hier wird der Transformationsprozess in fünf Schritten wie folgt beschrieben:

1. Modellieren von Daten aus Benutzersicht (Model the user`s view of data)
2. Definieren von Objekten und Beziehungen (Define objects and relationships)
3. Auswahl geographischer Entsprechungen (Select geographic representation)
4. Abgleich von Geodatenbank-Elementen (Match to geodatabase elements)
5. Geodatenbank-Struktur organisieren (Organize geodatabase structure)

Ergänzend dazu sei bemerkt, dass diese empfohlene Vorgangsweise schon für die Modellierung von Coverages vorgeschlagen wurde. Sie ist offensichtlich so allgemein gehalten, dass sie auch heute noch (in diesem Fall für die Ausarbeitung von Datenmodellen für Geodatenbanken) Gültigkeit hat. Die Vorgangsweise wurde wie folgt beschrieben [ESR94]:

1. Modellieren von Daten aus Benutzersicht (Model the users` view)
2. Definieren von Einheiten und ihren Beziehungen (Define entities and their relationships)
3. Identifizieren der Entsprechungen von Einheiten (Identify representation of entities)
4. Abgleich mit dem ARC/INFO Datenmodell (Match to ARC/INFO data model)
5. In geographische Datensätze zusammenfassen (Organize into geographic data sets)

Die Erkenntnis aus dieser Gegenüberstellung ist, dass die Modellierung von Datenbeständen, ob nun Coverage oder Geodatabase grundsätzlich ähnlichen Vorgangsweisen gehorchen mag, diese aber dennoch sehr allgemein gehalten sind. Die Modellierung ist sehr stark von den zu behandelnden Datenbeständen abhängig und deshalb sehr stark an diese anzupassen.

Um einer geeigneten an einer „best practise“ angelehnten Vorgangsweise näher zu kommen werden nachfolgend Beispiele für die Umsetzung untersucht. Diese use cases sollen die Vorgangsweise besser verstehen helfen.

4.1.2 Case studies

ESRI stellt den Benutzern eine Anzahl an Case studies über die Anlage von Geodatabases im Internet zur Verfügung. Allen diesen Dokumentationen ist gemeinsam, dass sie die Gründe für die Erstellung, die technischen Grundlagen und Informationen über die Organisation zur Verfügung stellen. Der eigentliche Entstehungsprozess von Geodatenbanken, d.h. die Modellierung von Datenbanken und die Umsetzung des Modells in eine produktive Geodatenbank werden jedoch nicht oder nicht im notwendigen Detaillierungsgrad dargestellt. Diese Informationen werden vor allem in den technischen Dokumentationen angeboten, welchen in dieser Arbeit dargestellt wurden bzw. noch werden. Dennoch darf ich auf die entsprechende Informationsquelle hinweisen. Sie kann in jedem Fall dazu dienen, Erstinformationen zu erhalten: siehe [\[ESR04CS01\]](#), [\[ESR04CS02\]](#)

4.2 Planung: technische Aspekte des Umstiegs

Die technische Umsetzung des Umstieges der filebasierten Datenhaltung zur Geodatabase umfasst einige wichtige Bereiche. Das sind einerseits Hard- und Software, andererseits die Aufbereitung der Daten, sprich die Definition von konzeptionellem, logischem und letztlich physikalischem Modell. Der Schwerpunkt der nachfolgenden Ausführungen liegt eindeutig auf dem konzeptionellen Modell. Hard- und Softwareausstattung hängen einerseits von der jeweiligen EDV-Umgebung des Anwenders und dessen Ansprüchen ab und sind in weiten Bereichen vorgegeben. Viel wichtiger für die Planung ist in diesem Zusammenhang die entsprechende Vorbereitung der Daten, sprich das konzeptionelle Modell und der logische Datenbankentwurf. Auf diesen technischen Aspekt werden wir hier genauer eingehen.

Nach der Erläuterung der Datenmodellierung wird eine Vorgangsweise zur Erstellung logisch zweckmässiger Datenmodelle beschrieben um anschliessend näher auf UML (Unified Modelling Language) einzugehen, welche als Hilfestellung für die Definition eines Datenmodells dienen soll.

Zweck der Beschäftigung mit der Datenmodellierung ist die Erarbeitung einer tragfähigen Basis für den Aufbau des Informationssystems.

Wegen des hohen Zeitaufwandes und der geringen direkt umsetzbaren Ergebnisse wird dem Entwurfprozess generell zu wenig Augenmerk geschenkt. Diese Vernachlässigung kann schwerwiegende Nachteile mit sich bringen:

Eine Datenbank, welche die in sie gestellten Ansprüche nicht erfüllt weil z.B. Daten doppelt oder gar nicht vorhanden sind, Daten unzweckmässig repräsentiert sind, Aktualisierungsmaßnahmen nicht berücksichtigt sind u.ä., verursacht hohe Kosten und wird von den Nutzern nicht akzeptiert [ZEI99].

4.2.1 Theorie der Datenmodellierung

Der Begriff des Datenmodelles steht zwar in der Umsetzung nicht im Vordergrund, spielt aber für die Vorarbeiten und die Ausarbeitung einer künftigen Struktur eine bedeutende Rolle.

„Ein **Datenmodell** (DM...data model) ist ein formales Konzept zur Beschreibung von Datenbankstrukturen, die die Informationsstrukturen des UoD (Anm.: Universe of Discourse = Informationsbereich, den die Datenbank darstellt) repräsentieren. Ein Datenmodell ist somit eine Beschreibungsform für die aus dem konzeptuellen Schema ableitbaren logischen Datenbankstrukturen.“ [SW01]

Konkret ausgedrückt heisst dies:

- Das Ziel einer Modellierung ist die Abbildung der realen Welt in einem Modell.

- Durch thematische Einschränkung wird ein Teilbereich gewählt, welcher Grundlage für das konzeptionelle Modell ist.
- Das konzeptionelle Modell beschreibt die einzelnen Objekte der realen Welt, ihre Beziehungen untereinander und ihr Verhalten über Attribute und Methoden.
- Das logische Modell wird aus dem konzeptionellen Modell entwickelt und zwar für eine bekannte Systemumgebung.
- Das physikalische Datenmodell entsteht durch Speicherung der Daten entsprechend dem logischen Modell im System.

Der Ablauf kann wie in nachfolgender Abbildung skizziert werden.

3- Stufenmodell der Modellierung

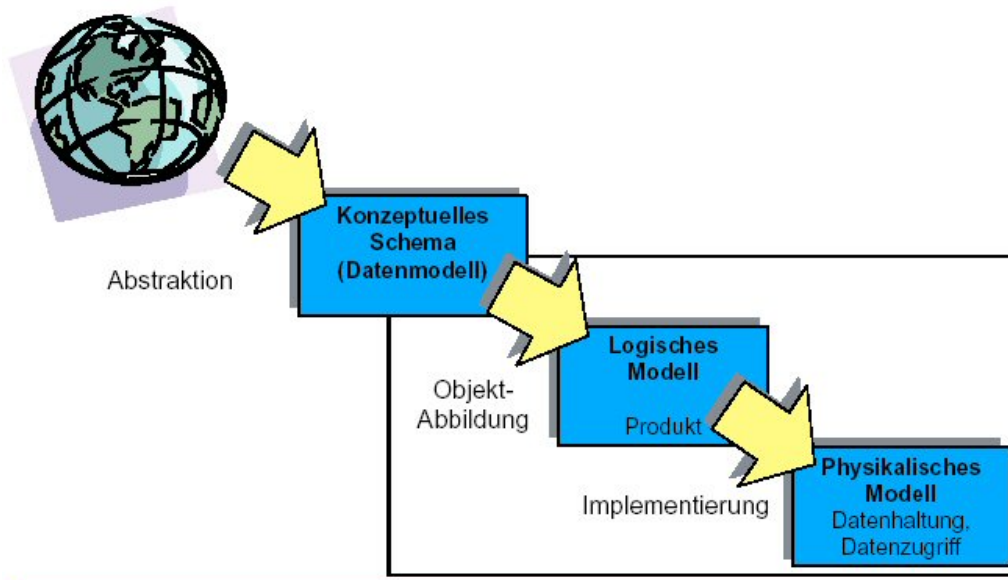


Abb. 4.3: Dreistufenmodell der Datenmodellierung [SCH103a]

Der logische Datenbankentwurf führt dann seinerseits durch Transformation zum eigentlichen Zieldatenmodell; heutige kommerziell verfügbare Systeme lassen eine direkte Überführung nicht zu. Dies kann dann ein relationaler oder ein objektorientierter Datenbankentwurf sein.

Gebräuchlich ist heute eine abgestufte Vorgangsweise, mit welcher in einem Abstraktionsprozess die Abbildung von Inhalten aus der realen Welt in einer Datenbank ermöglicht werden.

Neben den oben genannten wird noch ein weiteres bedeutendes Datenmodell unterschieden, welches zur Kategorie der Semantischen Datenmodelle gehört. Der wichtigste Vertreter der Semantischen Datenmodelle ist das Entity-Relationship-Modell (ERM).

Als Modellierungstechniken haben sich die ER-Modellierung (Entity-Relationship-Modellierung) und die UML-Notation (Unified Modelling Language) durchgesetzt [DFH03].

Die Unified Modeling Language (UML) ist eine Modellierungssprache zur Beschreibung von Softwaresystemen. Im Unterschied zu der Vorgängermethoden fehlen der UML eine Notation und Beschreibung für Prozesse. Der Grundgedanke bei der UML bestand darin, eine einheitliche Notation für viele Einsatzgebiete zu erstellen. Die UML dient der Beschreibung von Datenbankanwendungen, Echtzeitsystemen, Grafikprogrammen, Workflow-Anwendungen usw. Es soll ermöglichen, dass alle Softwaresysteme mit der UML darstellbar werden.

„Die UML soll nach den Wünschen der Autoren eine Reihe von Aufgaben und Zielen verfolgen:

- Bereitstellung einer universellen Beschreibungssprache für alle Arten objektorientierter Softwaresysteme
- Vereinigung der bekanntesten Beschreibungsnotationen
- Setzen des Schwerpunkts auf die Produkte des Software-Engineerings und nicht auf den Prozeß.“ [UML_Dab00]

Sie ist ein Werkzeug zur Objektorientierten Analyse (OOA). In der OOA werden Daten und Funktionen als Einheit aufgefasst. Zentrale Begriffe von OOA sind Objekt und Klasse. Objekte und Klassen können wie nachfolgend dargestellt definiert werden (Zitate aus [SW01])

„Ein Objekt (object) ist ein problemrelevanter Gegenstand der Betrachtung, der über Eigenschaften (Attribute) verfügt und dessen Verhalten durch festgelegte Operationen beschrieben wird.

Beispiele für Objekte: Mitarbeiter, Auftrag, Polygon, Projekt.“

„Eine Klasse (class) definiert für eine Sammlung gleichartiger Objekte deren Struktur (Attribute), deren Verhalten (Operationen) und die Beziehungen zu anderen Klassen. Ferner besitzt eine Klasse einen Mechanismus, um neue Objekte zu erzeugen (object factory).

Beispiele für Klassen: Klasse der Mitarbeiter, Klasse der Polygone

Im Rahmen von OOA wird ein statisches Modell entwickelt, das die Klassen des Systems mit Attributen und Operationen, den Beziehungen zwischen den Klassen sowie den Vererbungsstrukturen beschreibt.“

Die UML Notation kann für alle in der objektorientierten Welt anzufertigenden Darstellungen verwendet werden. In der Regel werden Klassen als Rechtecke und Beziehungen

als Verbindungen dargestellt. Folgende Grafik verdeutlicht dies (vereinfacht nach [UML_QRC01]).

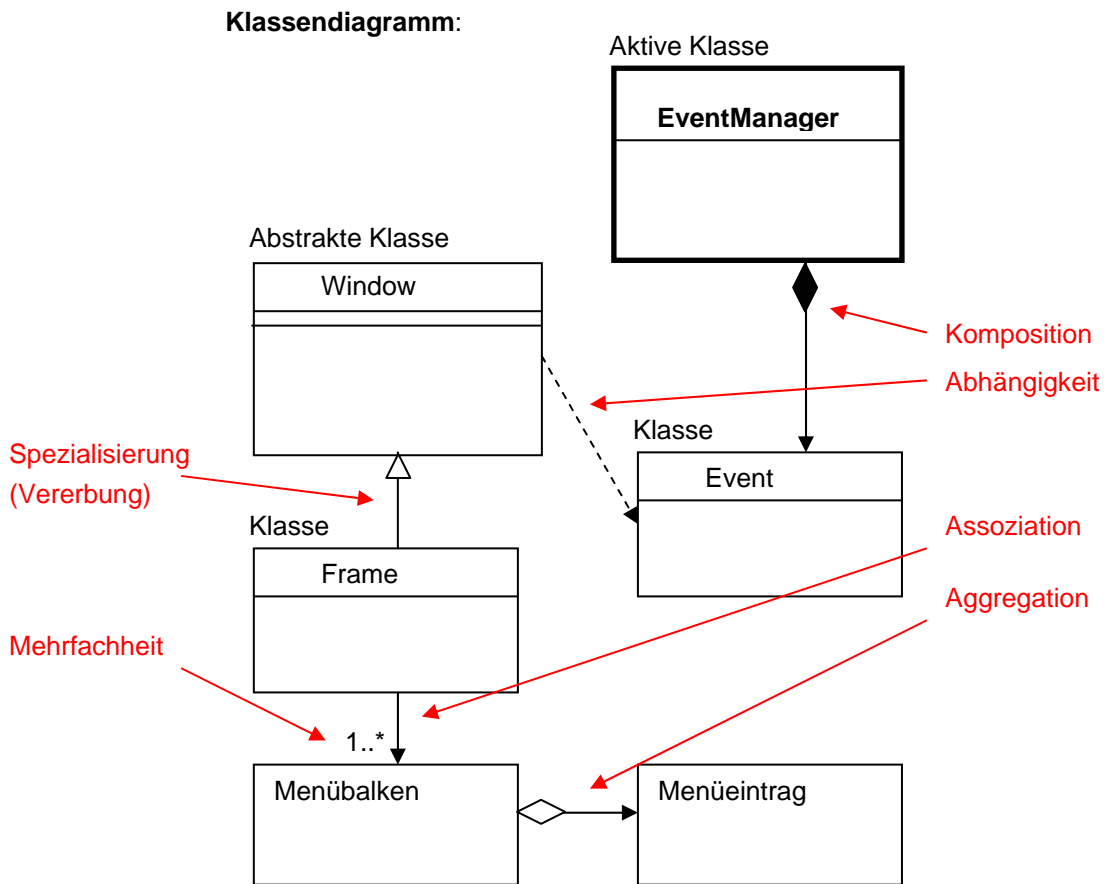


Abb.4.4: Schema der UML Notation

Mehrfachheit ist die Einschränkung, welche einem Objekt in der Beziehung zu einem anderen Objekt auferlegt wird. Die Darstellung kann folgende Werte annehmen.

1 ----- Ausschliesslich 1; dies ist der Standardwert

0..1 ----- 0 oder 1

M..N ----- Von M bis N (wobei M und N positive Ganzzahlwerte sind)

1..* ----- Von 1 bis zu einem beliebigen Ganzzahlwert.

Spezialisierung (Vererbung) beschreibt, dass Eigenschaften und Methoden von einem übergeordneten Objekt geerbt werden, bzw. dass eine Spezialisierung des Objektes erfolgt.

Aggregation stellt eine asymmetrische Beziehung dar, in welcher ein Objekt als Ganzheit verstanden wird und Objekte einer anderen Klasse als Teile. Die Teilobjekte können auch für sich allein bestehen.

Komposition ist eine stärkere Ausprägung der Aggregation. Die Teilobjekte können nicht für sich allein bestehen, sie sind vom zugeordneten Objekt abhängig.

Mit Hilfe dieser Werkzeuge kann zunächst ein logisches Datenbankmodell entwickelt werden, welches sich stark an die objektorientierte Struktur der Geodatabase anlehnt. Der Vorteil dieser Art der Modellierung ist, dass für die Datenbankmodellierung und für die Softwareentwicklung (gedacht wir hier in erster Linie an die Programmierung von Anwendungen zur Erweiterung von Funktionalitäten) dieselben Modellierungsmethoden verwendet werden. Wengleich dies mit großen Vorteilen bereits in der Modellierungsphase verbunden ist, soll auf den eigentlichen Gewinn dieser konzeptionellen Phase hingewiesen werden.

„Unabhängig von der benutzten Modellierungstechnik ist der entscheidende Vorteil, dass ein tieferes Verständnis für die Datenwelt eines Unternehmens entwickelt wird.“

[DFH03]

4.2.2 Vorgangsweise zur Erstellung logisch zweckmässiger Datenmodelle

An den Anfang dieser Ausführungen sei ein Zitat aus den Vorlesungsunterlagen gestellt, welches nicht oft genug wiederholt werden kann (und zumindest genauso oft seine Wirkung verfehlt):

„Man spricht zu wenig mit dem Auftraggeber, plant nur sehr oberflächlich und dokumentiert meist überhaupt nicht. Der Grund dafür liegt nur teilweise in mangelnder Fähigkeit, sich selbst zu organisieren.“ [UGIS02b]

Die Modellierung bildet, wie schon erwähnt, die Grundlage für den Aufbau eines GIS und bestimmt wesentlich dessen Leistungsfähigkeit. Wie die Daten genutzt werden können, wird in erster Linie durch das Modell festgelegt. So ist z.B. bereits im Rahmen der Modellbildung zu berücksichtigen, welche Informationen durch Abfragen und Analysen aus dem Datenbestand gewonnen werden können bzw. sollen.

Die Modellierung bezieht sich auf **Daten, Funktionen und Methoden**. Aufgabe der **Datenmodellierung** ist es, die Daten, die die Anwendung beschreiben, in einem geeigneten Modell darzustellen.

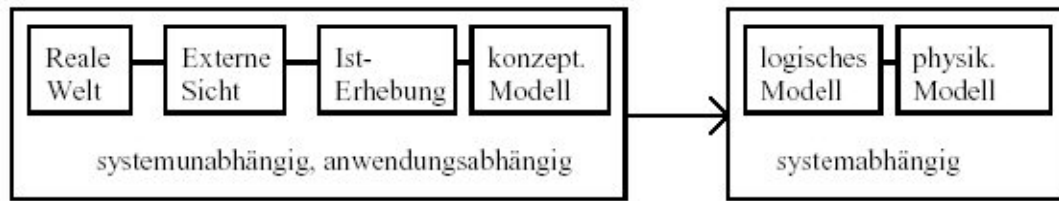


Abb. 4.5: Ablauf der Modellbildung

Im Rahmen einer Isterhebung bzw. Istanalyse werden zunächst Informationen über die Anwendung und die künftige GIS-Lösung erhoben. Es wird beispielsweise festgelegt, welche Daten erfaßt werden müssen, welche Datenquellen vorhanden sind und insbesondere welche Ziele mit einer GIS-Einführung verbunden sind. Dieser Schritt erfüllt somit die Funktion der Informationserhebung, die Strukturierung und Modellierung erfolgt im nächsten Schritt.

Im Rahmen der **konzeptionellen Modellbildung** erfolgt eine Strukturierung der Daten unter Einsatz bestimmter Modellierungs- und Beschreibungstechniken.

Der Übergang von der realen Welt zum Modell erfolgt durch einen, mit einem bestimmten Detaillierungsgrad durchzuführenden Abstraktionsprozeß.

In dieser Phase wird u.a. festgelegt, welche Objekte der realen Welt im einzelnen zu berücksichtigen sind, welche Eigenschaften der Objekte erfaßt werden müssen und welche Beziehungen zwischen den Objekten vorhanden sind. Weiterhin werden gleichartige Objekte in Klassen oder Gruppen zusammengefaßt, um eine systematische Strukturierung zu erhalten.

Die Topologie beschreibt die nichtmetrischen Beziehungen zwischen Objekten. Einige grundlegende topologische Beziehungen sind notwendig, um Geodaten (z.B. Flächen, Leitungsnetze) strukturiert zu speichern. Je umfassender die Topologie raumbezogener Daten erfaßt wird, desto größer sind die Möglichkeiten für Datenanalysen.

Im Rahmen der konzeptionellen Modellierung wird festgelegt, in welcher Form die dritte Dimension (Höhe) berücksichtigt wird.

Auch die Aktualisierung und Fortführbarkeit (Faktor Zeit) der künftigen GIS-Datenbestände sollte bereits in der Konzeptionsphase ein Kriterium sein. Hier wird weiterhin festgelegt, in welcher Form die Geometrie im GIS abgebildet und dargestellt wird: In Vektor- oder Rasterform oder kombiniert (hybrides GIS).

Im Rahmen des **objektorientierten Modellierens** wird die reale Welt als eine Vielfalt

miteinander in Beziehung stehender Objekte aufgefasst. Der Anwender findet die für ihn relevanten Objekte im GIS wieder (z.B. Haus, Flurstück, Leitung usw). Weitere Merkmale sind u. a.:

Klassenbildung: Jedes Objekt ist ein Exemplar einer Klasse. Der Klasse sind durch die Klassendefinition Eigenschaften und Methoden zugeordnet. Durch Bildung von Unter- und Oberklassen entsteht ein Objektklassenmodell.

Vererbung: Objekte einer Unterklasse erben Eigenschaften und Methoden der Oberklasse.

Kapselung: Ein Objekt ist eine abgeschlossene Einheit, der Eigenschaften und Methoden zugeordnet sind. Durch das Senden einer Botschaft wird ein Objekt veranlaßt, eine bestimmte Methode auszuführen.

Die Bildung **abstrakter, benutzer- bzw. anwendungsspezifischer Datentypen** wird unterstützt.

4.2.3 Vorschlag für ein Ablaufschema auf Basis der UML-Notation

Wie bereits erwähnt, wird die UML Notation einerseits verwendet, weil es sich bei der Geodatabase um ein objektorientiertes Modell handelt und daher eine Modellierungshilfe verwendet wird, welche ebenfalls auf der Basis der Objektorientierung beruht und andererseits für viele Modellierungsaufgaben verwendet werden kann, somit zu Recht als Modellierungsstandardsprache bezeichnet wird. Allerdings sind bestimmte Vorgaben einzuhalten, um UML richtig einzusetzen. Folgendes Beispiel baut auf einer ER Analyse auf, welche nach wie vor in der Relationalen Datenbankwelt als Standard für den Aufbau der Datenbanklogik verwendet wird. Anschließend wird das ER Modell (in UML Notation) in ein UML Modell umgewandelt. Dieses Modell kann dann über verschiedene Methoden in einem RDBMS implementiert werden. Die nachfolgend dargestellten Grafiken stammen aus [GOR03a].

Zunächst wird eine Entität (in diesem Beispiel die der Tabelle der Angestellten (employees) definiert, welche wie folgt dargestellt wird:

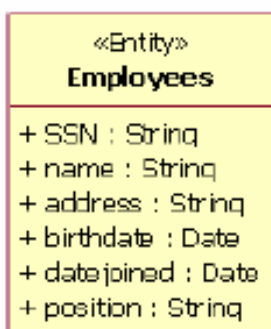


Abb 4.6.: Entität Angestellte in UML Notation

Dabei werden die das Objekt beschreibenden Inhalte in Abteile (sogenannten compartments) gegliedert; die erste Abteilung enthält zunächst die Klassifizierung des Typs als sogenannten „Stereotyp“ (in diesem Fall <<Entity>>). Ein Stereotyp wird zwischen doppelten Spitzklammern dargestellt und bezeichnet zusammengehörige Klassen. Die zweite Abteilung enthält die einzelnen Felder der Entität und getrennt durch einen Doppelpunkt den Datentyp. Das „+“ Zeichen bedeutet, dass die Elemente sichtbar („public“) sein sollen.

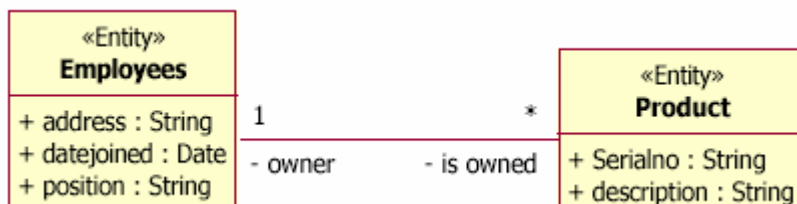


Abb. 4.7.: Darstellung der Beziehungen zwischen Entitäten; die Beziehung ist über die Angabe der Rollen näher bezeichnet (1 Angestellter ist Eigentümer von beliebig vielen Produkten)



Abb.4.8: Darstellung der Beziehungen zwischen Entitäten; die Beziehung ist über die Angabe des Namens näher bezeichnet.

Beziehungen zwischen den Entitäten werden als Linien dargestellt. Die an beiden Enden vorhandenen Zahlen oder Symbole bezeichnen die Kardinalität der Beziehung. Die Beziehungen werden entweder direkt benannt (d.h. die Verbindungslinie beschriftet) oder es werden die Rollen der Entitäten (im obigen Beispiel „owner“ (Eigentümer), „is owned“ (ist im Eigentum von)) angegeben.

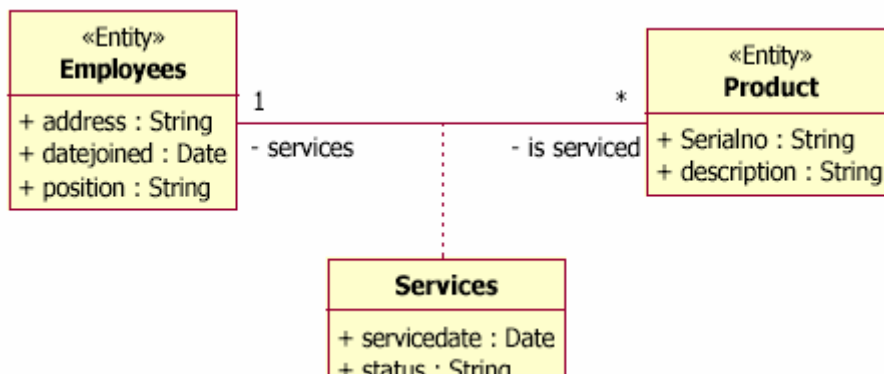


Abb.4.9: Eigenschaften der Beziehung „Services“ werden in einer assoziierten Klasse beschrieben.

Auch eine Beziehung kann Eigenschaften aufweisen, welche mit einer assoziierten Klasse beschrieben werden können. Die Darstellung erfolgt in einem eigenen Rahmen, welcher mit einer gepunkteten Linie mit der eigentlichen Beziehung verbunden ist. Es ist wichtig festzuhalten, dass Beziehung und Rahmen für die Beschreibung keine unabhängigen Elemente sind, sondern zusammengehören.

Abhängigkeiten können ebenfalls mit der Notation für Beziehungen abgebildet werden. Es werden in UML grundsätzlich zwei Formen unterschieden. Die **Aggregation** und die **Komposition**. Die folgende Grafik zeigt ein Beispiel für die erste Form.

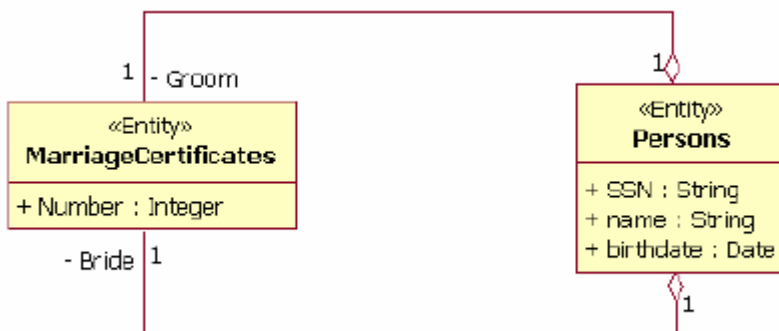


Abb.4.10: Jede Heiratsurkunde (Marriage Certificate) hängt von zwei Personen ab; der Braut und dem Bräutigam.

Eine **Aggregation** ist eine Abhängigkeit zwischen zwei Entitäten, welche für die Existenz der abhängigen Entität benötigt wird. Dargestellt wird die Aggregation mit einem hohlen Rautensymbol auf der Seite des Aggregates.



Abb.4.11: Die Entität OrderPositions ist durch die Entität Orders, wie durch die Komposition spezifiziert, vollständig definiert .

Eine **Komposition** bildet ebenso die Abhängigkeit zwischen zwei Entitäten ab, ist jedoch eine strengere Form davon. Sie wird dann verwendet, wenn die abhängige Entität nicht für sich alleine existieren kann. Somit bestimmt die die Komposition bestimmende Entität die Lebensdauer der abhängigen Entität. Sie wird durch ein gefülltes Rautensymbol auf der Seite der Seite der Aggregation dargestellt.

Ein weiteres wichtiges Merkmal der UML Notation in der Entity Relationship Analyse ist die **Spezialisierung** und **Generalisierung**.

Diese wird dann angewendet, wenn mehrere Entitäten eine gemeinsame Ausgangsentität aufweisen und sich voneinander in definierbaren Punkten unterscheiden.

Generalisierung wird der Schritt genannt, welcher von den Entitäten Gemeinsamkeiten extrahiert, **Spezialisierung** bezeichnet den Vorgang der Ausformulierung von Detailcharakteristika zur Bezeichnung von Entitäten.

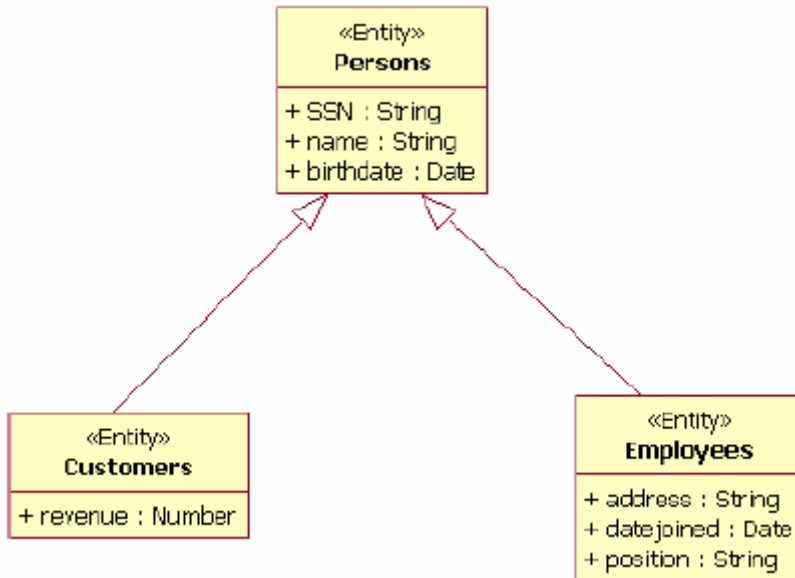


Abb.4.12: Generalisierung definiert die Entitäten Kunden (Customers) und Beschäftigte (Employees) als Typen der Entität Personen, welche ihre Eigenschaften erben.

Generalisierung (und in der Gegenrichtung die Spezialisierung) werden in UML durch einen hohlen Pfeil auf der „Elternseite“ der Beziehung dargestellt. Die Darstellung von Kardinalitäten ist bei der Generalisierung nicht zulässig.

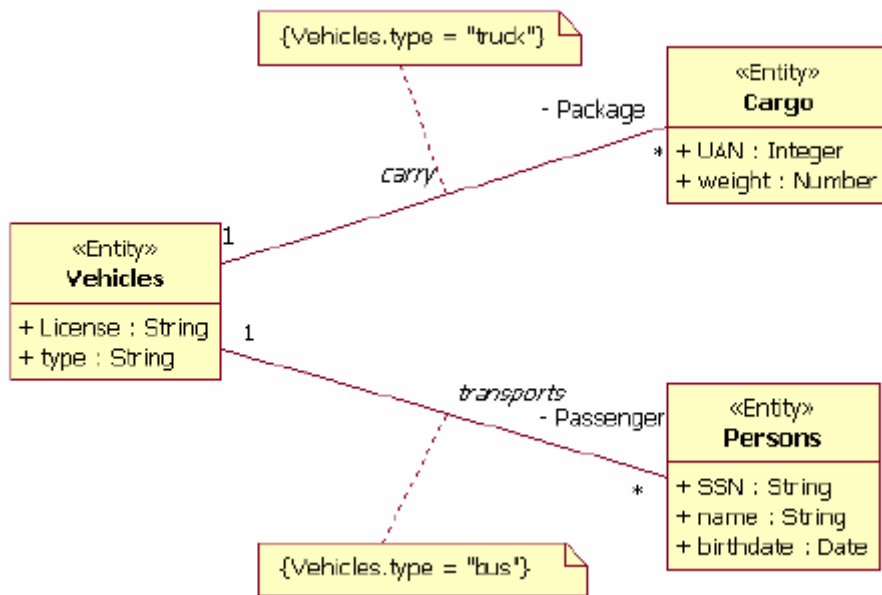


Abb.4.13: Kategorisierung von Entitäten definiert die Einschränkungen für bestimmte Beziehungstypen. Der Begriff Fracht (Cargo) ist wichtig für Fahrzeuge des Typs „LKW“, der Begriff Passagier (Passenger) ist von Bedeutung für Fahrzeuge des Typs „Bus“.

Beziehungen zwischen Entitäten können demselben Typ entsprechen, aber dennoch unterschiedliche Eigenschaften aufweisen. Diese Art von Beziehungen können als Kategorien zusammengefasst werden. Sie werden als Kommentar (in UML Notation eine strichlierte Linie und eine Box mit einer umgeknickten Ecke) dargestellt.

Bis zu hierher wurde die Entwicklung des logischen Modells (Entity-Relationship Analyse) mit der UML Notation ausgeführt. Nun muß das logische Modell in ein Relationales Modell übergeführt werden. Dies geschieht ebenfalls in der UML Notation.

Um den Unterschied noch einmal ausdrücklich herauszustreichen:

ER Modellierung dient dazu, das Problem oder die Aufgabenstellung zu analysieren, während die Relationale Modellierung auf die Implementierung abzielt.

Obwohl die Unterschiede zwischen diesen beiden Bearbeitungsschritten marginal erscheinen, bedeuten sie eine Konkretisierung und damit eine Annäherung an die Implementierung des Modelles in ein Datenbank-Artefakt. Dies soll anhand von einem einfachen Beispiel konkret dargestellt werden.

Die nachfolgend dargestellte Tabelle enthält mehrere Felder, welche Eigenschaften von Fahrzeugen enthalten. Diese Tabelle entspricht dem Endprodukt, welches in einer Relationalen Datenbank vorliegen soll.

LicenseNumber	Brand	Model	Year	Color
A17046	BMW	X5	2001	black
A13921	HONDA	ODYSSEY	2002	gold
A13502	VOLVO	S60	2002	silver
A18342	PEUGEOT	406	2000	red
A15022	INFINITY	Q45	2002	silver

Abb.4.14: Tabelle mit den Eigenschaften von Automobilen. Die gelb markierte Spalte mit der Zulassungsnummer enthält den Primärschlüssel.

Zunächst wird im Rahmen einer Entity-Relationship-Modellierung in UML Notation eine entsprechende Entität definiert. Diese enthält im Kopfbereich den Stereotyp <<Entity>> und darunter den Namen der Entität.

Im zweiten Kompartiment werden die Bezeichnungen der vorzusehenden Eigenschaften (d.h. die Feldnamen) mit dem vorzusehenden universellen Datentyp. Hier wird allerdings noch nicht definiert, welche Ausprägung dieser Datentyp in der Implementierung erhalten wird (z.B. Char(20) für einen String, welcher vom Typ Charakter ist und 20 Zeichen lang ist).

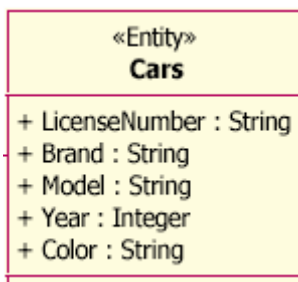


Abb.4.15: Darstellung einer Entität mit zugehörigen Eigenschaften.

Aus dem ER-Diagramm wird anschließend das Relationale Diagramm entwickelt. Es enthält weitere Informationen darüber, wie die Tabelle implementiert werden soll. Im Unterschied zum ER-Diagramm erhält das Objekt eine Stereotypbezeichnung, welche dem Inhalt entspricht. Es kann neben der Tabelle (wie in diesem Fall) auch eine virtuelle Tabelle (view) sein. Neben dem Namen wird weiters festgelegt, welches Feld den Primärschlüssel enthält; dieser wird mit dem Bezeichner „PK“ für PrimaryKey versehen. Der in der Darstellung der Entität angegebene Datentyp wird hier konkretisiert und seine Ausprägung definiert (CHAR(8)).

Der Bezeichner „NN“ vor dem Feldnamen bedeutet NOT NULL, es muß also ein Wert eingetragen werden; „N“ (NULL) bedeutet, dass das Feld auch leer bleiben kann. Fremdschlüssel werden mit dem Buchstaben „F“ für ForeignKey gekennzeichnet.

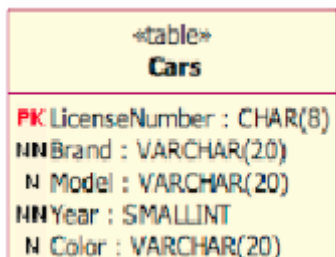


Abb.4.16: Relationales Modell, in welchem die Tabellenstruktur in UML Notation angegeben wird.

Neben den Entitäten müssen auch Beziehungen aus der ER-Modellierung im Relationalen Modell abgebildet werden. Als Beispiel dafür dient die Erweiterung des oben angeführten Beispiels um zwei weitere Tabellen, nämlich Einwohner als Eigentümer der Fahrzeuge (Residents) und genehmigte Änderungen an den Fahrzeugen (Approved Changes). Die Vorgangsweise zur Abbildung der Beziehungen zwischen diesen Tabellen zeigt analog zu oben folgendes Schema.

SocialSecurityNumber	FirstName	MiddleName	LastName	DateOfBirth	Sex
232483654	Jim	Evans	Bellevue	7/13/1971	M
254865482	Eva	Weber	Adams	8/16/1974	F
865947523	Ken		Dietrich	2/3/1952	M
859647521	Barbara		Mc Kenna	11/2/1951	F

Abb.4.17: Tabelle Einwohner (Residents).

LicenseNumber	ChangeName	ChangeDescription	ApprovalDate
A17046	Sports Exhaust	Brand Loud Devil, Model Polished Tiger	3/19/2002
A13921	New Engine	New Serial No: 458D093847233-0231	8/11/2002
A13502	New Engine	New Serial No: 93820XJ03988742BDD	2/1/2003

Abb.4.18: Tabelle genehmigte Änderungen (Approved Changes).

Die Beziehungen zwischen den einzelnen Entitäten in der ER-Modellierung werden durch Verbindungslinien repräsentiert. Die Kardinalität wird durch Symbole auf beiden Seiten der Entität dargestellt. Zwischen den Einwohnern (Residents) und den Fahrzeugen (Cars) besteht folgende Beziehung:

Ein Einwohner kann 0 oder beliebig viele Fahrzeuge besitzen (gekennzeichnet durch das *-Symbol auf der Seite der „Cars“-Entität). Ein Fahrzeug kann 0 oder einem Einwohner zugeordnet sein (0..1).

Zwischen den Fahrzeugen und den genehmigten Veränderungen bestehen folgende Beziehungen:

An einem Fahrzeug können beliebig viele genehmigte Änderungen vorgenommen werden. Allerdings können die Änderungen an einem Fahrzeug nicht für sich allein existieren. Deshalb ist die Beziehung als Pfeilsymbol ausgeführt (siehe oben).

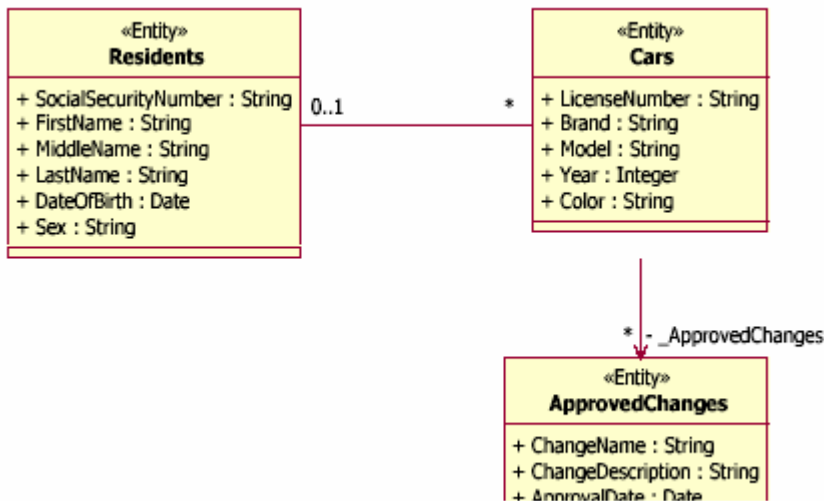


Abb.4.19: Eigenschaften der Beziehung „Services“ werden in einer assoziierten Klasse beschrieben.

Die Überführung diese ER Schemas in ein Relationales Modell wird wie folgt durchgeführt (beschrieben werden vor allem die Elemente, welche noch nicht erläutert wurden).

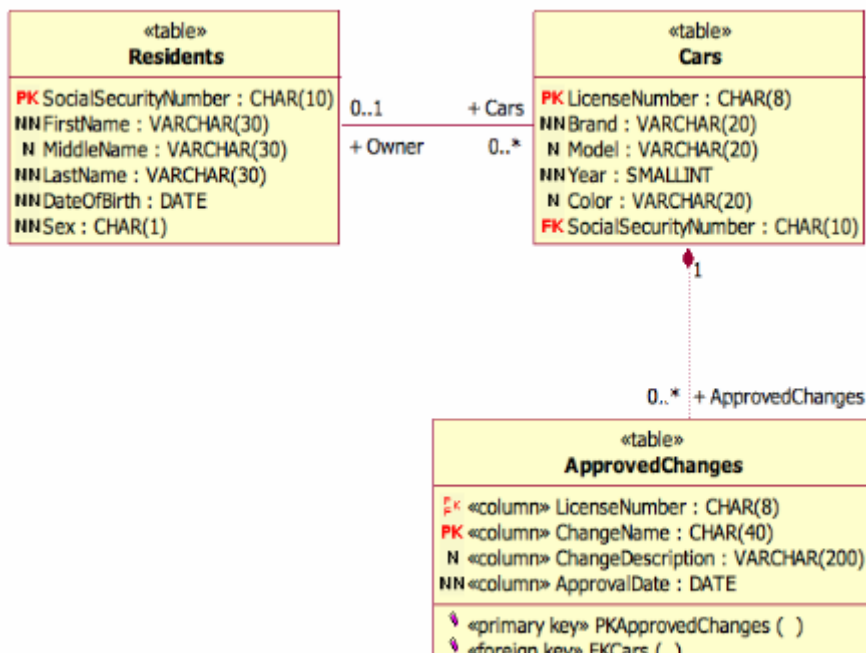


Abb.4.20: Darstellung der Beziehungen zwischen den einzelnen Relationen in der UML Notation.

Die umgeformte relationale Entsprechung des ER-Diagramms verwendet eine sehr ähnliche Syntax. Allerdings enthält das relationale Diagramm zusätzlichen Informationen, welche die Umsetzung in eine Datenbank wesentlich erleichtert.

Die Beziehungen zwischen den Tabellen erhalten zunächst die für die UML Notation vorgesehene Symbolik. Die Angabe der Rolle (Eigentümer bzw. owner) unter der Bezeichnung für die Kardinalität legt die Art der Beziehung genauer fest. Die Symbolik für die Beziehung zwischen den Fahrzeugen und den genehmigten Veränderungen unterscheidet sich zum ER-Diagramm insofern, als über ein eigenes Symbol (die gefüllte Raute) eine Komposition dargestellt wird (die über eine Komposition verbundene Relation kann nicht ohne die „Eltern“-Relation existieren).

Anzumerken ist noch die Ergänzung des Tabellen-Objektes durch einen dritten Abschnitt unter der Definition der Felder: In diesem Abschnitt wird dokumentiert, zu welchem Primärschlüssel in der über eine Beziehung verbundenen Tabelle ein Fremdschlüssel existiert. Der Fremdschlüssel wird in der Tabelle mit P_{FK} (Primary-Foreign-Key) gekennzeichnet.

4.2.4 Ablaufschema:

In den vorangegangenen Ausführungen haben wir uns mit den Grundlagen und der Technik der Modellierung befasst. Jetzt stehen uns alle Werkzeuge zur Verfügung, um ein Ablaufschema aufzustellen.

Dieses Ablaufschema gliedert sich in drei Gruppen. Die organisatorischen Aspekte, das konzeptuelle Schema, das logische Modell. Jede einzelne Maßnahme kann einer dieser drei Gruppen zugeordnet werden.

I.) Allgemeine Datenbankverwaltung:

- 1) Vollständige Sicherung des letztgültigen Datenbestandes auf ein aktuelles Speichermedium (DVD)
- 2) Erstellen einer Liste, welche Daten (geordnet nach Themenebenen) in die GeoDB übernommen werden sollen
- 3) Abklärung und Festlegung der benötigten Datenbankstrukturen (abhängig vom Nutzerkreis und den notwendigen Zugriffsberechtigungen); das sind GeoDatenbanken, feature datasets und feature classes.
- 4) Festlegung und Dokumentation der Benutzergruppen und -rechte

II.) Konzeptionelles Schema:

- 5) Erstellen der Datenmodelle zu jedem Datensatz (bzw. zu jeder Themenebene). Dazu zählt insbesondere die Definition der Objekte und deren Beziehungen untereinander.
- 6) Auswahl der tatsächlich zu übernehmenden Daten (z.B. Tics, arcs bei Routen etc.), externe Attributtabellen
- 7) Definition und Dokumentation der für die einzelnen Objekte benötigten Datentypen (insbesondere die Verwendung von einfachen Elementen (simple features) versus komplexen Elementen (complex features))
- 8) Definition und Dokumentation der zu verwendenden Projektionen, der Genauigkeit und der abzubildenden Ausdehnung (Extent)

III.) Logisches Modell:

- 9) Anlegen der zuvor festgelegten Datenbanken und Datenhüllen und Eingabe der vordefinierten Eigenschaften (Genauigkeit, Extent u.ä.)
- 10) Befüllen der Datenbank mit Hilfe von ArcCatalog oder batchgesteuerten ArcSDE Kommandos.
- 11) Definition und Eingabe der Regeln und der Attributdomänen (Beispiel DKM - Regeln: kein Grundstück ohne Nummer, keine Überlappungen; Attributdomänen)

4.3 Planung: organisatorische Aspekte des Umstiegs

In diesem Kapitel kommt jener Bereich zur Sprache, welcher für die Umsetzung unseres Vorhabens von entscheidender Bedeutung ist. Die Organisation ist von grösster Bedeutung, weil hier in sehr starkem Maße die menschliche Komponente zum Tragen kommt. Vor allem in der Betriebswirtschaftslehre hat sich diese Erkenntnis durchgesetzt – eine Entwicklung, vor der sich auch die technischen Disziplinen nicht verschliessen dürfen.

4.3.1 Exkurs – der Wechsel als Prinzip (Change Management)

Änderungen altbekannter Vorgangsweisen und Techniken fallen schwer. Insbesondere dann, wenn der Vorteil der Neuerungen nicht gleich offensichtlich zu Tage tritt in Form von Kostensenkung oder Zeitersparnis. Zudem kennt man sich im bewährten Schema aus, die Einführung neuer Technologien verursacht Aufwand in vielerlei Hinsicht, der Erfolg ist nicht garantiert.

Solche oder ähnliche Probleme haben zur Entstehung einer eigenen Wissenschaft geführt, welche unter dem Titel „Change Management“ in den Wirtschaftswissenschaften Einzug gehalten haben. Entstanden ist dieser Forschungszweig aus einer Vielzahl von Fehlschlägen in Projekten, welche zu einer wissenschaftlichen Aufarbeitung der Gründe geführt haben (Literaturhinweise: [DÖ01], [SVDW93]).

Die nun folgenden Vorschläge und Vorgangsweisen mögen trivial und offensichtlich erscheinen, dennoch bildet deren Vernachlässigung häufig den Grund des Scheiterns.

- Grundsätzlich kann man davon ausgehen, dass das Wohl und Wehe eines Veränderungsprojektes von den Betroffenen, d.h. allen handelnden Personen abhängt. Technische oder finanzielle Probleme sind in der Regel nur Vorwände.
- Die Betroffenen müssen noch vor der Durchführung des Technologiewechsels einbezogen und beteiligt werden.

Beispielhaft und zusammenfassend sollen die wesentlichen Punkte angeführt werden, welche dazu beitragen können, den Fehlschlag eines „Veränderungsprojektes“ zu vermeiden [DL00]:

- Klarheit der Ziele: Wie klar ist den Betroffenen, was mit dieser Veränderung konkret bezweckt wird?
- Informationsstand: Von welchem Wissensstand über das anstehende Thema kann man bei den Betroffenen ausgehen?
- Problembewusstsein: Empfinden die Betroffenen die Situation, um die es geht, überhaupt als Problem? Gibt es so etwas wie „Leidensdruck“?
- Glaubwürdigkeit des Vorhabens und der Initianten: Wie sehr nimmt man den Initiatoren ab, dass es ihnen tatsächlich um die Sache geht, die sie vorbringen?

- Energie und Engagement: Aus alledem ergibt sich das Ausmaß an Energie, mit dem sich die Beteiligten für die Problemanalyse und –lösung engagieren oder sich gegen sie sperren werden.

Erst wenn man spürt, dass die Betroffenen die Probleme erkennen und dass der Impuls zum kreativen Mitmachen vorhanden ist, macht es sinn, den nächsten Schritt zu tun, d.h. in die Phase der konkreten Problembearbeitung überzuleiten.“...

Aus diesen Ausführungen ist abzuleiten, dass die Betroffenen (das sind sowohl Systemverantwortliche, Datenlieferanten, als auch Nutzer des Informationssystems, kurz die „stakeholder“ (im anglo-amerikanischen Sprachraum werden so alle Interessenten an einer Sache bezeichnet) frühzeitig, das heißt bereits in die Konzeptionsphase, einzubeziehen sind. Ansonsten ist mit Widerständen, im schlimmsten Fall mit dem Scheitern des Projektes zu rechnen.

4.3.2 Organisation und Schulung

In diesem Zusammenhang sind auch die Punkte Organisation und Schulung zu sehen. Kann ein klares Konzept zur Umsetzung vorgestellt werden, ist die Akzeptanz der Betroffenen leichter zu gewinnen.

Organisation:

Dieser Aspekt scheint mir in dieser Beziehung deshalb von besonderer Bedeutung zu sein, weil die Organisationsstrukturen in vielen Unternehmen (insbesondere in der Öffentlichen Verwaltung) allen organisationstheoretischen Modeerscheinungen zum Trotz dem Liniensystem zuzurechnen sind. Charakteristika dieses Systems sind die hierarchische Ordnung (Chef – Abteilungsleiter – Sachbearbeiter) und das Vorhandensein eines „Dienstweges“ [WÖHE90].

Damit können Vorhaben von der Qualität des gerade behandelten nicht oder nur mit erheblichem Aufwand durchgeführt werden, weil Entscheidungen oft ad hoc und unter Mitarbeit aller Beteiligten zu treffen sind. Besser geeignet sind die Formen Team-Organisation bzw. die Projekt-Organisation, weil lange Entscheidungswege wegfallen, die Entscheidungskompetenz in der Gruppe vorhanden ist und die Fachexpertise durch Arbeitsgruppenmitglieder in der Regel abgedeckt ist.

Team-Organisation: „Unter einem Team im organisatorischen Sinne versteht man eine Stelle, deren Aufgabenbereich von einer Gruppe von Personen gemeinsam und weitgehend autonom bearbeitet wird.“ [THOM96]

Projekt-Organisation: „Projektaufgaben zeichnen sich durch die drei Merkmale Komplexität, Singularität und originäres Zielsystem aus.“ [THOM96]

Für die Erledigung der Projektaufgabe wird für eine begrenzte Zeit eine Projekt-Organisation zusammengestellt.

Der Unterschied zwischen den beiden Organisationsstrukturen: die Team-Organisation ist in der Regel dauerhaft, die Projekt-Organisation vorübergehend.

Schulung:

Der Themenbereich Schulung bezieht sich beim Umstieg von filebasierter zu datenbankbasierter Datenhaltung auf eine sehr weite zeitliche Spanne. Zwischen der Ausbildung des Endanwenders (so er überhaupt einen Unterschied zwischen vorher und nachher bemerkt) und dem Ersteller der Datenmodelle können Wochen oder Monate liegen. Das bedeutet, dass die Schulung auf neue Systeme mit einer zeitlichen Staffe- lung erfolgen kann und zweckmässigerweise auch soll. Auch in diesem Fall erscheint es angebracht, eine Dokumentation anzufertigen, welche sich mit den Schulungsmaßnah- men der verschiedenen Systemnutzer beschäftigt. Dabei sollen die Maßnahmen so ge- setzt werden, dass die Kenntnisse unmittelbar umgesetzt werden können.

Als bedeutend erkannt und daher eigens erwähnt wird die Tatsache, dass das Verständ- nis der geänderten Technologie und damit der Anwendung wesentlich wichtiger ist, als das rein technische Können. Daher ist das Augenmerk verstärkt auf die Vermittlung der Hintergründe und Prinzipien denn auf die Bedienung der einzelnen Befehlsschaltflächen zu legen.

Diese zwei Punkte, Organisation und Schulung, sollen bereits in der Vorbereitungspha- se energisch in Angriff genommen werden. Die personellen Strukturen müssen festge- legt sein, wenn das System seinen Betrieb aufnimmt. Zweckmässig ist es auch, die Um- stellung anhand eines Pilotprojektes zu testen; dieses wird von einer Kerngruppe durchgeführt und aufbauend auf den Erkenntnissen die Planung konkretisiert.

Die Schulung stellt einen Schlüsselaspekt in der Umsetzung dar. Die Mitarbeiter müs- sen in der Anwendung der neuen Technologien geschult werden und insbesondere die Unterschiede zwischen vorher und nachher kennen lernen. Datenbankwissen, insbeson- dere solches von DB Administratoren ist unabdingbar. Auch hier wird wieder die Trag- weite der Umstellung offensichtlich – die Standard-IT Technologien sind aus der GIS- Welt nicht mehr wegzudenken.

4.3.3 Ausstattung

Die Ausstattung ist stark von den Ansprüchen an das System hinsichtlich Geschwindig- keit, Daten- und Betriebssicherheit abhängig. An dieser Stelle soll lediglich darauf hin- gewiesen werden, dass der Einsatz relationaler Datenbanken, welche mit Massendaten

betrieben werden hohe Anforderungen an die Hardware stellt. Das geplante System ist insbesondere auf seine Leistungsfähigkeit zu untersuchen. Dazu gibt die einschlägige Literatur der Hersteller tiefere Einblicke und geeignete Anleitungen.

4.3.4 Zeitplan

Die Frage nach der terminlichen Planung umfangreicher Vorhaben (im engeren Sinne von Projekten) füllt Bände.

Dennoch sollen hier (sozusagen als Zusammenfassung) die wichtigsten Schlagworte im Zusammenhang mit der Zeitplanung angeführt werden: tlw. verändert nach [EUR03]

Projekttablauf:

- Strukturiert das Projekt in verschiedene Phasen
- Jede Projektphase beginnt und endet mit einem wesentlichen Projektereignis, einem sogenannten Meilenstein.

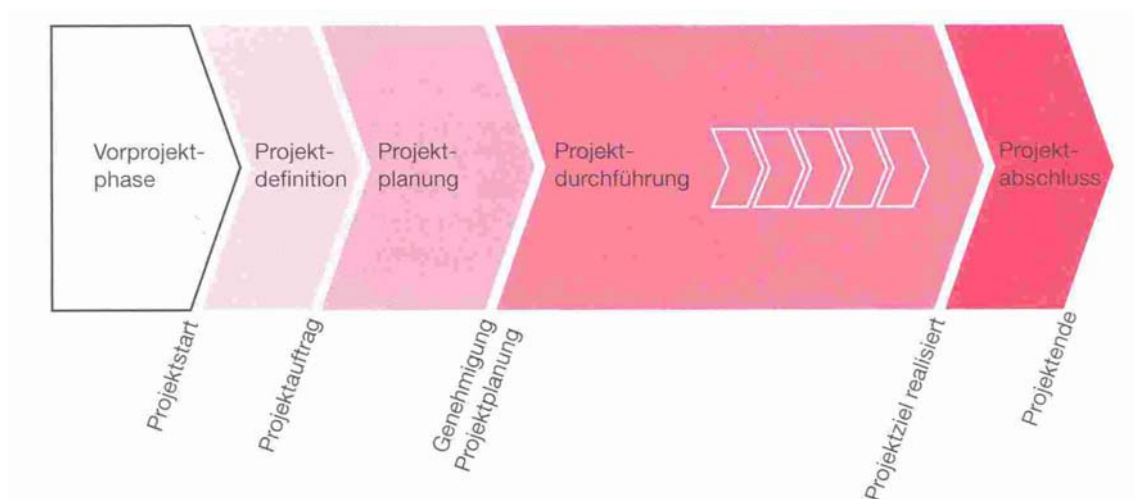


Abb.4.21: Schema eines idealtypischen Projekttablaufes

Meilensteinplanung:

- Um Projekte überschaubar und beherrschbar zu machen, werden sie mittels Meilensteinen in verschiedene Phasen unterteilt.
- Meilensteine sind wesentliche Projektereignisse
- Je Projekt sollen zwischen 5 und 7 Meilensteine definiert werden.

Terminplanung:

- Die Berechnung der Termine erfolgt auf der Basis des Projektstrukturplanes*
- Die logischen Verknüpfungen zwischen den Arbeitspaketen beeinflussen erheblich die Planung der Termine; Was muss fertig sein, damit die Aufgabe XY begonnen werden kann?
- Die Terminplanung kann durch die Festsetzung von Meilensteinen erleichtert werden; das sind zeitlich terminierte Zwischenziele, auf die alle Projektmitglieder hinarbeiten
- Effizientes und gut kommunizierbares Instrument der Terminplanung: Balkenplan

Der Zeitplan wird auf Basis der Ergebnisse des Pilotprojektes erstellt. Insbesondere die Vorgangsweise bei der Einrichtung der Hard- und Software und die Umstellung der einzelnen Datenbestände sind von großer Bedeutung.

Es ist zweckmässig, einen Zeitplan nach der oben skizzierten Vorgangsweise aufzustellen. Er ist eine wichtige Orientierungshilfe bei der erfolgreichen Umsetzung einer komplexen Aufgabe.

4.4 Die Umsetzung

Die Umsetzung des Vorhabens, wie es bis zum jetzigen Punkt angedacht wurde, dürfte keine grundsätzlichen Probleme mehr bereiten. Die Beteiligten wurden informiert, wenn nötig in ein Team einbezogen, die Ansprüche an das Informationssystem wurden formuliert, die Datenmodelle erstellt, ein Ablaufschema erstellt und last but not least ein Projekt definiert.

Es sind dennoch einige Fragen zu erörtern, die im Zuge der Umsetzung evident werden könnten. Zunächst sei noch ein Hinweis auf eine Sammlung ergänzender Werkzeuge erlaubt, welche verhindern sollen in der Euphorie des „endlich-machen-könnens“ das sprichwörtliche Kind mit dem Bade auszuschütten.

4.4.1 Exkurs – die Durchführung komplexer Vorhaben als Aufgabe (Systems Engineering)

Die Durchführung des Umstieges von filebasierter Datenhaltung zu datenbankbasierter Datenhaltung ist ein Schritt, welcher neben der detaillierten Ermittlung der notwendigen

* Der Projektstrukturplan (PSP) ist ein zentrales Planungsinstrument im Projektmanagement. Er strukturiert das Projekt in Arbeitspakete. Jedes dieser Arbeitspakete ist eine Teilarbeit und gibt darüber Aufschluss, was getan werden muss, um ein Projekt zu erfüllen.

Aufgaben auch die Konzeption des Systems im Blickfeld behalten muß. Dazu gibt es eine Vorgangsweise, welche sich als „Systems Engineering“ etabliert hat und sich mit Planungsmaßnahmen im generellen Sinne beschäftigen.

Der Hinweis auf die Vorgangsweisen des Systems Engineering soll nicht zur weiteren Komplexität des Vorhabens beitragen, sondern eine Hilfe für die Planung und Umsetzung. Jedenfalls trägt es dazu bei, möglicherweise wichtige Elemente nicht zu übersehen. Das im vorigen Kapitel kurz behandelte Planen in Projekten greift dabei zu kurz, weil es sich auf die eigentliche Fragestellung beschränkt. Im Systems Engineering (kurz SE) wird auch das Umfeld beleuchtet. Es wird deshalb an dieser Stelle behandelt, weil insbesondere der zweite Teil, das SE-Vorgehensmodell, Handlungsanleitungen für die Durchführungsphase von Vorhaben bereitstellt.

„Auf eine knappe Formel gebracht, soll Systems Engineering (SE) als eine auf bestimmten Denkmodellen und Grundprinzipien beruhende Wegleitung zur zweckmäßigen und zielgerichteten Gestaltung komplexer Systeme betrachtet werden.“ [HNBBM94]

...

„SE befasst sich mit der Aufgliederung komplexer Probleme in überblickbare und damit einzeln lösbare Teilkomplexe, mit der zweckmäßigen Abfolge von Tätigkeiten und mit der Koordination verschiedener beteiligter Personen.“ [HNBBM94]

Die SE – Philosophie lässt sich dabei in zwei Grundkomponenten zerlegen:

1. Systemdenken (d.h. das Denken in Systemen)
2. Das SE-Vorgehensmodell (mit den Teilbereichen: vom Groben zum Detail, Prinzip der Variantenbildung, Phasengliederung von Projekten, Problemlösungszyklus)

Zu 1: Ein System kann auf vielfältige Weise analysiert werden, um seine Wirkprinzipien zu erfassen und keine bedeutenden Teilbereiche zu vergessen.

Auch beim Aufbau von Systemen sind die Analyseergebnisse wichtig. Als Faustregeln für den Aufbau von Systemen sollen jedenfalls folgende Aspekte berücksichtigt werden:

- Prinzip der Minimierung von Schnittstellen (möglichst wenige und einfache Beziehungen sollen nach außen bestehen)
- Prinzip des modularen Aufbaues (Systembausteine sollen so gebildet werden, dass sie mehrfach zu verwenden sind)

- Prinzip des „Piecemeal Engineerings“ (in großen und komplexen Systemen soll man sich davor hüten, Veränderungen, deren Auswirkungen nicht durchschaut werden können, in großen Schritten zu vollziehen).
- Prinzip der „minimalen Präjudizierung“ (im Zweifelsfall ist jener Lösung der Vorzug zu geben, welche die meisten Freiräume für die weitere Entwicklung offenhält).

Zu 2: Dem Vorgehensmodell liegen 4 Grundgedanken zu Grunde:

- Vom Groben zum Detail -
- Denken in Varianten
- Prozeß der Systementwicklung und –realisierung nach zeitlichen Gesichtspunkten gliedern (Phasengliederung)
- Anwenden eines Problemlösungszyklus

Damit soll das an dieser Stelle recht kurz angeschnittene Thema des SE seinen Abschluß finden. Eine ausführliche Darstellung findet sich in [HNBBM94].

Einige Grundgedanken des SE sollen in der nachfolgenden Anwendung des Ablaufschemas berücksichtigt werden.

4.4.2 Anwendung des Ablaufschemas

Das unter Punkt 4.2.4 skizzierte Ablaufschema dokumentiert nur eine von vielen Möglichkeiten, den Umstieg in der Technologie der Datenhaltung zu realisieren.

Anhand von Beispieldatensätzen (konkret waren dies die DKM und das Fliessgewässernetz von Vorarlberg) wurde die Relevanz der Vorgabe getestet. Detaillierte Angaben dazu können weiter unten nachgelesen werden. Zur Anwendung des Ablaufschemas sollen an dieser Stelle, sozusagen vorab, folgende Bemerkungen angebracht werden:

- Das Ablaufschema kann nicht ohne Grundlagenwissen angewendet werden. Kenntnisse über die Modellierung von Daten, die einzelnen Fachthemen und die entsprechenden Anwendungen sind notwendig.
- Der Ablaufplan kann uns muss an die entsprechende Aufgabe angepasst werden. Dies kann Verringerung aber auch Ausbau der einzelnen Schritte bedeuten.
- Es hat sich als zweckmässig herausgestellt, die Vorgangsweise auf breiter Basis zu diskutieren (siehe dazu oben: Teambildung)
- Die Anwendung des Systems Engineering ist ein wesentlicher Aspekt hinsichtlich der Plastizität des Umwandlungsprozesses. Änderungen sind trotz bester Planung die Regel, nicht die Ausnahme.

4.4.3 Dauer der Umsetzung

Während die Vorbereitung aufgrund vieler zu berücksichtigender Aspekte (Modellierung der Daten, Rangreihung, Planung des Betriebes, ...) längere Zeit in Anspruch nehmen kann, liegt es im Interesse des Betreibers des Informationssystems, die Umsetzung der Planung in möglichst kurzer Zeit durchzuführen. Die parallele Führung von alten und neuen Beständen ist aufwändig und somit teuer. Sie ist außerdem fehleranfällig und damit gefährlich für Konsistenz der Daten.

Die Angabe einer konkreten Dauer ist an dieser Stelle nicht möglich, weil diese wiederum von vielen Faktoren (siehe oben) abhängig und daher nicht pauschal angegeben werden kann.

Aufgrund der Erfahrungen aus den Testläufen ist jedoch mit einer Dauer in der Größenordnung von Wochen denn von Tagen auszugehen.

4.4.4 Betrieb während der Umsetzung

Eine zumindest vorübergehende parallele Führung der Systeme lässt sich trotz bester Vorbereitung nicht vermeiden, weil einerseits die Anzahl der verschiedenen Themen groß ist, die Einspielung, Kontrolle, Test der Daten zu viel Zeit in Anspruch nimmt. Die Systeme sollen daher unabhängig voneinander sein (d.h. Systeme nicht auf derselben Hardware), um auch fatale Fehler nicht in einen GAU ausarten zu lassen. Es ist jedenfalls mit Problemen zu rechnen.

Zudem sollen hier zwei wichtige Faktoren des Systems Engineering rekapituliert werden, deren Berücksichtigung zwar Zeit kostet, die Verlässlichkeit eines Systems aber wesentlich erhöht. Das Prinzip des „Piecemeal Engineerings“ und das Prinzip der „minimalen Präjudizierung“ (siehe oben).

4.5 Die Nachbereitung

Die Nachbereitung des gegenständlichen Vorhabens erfordert ebensoviel Aufmerksamkeit wie die Vorbereitung. Gleichwohl nicht alle zweckmässigen Arbeiten in der wünschenswerten Tiefenschärfe durchgeführt werden können (zum Beispiel eine lückenlose Dokumentation der gesetzten Schritte), sollen die wichtigsten Maßnahmen kurz skizziert werden. Das sind die abschliessenden Arbeiten, die Dokumentation und die Kritik der Anwendung des Ablaufschemas.

4.5.1 Abschliessende Arbeiten (final Clean Up)

Unter dem Begriff „abschliessende Arbeiten“ sollen jene Vorgänge subsumiert werden, welche in diesem Kapitel nicht mit einem eigenen Punkt abgehandelt werden. Dabei

handelt es sich hauptsächlich um Fragen, welche während oder nach dem Umstieg aufgetaucht sind und in den bisherigen Erläuterungen keine Erwähnung gefunden haben. Zu diesen Punkten gehören:

- Die Dokumentation von Fragen und Problemen: Dabei hat sich bewährt, die einzelnen Fragen in chronologischer Reihenfolge (nach Datum) zu erfassen und die Lösung im selben Journal zu dokumentieren. Eine solche Liste kann bei nachfolgenden Tätigkeiten ähnlicher Art von großem Nutzen sein.
- Die Funktionskontrolle des neuen Systems – Programmfunktionalität, Sicherungsfunktionalität etc.: Dieser Punkt ist nur insoweit Gegenstand dieser Arbeit, als er die Datenhaltung selbst betrifft. Dennoch soll der Vollständigkeit halber darauf hingewiesen werden.
- Die offizielle Übernahme und die Bekanntmachung des neuen Systems im Benutzerkreis: Diese Maßnahme ist zumindest in zweierlei Hinsicht wichtig. Zum einen soll der Startschuß für die Nutzung der Datenbestände gegeben werden, zum anderen findet damit eine förmliche Übergabe statt und damit ein Zeichen, dass der Umstieg abgeschlossen wurde.

4.5.2 Kritik der Anwendung des Ablaufschemas

Fehler, die begangen werden, Probleme, die auftreten sollen nicht unter den Tisch gekehrt, sondern in geeigneter Form behandelt, diskutiert und dokumentiert werden. Eine selbstkritische Reflexion des Handelns führt u.a. zu verbesserten Prozeduren.

Das Ablaufschema ist zugegeben grob gestrikt (und daher ohne Vorkenntnisse nicht anzuwenden). Allerdings würde einerseits der Rahmen eines Schemas bei einer detaillierten Behandlung gesprengt, andererseits können bei der Vielzahl an verschiedenen Konfigurationen unmöglich alle Varianten durchgespielt werden. Der Gewinn an Vollständigkeit würde auf Kosten der Klarheit erkaufte.

4.5.3 Dokumentation des Umstieges

Die Dokumentation wird hier an den Schluß der Ausführungen gestellt. Nicht, weil alle anderen Inhalte wichtiger wären, sondern ganz im Gegenteil. Die Dokumentation ist neben dem Nachweis der durchgeführten Arbeiten (und der dafür verwendeten Zeit) ein wichtiges Instrument für die Planung nachfolgender Projekte und eine wichtige Informationsquelle für später in Erscheinung tretende Fragen oder Probleme.

Ein weiterer nicht zu vernachlässigender Aspekt der Dokumentation ist ihre Funktion beim personellen Wechsel im Bearbeiter-Team. Damit können unangenehme Wissenslücken von vorne herein vermieden werden.

Sie dient ausserdem dazu, Auftraggebern und anderen Betroffenen den Nachvollzug der Erarbeiteten Lösungen zu ermöglichen.

Als Dokumentation eignet sich am besten eine nach Datum geordnete Tabelle mit den chronologischen Einträgen der durchgeführten Tätigkeiten und dem Namen der durchführenden Personen.

4.6 Erfahrungsberichte

An dieser Stelle hätten Erfahrungsberichte anderer Anwender zum Umstieg vom „Coverage zur Geodatabase“ stehen sollen. Trotz Recherche in den entsprechenden Ressourcen des Internets konnten (mit Ausnahme der „case studies“ bekannter Softwareunternehmen) keine Anwendererfahrungsberichte gefunden werden, welche in den Rahmen dieser Arbeit gepasst hätten.

Daher kann lediglich auf Kapitel 6 verwiesen werden, in welchem in kurzer Form auf die Entwicklungen in den österreichischen Bundesländern bezug genommen wird.

4.7 Zusammenfassung

In den Dokumentationen zur Migration von Datenbeständen in eine Geodatenbank werden drei im Effekt gleichwertige Möglichkeiten aufgezeigt. Allen Varianten liegt zugrunde, dass zunächst das Design der Datenbank, d.h. die Modellierung durchzuführen ist. Die Anleitungen zur Modellierung einer Geodatenbank sind sehr allgemein gehalten. Der Grund liegt darin, dass die Modellierung sehr stark von den zu bearbeitenden Datenbeständen abhängt. Es ist daher wichtig, diesem Thema große Aufmerksamkeit zu schenken. Denn eine Datenbank, welche die in sie gestellten Ansprüche nicht erfüllt, verursacht hohe Kosten und wird von den Nutzern nicht akzeptiert. Die Vorgangsweise zur Erlangung eines geeigneten Datenmodells ist eine abgestufte und umfasst das konzeptuelle Schema, das logische Modell und das physikalische Modell.

Als Modellierungstechniken haben sich die ER-Modellierung (Entity-Relationship) und die Relationale Modellierung durchgesetzt. Durch die UML-Notation (Unified Modeling Language) wird die Modellierung von Objekten (wie sie in der objektorientierten Geodatabase zur Anwendung kommen) vereinfacht.

Unabhängig von der benutzten Modellierungstechnik ist der entscheidende Vorteil der Modellierungsphase, dass ein tieferes Verständnis für die Datenwelt eines Unternehmens entwickelt wird.

Aufgabe der Modellierung ist es, die Daten, welche in eine Anwendung beschreiben, in einem geeigneten Modell darzustellen. Im Rahmen des objektorientierten Modellierens wird die reale Welt als eine Vielfalt miteinander in Beziehung stehender Objekt aufgefasst.

Konkret wird zunächst ein ER Modell erstellt, welches dann in ein Relationales Modell (beide in UML Notation) umgewandelt wird. Die Objekte und ihre Eigenschaften werden in UML in Kästchen dargestellt (sogenannten compartments), Beziehungen zwischen diesen als Linien. Die Beziehungen können ebenfalls Eigenschaften aufweisen, welche mit eigenen Notationen dargestellt werden. Die wichtigsten Eigenschaften von Beziehungen zwischen Objekten sind Aggregation, Komposition, Spezialisierung und Generalisierung.

Die ER Modellierung dient dazu, das Problem oder die Aufgabenstellung zu analysieren, während die Relationale Modellierung auf die Implementierung abzielt.

Auf der Grundlage der Modellbildung wird ein generelles Ablaufschema aufgestellt, welches sich aus drei Teilbereichen zusammensetzt:

Die allgemeine Datenbankverwaltung, das konzeptionelle Schema und das logische Modell. Neben der Frage der Modellierung sind beim Thema Umstieg in der Datenhal-

tung die Bereiche Organisation, Schulung, Ausstattung und Zeitplan von besonderer Bedeutung.

Bei der Umsetzung des Umstiegs muss, vereinfacht gesagt, mit Bedacht und Augenmaß vorgegangen werden. Hilfestellung liefert dabei das Systems Engineering. Nach der Umsetzung müssen die gesetzten Schritte entsprechend dokumentiert und einer kritischen Überprüfung unterzogen werden.

5 Fallstudien

In den nächsten beiden Unterkapiteln werden die bisher erarbeiteten Ergebnisse auf ihre Praxistauglichkeit überprüft und der Nachweis angetreten, dass die angeführten bzw. vorgeschlagenen Vorgangsweisen auch durchaus angebrachte Ansätze für die Lösung des Problems darstellen.

5.1 Grundlagenthema: DKM

Die DKM (Digitale Katastral Mappe) gehört unbestrittenerweise zu den Grunddatenbeständen in den vielen Geographischen Informationssystemen. Sie spiegelt die Grundstückskonfiguration der Katastralgemeinden wider und ist somit ein unverzichtbares Mittel für raumbezogene Planungen, besonders im Fachbereich Raumplanung.

Die DKM setzt sich aus mehreren thematischen Inhalten zusammen, welche sich zweckmässigerweise in der Modellierung wiederfinden sollen.

An dieser Stelle soll das in Kapitel 4.2.4 aufgestellte Ablaufschema bzw. die in diesem Zusammenhang zweckmässigen Teile desselben zur Anwendung kommen.

5.1.1 Datenmodellierung

Die im Ablaufschema festgelegte Vorgangsweise wird ab Punkt 5 der Reihenfolge nach abgearbeitet. Die Modellierung erfolgt mit Hilfe von UML in der Form, wie sie in Kapitel 4.2.3 dargestellt wurde.

Nachfolgend eine graphische Darstellung über die Struktur des ursprünglichen DKM Datensatzes und ein Beispiel dafür, wie die einzelnen Inhalte dargestellt werden.

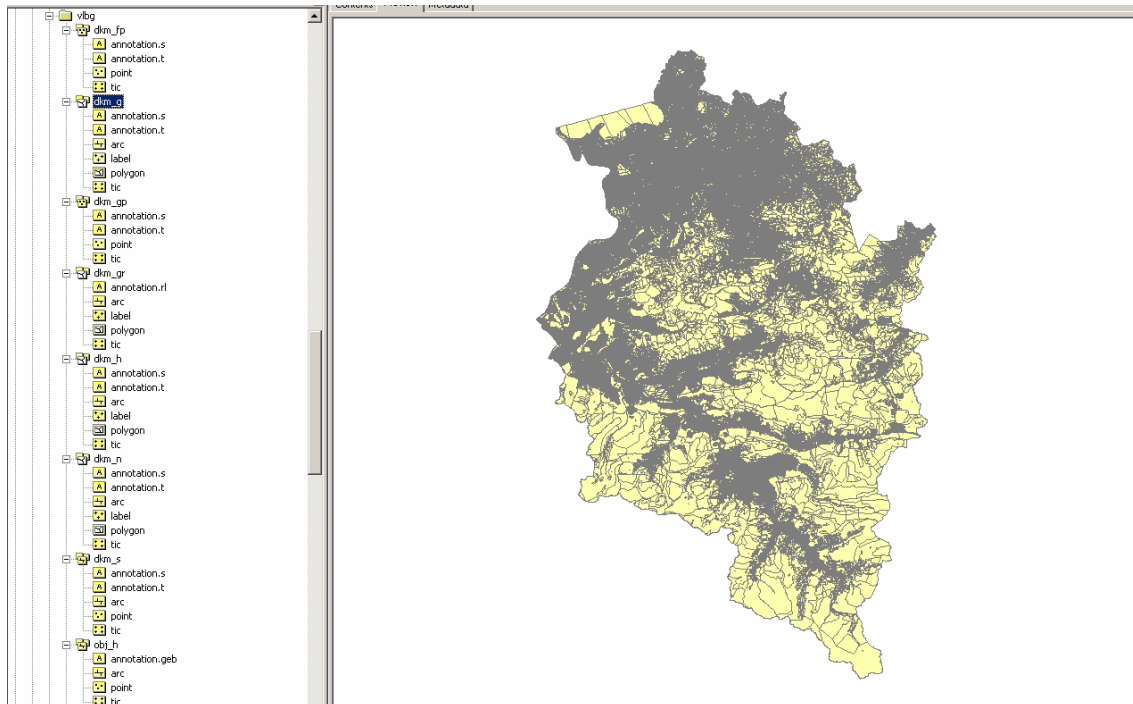


Abb.5.1: Darstellung des DKM Datenbestandes (Coverage) in ArcCatalog.

II.) Konzeptionelles Schema:

- 5) Erstellen der Datenmodelle zu jedem Datensatz (bzw. zu jeder Themenebene). Dazu zählt insbesondere die Definition der Objekte und deren Beziehungen untereinander.

Name	Type
dkm_fp	Coverage
dkm_g	Coverage
dkm_gp	Coverage
dkm_gr	Coverage
dkm_h	Coverage
dkm_n	Coverage
dkm_s	Coverage
obj_h	Coverage
dkm_i.dbf	dBASE Table

Abb.5.2: Ebenen (Coverages), welche das Thema DKM aufbauen (in ArcCatalog Darstellung).

Zum besseren Verständnis werden die einzelnen Datenebenen anhand von Beispielen in einem Kartenausschnitt markiert:

- ① dkm_fp...Festpunkte
- ② dkm_g...Grundstücksflächen mit Grundstücksnummern
- ③ dkm_gp...Grenzpunkte

- ④ dkm_gr...Katastralgemeindegrenzen
- ⑤ dkm_h...Gebäudegrenzen
- ⑥ dkm_n...Nutzungsgrenzen
- ⑦ dkm_s...sonstige Grenzen



Abb.5.3: DKM Ausschnitt mit Darstellung der unterschiedlichen Themenebenen.

Die einzelnen Coverages (bzw. deren Elemente) stehen in bestimmten Beziehungen zueinander, welche im Datenmodell entsprechend berücksichtigt werden müssen. Diese Beziehungen dienen als Grundlage für die Regeln, welche erstellt und implementiert werden müssen, um die topologische Integrität des Datenbestandes zu gewährleisten. Die Regeln könnten folgendermaßen lauten:

- Grundstücke liegen entlang ihrer Grundstücksgrenzen immer bündig aneinander. Es treten keine Überlappungen und keine Klaffungen auf.
- Für jene Elemente, welchen innerhalb der Modellierung topologische Funktionen zukommen sollen, werden solche Regeln definiert:

- Grenzpunkte liegen immer an jener Stelle, wo sich zwei oder mehr Teilgrenzen treffen. Sie liegen immer an den Enden der jeweiligen Teilgrenzen.
- Ein Grundstück weist immer eine Grundstücksnummer auf. Diese ist innerhalb einer Katastralgemeinde eindeutig.
- Katastralgemeinden liegen entlang ihrer Katastralgemeindegrenzen immer bündig aneinander. Es treten keine Überlappungen und keine Klaffungen auf.
- Gebäudegrenzen liegen immer innerhalb eines Grundstückes. Gebäudeteile, welche auf ein Nachbargrundstück reichen, werden rechtlich diesem Nachbargrundstück zugerechnet.

Nutzungsgrenzen und sonstige Grenzen weisen hinsichtlich des Grundstückskatasters nur untergeordnete rechtliche Bedeutung auf und werden daher in der Definition von Regeln nicht berücksichtigt.

Die Regeln werden in weiterer Folge für die Umsetzung vom konzeptionellen Schema in das logische Schema zur Implementierung der topologischen Information benötigt.

Nachfolgend ist das Entity-Relationship-Modell der DKM in UML Notation angeführt. Dieses Modell wird anschließend in das Relationale Modell (es baut auf Tabellen auf, daher Relational und ist damit die Basis für die Umsetzung) überführt.

Siehe dazu auch die Abbildungen 4.19 und 4.20.

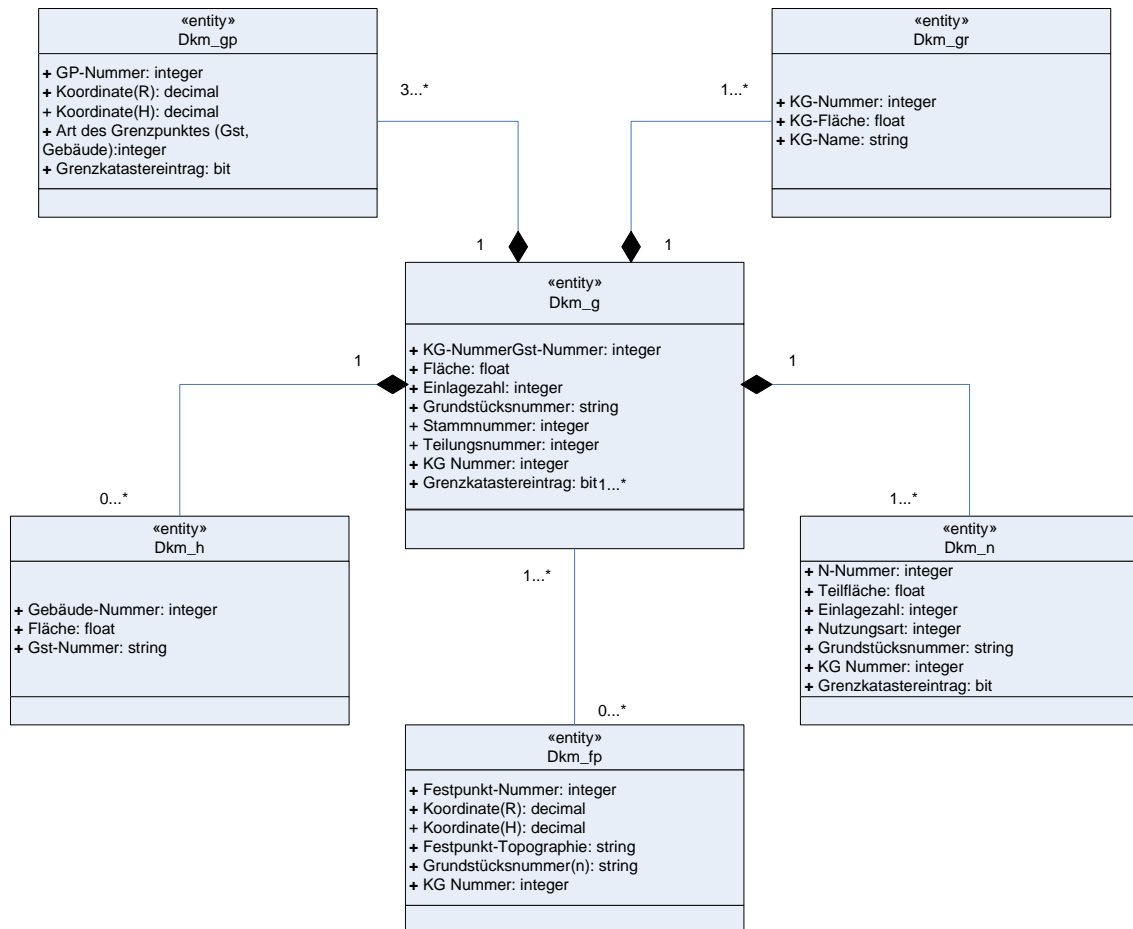


Abb.5.4: ER-Modell in UML Notation (PK...Primary Key, NN...Not Null).

6) Auswahl der tatsächlich zu übernehmenden Daten (z.B. Tics, arcs bei Routen etc.), externe Attributtabellen

Wie bereits erwähnt, werden in den neuen Datenbestand nur Daten überführt, welche aus thematischer Sicht, und nicht aus Gründen der Notwendigkeit für die Datenstruktur, notwendig sind. Dazu gehören „tics“ oder auch „labelpoints“ und „arcs“ für die Darstellung von Flächen, weil diese im Geodatabase-Format mit einer anderen Struktur dargestellt werden. Grundsätzlich muß diese Entscheidung jeder Nutzer, abhängig vom Inhalt und der Struktur seiner Daten, für sich treffen.

7) Definition und Dokumentation der für die einzelnen Objekte benötigten Datentypen.

Hier erfolgt der nächste wesentliche Schritt in der Modellierung. Zu diesem Zweck kann das oben erarbeitete Schema verwendet werden. Es wird (diesmal in Relationaler Schreibweise mit UML) um die Datentypen ergänzt.

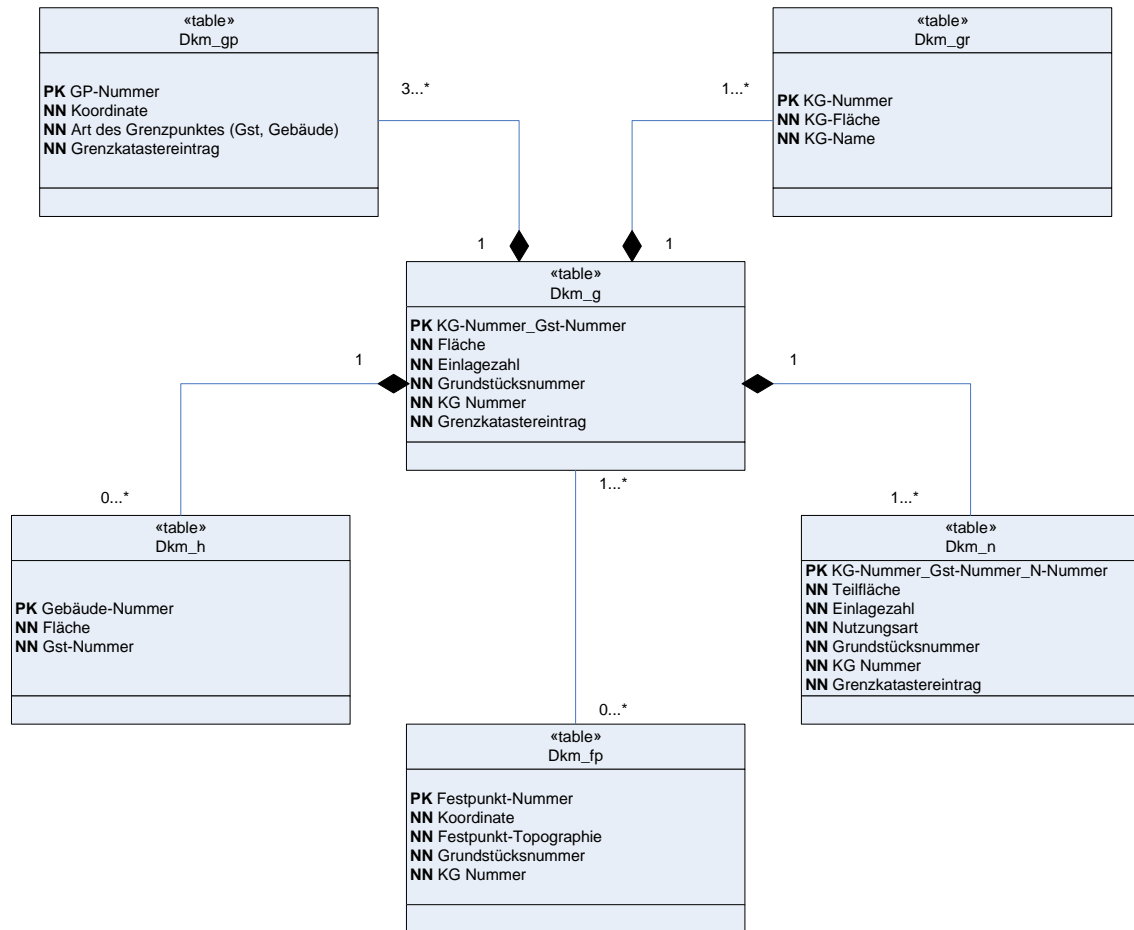


Abb.5.5: Relationales Modell in UML Notation (+...Sichtbarkeitsspezifikation für die Entität; + bedeutet public).

8) Definition und Dokumentation der zu verwendenden Projektionen, der Genauigkeit und der abzubildenden Ausdehnung (Extent)

Diese Festlegungen hängen von den jeweiligen Anforderungen an den Datenbestand ab und sind für den Datenbestand (bzw. für jedes geographische Informationssystem) gesondert festzulegen. In der Regel wird jedoch diese Festlegung nur einmal getroffen und für alle Datenbestände derselbe Extent (d.h. die räumliche Ausdehnung, welche durch die Daten abgedeckt werden), dieselbe Projektion und dieselbe Genauigkeitsfestlegung verwendet. Diese werden in einem sogenannten feature-dataset (siehe unten) vorgegeben und sind für all jene Daten gültig, welche sich in diesem befinden.

III.) Logisches Modell:

9) Anlegen der zuvor festgelegten Datenbanken und Datenhülsen und Eingabe der vordefinierten Eigenschaften (Genauigkeit, Extent u.ä.)

Für das DKM Beispiel wird in ArcCatalog ein sogenanntes „feature-dataset“ angelegt, welches die einzelnen Datenbestände, die „feature-classes“ enthält. Auf der Ebene des feature-dataset werden auch die oben genannten Ausdehnungen, die Genauigkeiten und die verwendete Projektion festgelegt.

10) Befüllen der Datenbank mit Hilfe von ArcCatalog oder batch-gesteuerten ArcSDE Kommandos.

Dieser Prozess kann entweder manuell oder scriptgesteuert durchgeführt werden. Auf eine detaillierte Beschreibung wird an dieser Stelle verzichtet. Die einschlägigen Anleitungen enthalten ausführliche Informationen zur Durchführung dieses Schrittes. Im Anhang dieser Arbeit befinden sich Beispiele der wichtigsten Kommandos für den batchgesteuerten Modus.

11) Definition und Eingabe der Regeln, der Attributdomänen und der Rangreihung (rank)

Folgende Regeln wurden für den DKM Datenbestand angegeben:

- Für dkm_g (Grundstücke): Must not overlap (die Grundstücke dürfen sich nicht überlappen), Must be covered by (die Grundstücke überdecken die Katastralgemeinde vollständig), Must not have gaps (die Grundstücke dürfen keine Lücken bilden)
- Für dkm_gp (Grenzpunkte): Must be covered by boundary of (Grenzpunkte müssen auf einer Grundstücksgrenze liegen)
- Für dkm_gr: Must be covered by feature class of dkm_g (Grundstücke müssen die Katastralgemeinde vollständig überdecken)
- Für dkm_h: Must be covered by
- Für dkm_n: Must be covered by (Nutzungsgrenzen überdecken Grundstücke ganz oder teilweise)
- Für dkm_s: Must be covered by boundary of (Sonstige Grenzen müssen innerhalb der Grenzen eines Grundstückes liegen).

Ein ausgesprochen wichtiger Parameter in der Definition der Beziehungen zwischen den einzelnen Datenbeständen ist die Angabe der Rangreihung (rank). Ein Datenbestand mit hoher Rangreihung (d.h. mit niedrigem Wert)

ist ein Datenbestand, dessen geometrische Genauigkeit hoch ist. Ein Datenbestand mit niedriger Rangreihung (das ist ein hoher Wert) weist eine geringere geometrische Genauigkeit auf. Bei Verarbeitungsprozessen werden daher Daten mit niedrigerem Rang, welche innerhalb der sogenannten „Cluster“-Toleranz zu liegen kommen, zu jenen mit hohem Rang hingezogen.

5.1.2 Schlussfolgerungen

In diesem Unterkapitel wird beschrieben, wie mit Hilfe eines Ablaufschemas zunächst die Modellierung und anschliessend die Umarbeitung eines Datenbestandes von der filebasierten Datenhaltung (Coverage) in die datenbankbasierte Datenhaltung (Geodatabase) vonstatten gehen kann. Im Ablaufschema werden die einzelnen Schritte von der konzeptionellen Ausgestaltung bis zur logischen Umsetzung behandelt.

Es muß festgehalten werden, dass diese Anleitung aufgrund der Komplexität des Umfeldes sehr grob und oberflächlich bleiben muss. Viele Details müssen in Zusammenarbeit mit den Nutzern der Daten und durch iteratives Durchspielen von Möglichkeiten hinsichtlich der Beziehungen zwischen denselben ermittelt werden.

Insbesondere die Festlegung der Topologieregeln erfordert eine gewissenhafte Vorgangsweise und letztlich viel Erfahrung.

5.2 Fachthema: Fließgewässer

5.2.1 Datenmodellierung

Das Thema Fließgewässer wurde gewählt, weil es sich um ein Routensystem ergänztes Linienthema handelt, welches sich mit einer Vielzahl von ergänzenden Informationen verknüpfen lässt. In dem nachfolgend angeführten Beispiel werden die Ordnungszahlen und die Kilometrierung in unterschiedlichen Ausprägungen für die Darstellung der Möglichkeiten herangezogen.

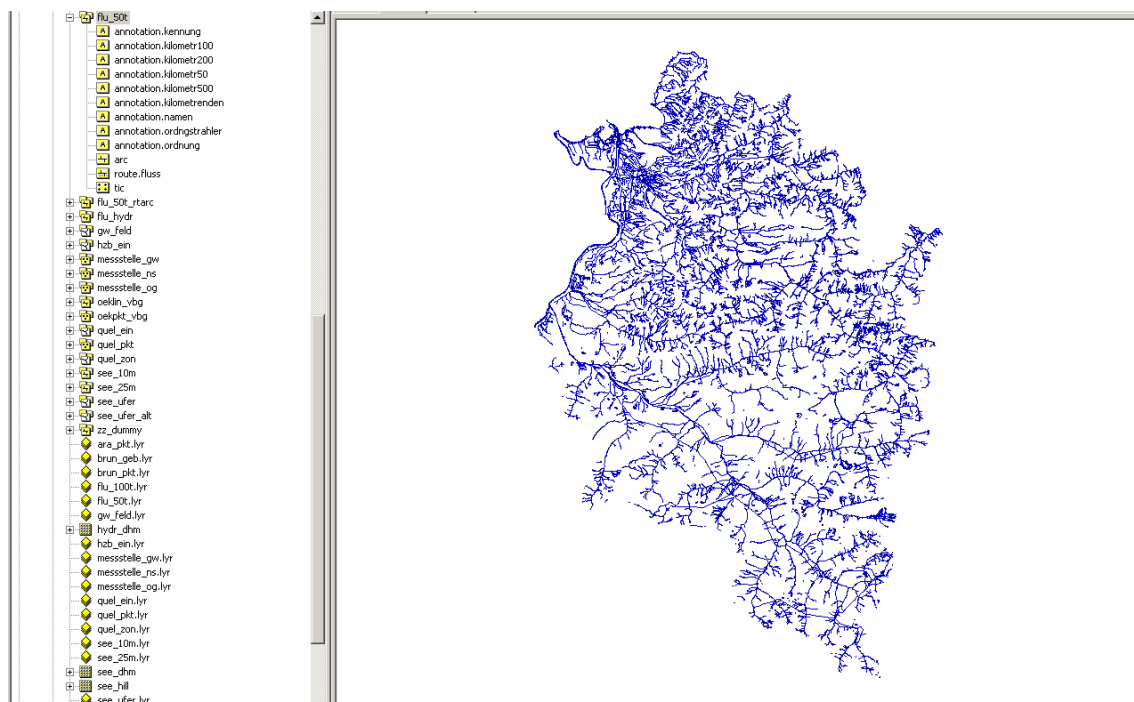


Abb.5.6: Darstellung des Fließgewässer Datenbestandes (Coverage) in ArcCatalog

II.) Konzeptionelles Schema:

- 5) Erstellen der Datenmodelle zu jedem Datensatz (bzw. zu jeder Themenebene). Dazu zählt insbesondere die Definition der Objekte und deren Beziehungen untereinander.

Dieser Datenbestand bildet eine Besonderheit bei der Überführung vom Coverage zur Geodatabase, weil er (bedingt durch die Darstellungsmöglichkeiten im Coverage) sehr viele Annotationen enthält. Verschiedenste Inhalte wie Flußbezeichnungen, Kilometrierungen, Darstellung der Ordnungszahlen etc. wurden als Beschriftungen („annotation“)

umgesetzt. Im Geodatabase-Modell wird diese Variante weitgehend durch die Möglichkeit die Beschriftung aus dem Attributdatenbestand („labelling“) zu realisieren, ersetzt. Es ist aber notwendig, die entsprechenden externen Attributtabellen zu generieren und diese über Eventthemen zum Grunddatenbestand zu verknüpfen.













Name	Type
 annotation.kennung	Annotation Feature Class
 annotation.kilometr100	Annotation Feature Class
 annotation.kilometr200	Annotation Feature Class
 annotation.kilometr50	Annotation Feature Class
 annotation.kilometr500	Annotation Feature Class
 annotation.kilometrenden	Annotation Feature Class
 annotation.namen	Annotation Feature Class
 annotation.ordngstrahler	Annotation Feature Class
 annotation.ordnung	Annotation Feature Class
 arc	Arc Feature Class
 route.fluss	Route Feature Class
 tic	Tic Feature Class

Abb.5.7: Ebenen (Coverages), welche das Thema Fließgewässer aufbauen (in ArcCatalog Darstellung).

Nachfolgende Darstellung zeigt die die ER-Modellierung in UML Notation zum Datenbestand Fließgewässer. Dabei sind beispielhaft zwei Attributtabellen mit dem Routenthema verknüpft. Über diese Attributtabellen werden in der Darstellung mit Hilfe von Eventthemen die entsprechenden Inhalte (sprich die Ordnungsnummern und die Hektometerwerte) visualisiert.

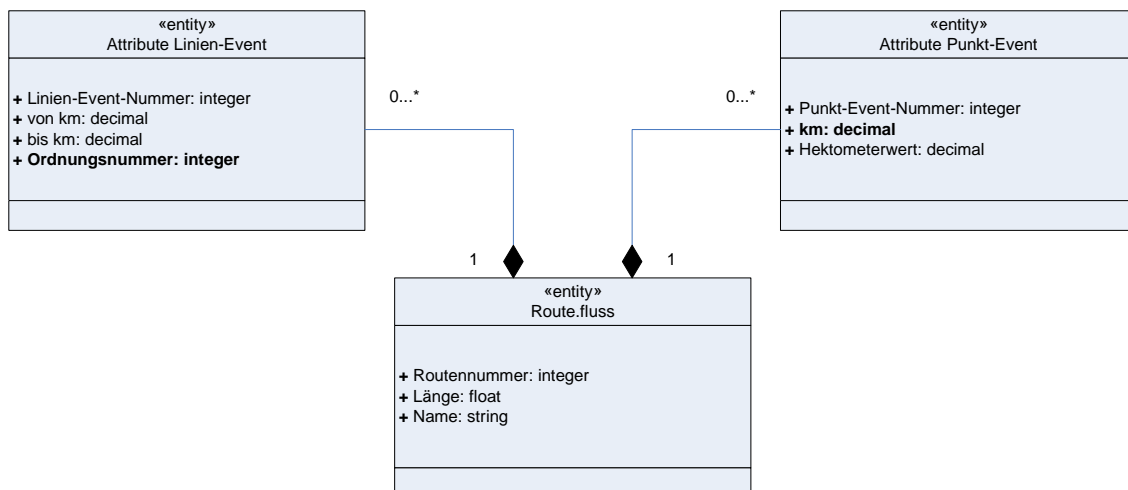


Abb.5.8: ER-Modell in UML Notation (PK...Primary Key, NN...Not Null).

6) Auswahl der tatsächlich zu übernehmenden Daten (z.B. Tics, arcs bei Routen etc.), externe Attributtabelle – siehe 5.1.1

7) Definition und Dokumentation der für die einzelnen Objekte benötigten Datentypen

Ausgehend vom ER-Modell werden im Relationalen Modell die einzelnen Datentypen festgelegt. Die angeführten Attribute sind als Beispiele zu verstehen, welche in beinahe beliebiger Anzahl definiert werden können. Als zweckmässig hat sich jedoch eine Anzahl herausgestellt, welche leicht erfasst werden kann; im Zweifelsfall sollte auf nicht ständig benötigte Attribute verzichtet werden.

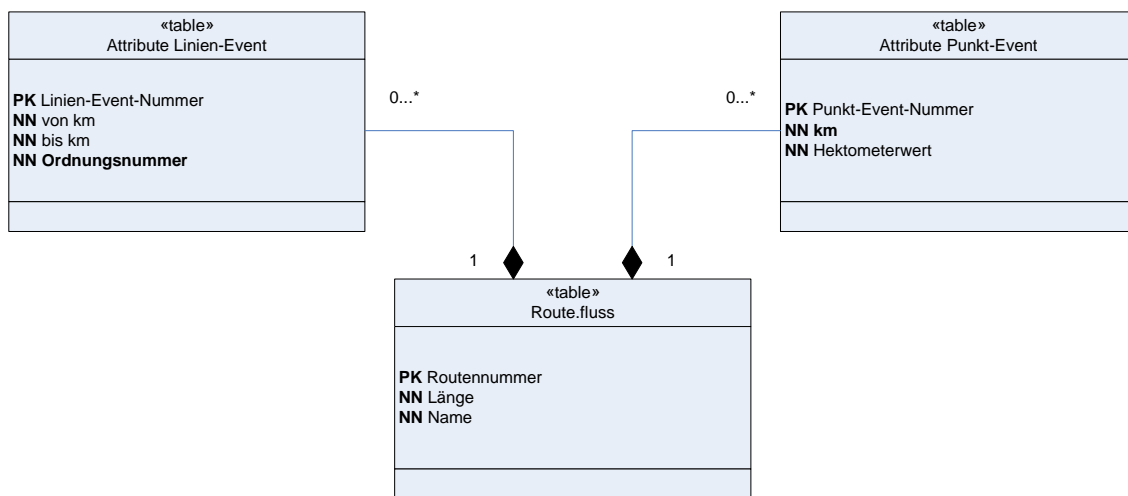


Abb.5.9: Relationales Modell in UML Notation (+...Sichtbarkeitspezifikation für die Entität; + bedeutet public).

8) Definition und Dokumentation der zu verwendenden Projektionen, der Genauigkeit und der abzubildenden Ausdehnung (Extent) - siehe 5.1.1

III.) Logisches Modell:

9) Anlegen der zuvor festgelegten Datenbanken und Datenhüllen und Eingabe der vordefinierten Eigenschaften (Genauigkeit, Extent u.ä.)

10) Befüllen der Datenbank mit Hilfe von ArcCatalog oder batch-gesteuerten ArcSDE Kommandos.

Dieser Prozess kann entweder manuell oder scriptgesteuert durchgeführt werden. Auf eine detaillierte Beschreibung wird an dieser Stelle verzichtet. Die einschlägi-

gen Anleitungen enthalten ausführliche Informationen zur Durchführung dieses Schrittes.

11) Definition und Eingabe der Regeln und der Attributdomänen:

- Must not overlap (die Fliessgewässer dürfen sich nicht überlappen)
- Must not self overlap (die Fliessgewässer dürfen sich nicht selbst überlappen)
- Must not intersect (die Fliessgewässer dürfen sich nicht überschneiden – eine Einmündung ist dagegen möglich).

Bezüglich der Rangreihung (rank) ist es nicht ganz so einfach, die richtige Einordnung zu finden. Es müssen jedenfalls alle im Datenbestand mit dem gegenständlichen Thema in potentieller Beziehung stehende Ebenen auf ihre Rangreihung hin untersucht werden, um unbeabsichtigte geometrische Verzerrungen zu verhindern.

5.2.2 Schlussfolgerungen

Dieses Beispiel zeichnet sich im Vergleich mit dem DKM Beispiel nicht nur dadurch aus, dass es sich um ein Linienthema handelt, sondern und vor allem dadurch, dass es sich durch das Routensystem dazu eignet, allein durch ergänzende Attributdaten neue Visualisierungen (und damit die Generierung neuer Datenbestände) zu ermöglichen. Das Datenmodell wird dadurch leichter und einfacher konzipier- und umsetzbar. Dennoch ist auch hier die Notwendigkeit gegeben, durch die Erstellung von Regeln und die Zuordnung einer Rangreihung die Beziehungen zu anderen Datenbeständen im Gesamtdatensatz zu definieren. Hier offenbaren sich auch die Schwierigkeiten in der Praxis, schon im Vorhinein geeignete Regeln zu definieren und eine im Hinblick auf die spätere Nutzung zweckmässige Zuteilung der Rangreihung zu finden.

In der einschlägigen Literatur [ESR02bg] wird darauf hingewiesen, dass diese Topologie, die als Verhalten von Elementen und Regeln implementiert ist, einen flexibleren Satz geometrischen Beziehungen zu modellieren erlaubt, als jene Topologie, die in der Datenstruktur bereits enthalten ist.

Es sei angemerkt, dass erstere sicher zu einer reicheren und vielgestaltigeren Definition von Beziehungen zwischen den Datenebenen führen kann als letztere, jedoch für den Preis einer Komplexität, welche sogar bei eingehender Beschäftigung mit dem Thema noch schwer zu fassen ist.

Bislang war Logik des Datenbestandes in der Datenstruktur selbst gegeben, nunmehr muß diese durch Maßnahmen wie Regeln, Domänen und Vorgabewerte selbst definiert werden – dies erfordert mehr Wissen und größere Gewissenhaftigkeit.

5.3 Erfahrungen aus der Anwendung des Ablaufschemas

Die Erfahrungen aus der Anwendung des Ablaufschemas (für die Umstellung von Datenbeständen) lassen sich wie folgt zusammenfassen:

- Das Ablaufschema kann nicht ohne Grundlagenwissen angewendet werden. Kenntnisse über die Modellierung von Daten, die einzelnen Fachthemen und die entsprechenden Anwendungen sind notwendig.
- Ein Pauschalrezept existiert nicht – die umfassende Verallgemeinerung einer Vorgangsweise würde sehr umfangreich werden und damit wieder den Charakter einer Dokumentation erhalten
- Es gibt mehrere Varianten, die Konzeption umzusetzen
- UML ist eine zukunftsweisende Vorgangsweise in der Modellierung von Geodatenbanken.
- Der Ablaufplan kann uns muss an die entsprechende Aufgabe angepasst werden. Dies kann Verringerung aber auch Ausbau der einzelnen Schritte bedeuten.
- Es hat sich als zweckmässig herausgestellt, die Vorgangsweise auf breiter Basis zu diskutieren (siehe dazu oben: Teambildung)
- Die Anwendung des Systems Engineering ist ein wesentlicher Aspekt hinsichtlich der Plastizität des Umwandlungsprozesses. Änderungen sind trotz bester Planung die Regel, nicht die Ausnahme.

5.4 Zusammenfassung

Die in den vorigen Kapiteln erarbeiteten Vorschläge für eine zweckmässige Umstellung von Coverages zu einer Geodatabase wird anhand von Beispielen aufgezeigt.

Dazu werden zwei konkrete Datenbestände, nämlich die Digitale Katastralmappe (kurz DKM) und ein Fliessgewässernetz (hier jenes des Bundeslandes Vorarlberg) verwendet. Zunächst wird mit Hilfe des im Kapitel 4.2.4 erarbeiteten Ablaufschemas die Datenmodellierung durchgeführt. Ausgehend von einer mit UML Notation durchgeführten ER-Modellierung wird ein Relationales Modell (ebenfalls in UML Notation) erstellt. Dieses Relationale Modell dient als Vorlage für die Erstellung der einzelnen Tabellen, welche mit den entsprechenden Werkzeugen direkt in der Geodatabase angelegt werden können. Weiters werden die für die Erstellung topologischer Beziehungen notwendigen Regeln, Attributdomänen und Rangreihungen definiert. Die Festlegung dieser Topologieregeln erfordert eine gewissenhafte Vorgangsweise, weil dadurch die Funktionalitäten des Datenbestandes wesentlich beeinflusst werden.

Unbestritten ist, dass es durch die Möglichkeiten einer Geodatabase zu einer reicheren und vielgestaltigeren Definition von Beziehungen zwischen Datenebenen kommen kann; der Preis dafür ist jedoch eine Komplexität, welche sogar bei eingehender Beschäftigung mit dem Thema nur schwer zu fassen ist.

Die Erfahrungen aus der Anwendung des Ablaufschemas lassen sich in den wichtigsten Punkten wie folgt zusammenfassen:

- Grundlagenwissen zum Thema Geodatabase muss zur Anwendung des Ablaufschemas vorhanden sein
- Ein Pauschalrezept existiert nicht
- Es gibt mehrere Varianten, die Konzeption umzusetzen
- Es hat sich als zweckmässig herausgestellt, die Vorgangsweise auf breiter Basis zu diskutieren.

6 Status in den neun Bundesländern

6.1 Derzeitiger Stand der Entwicklung

Die Erfahrungen aus der in dieser Arbeit skizzierten Tätigkeiten zur Umsetzung eines Datenbestandes vom Coverage zur Geodatabase haben gezeigt, dass neben einer eingehenden Beschäftigung mit dem Datenbestand selbst, auch das gesamte Umfeld einer detaillierten Betrachtung unterzogen werden muß. Dabei handelt es sich insbesondere um die Ansprüche der Nutzer und die von ihnen gepflegten Anwendungen.

Alles in allem sind die Aufwendungen trotz der vielen Assistenten, welche die Softwarehersteller in ihre Produkte integriert haben, ausgesprochen hoch.

Es kann daher vermutet werden, dass die Umstellung, so sie nicht aus einer Notwendigkeit heraus in Angriff genommen werden muss, nicht zu den prioritär behandelten Aufgaben zu zählen ist.

Eine Nachfrage nach dem Stande der Entwicklungen in den einzelnen Bundesländern hat unterschiedliche (und unergiebig) Ergebnisse gebracht.

Während in einigen Bundesländern (darunter auch Vorarlberg) zunächst über mögliche Szenarien nachgedacht wird, ist in Tirol derzeit ein Projekt in Ausarbeitung, welches die Umstellung des gesamten Tiris Datenbestandes zu einer Geodatabase zum Ziel hat.

Die Umstellungsphase ist auf mehrere Monate ausgelegt, wobei die Datenübernahme parallel zum Betrieb des bisherigen Systems vorgenommen wird.

In den meisten anderen Bundesländern wird derzeit die Geodatabase auf Grundlage von ArcSDE als Server von Web-Diensten verwendet.

Eine vollständige Umstellung der Datenbestände von filebasierter auf datenbankbasierte Datenhaltung ist nach den vorliegenden Informationen noch in keinem Bundesland erfolgt!

7 Schlußfolgerungen

„Vom Coverage zur Geodatabase“ behandelt wesentliche Gesichtspunkte, welche bei der Umstellung von Daten von einem Datenmodell in das andere zur berücksichtigen sind. Dazu gehören die Untersuchung des Coverage-Datenmodells mitsamt seinen wesentlichen Merkmalen und die Erörterung des Geodatabase-Modells.

Dabei stellen sich einige Besonderheiten heraus, welche am übersichtlichsten in einer Liste aus **Merksätzen** dargestellt werden können:

- Die Kenntniss der beiden Datenmodelle (coverage und geodatabase) sind von großem Vorteil bei der Konzeption des Umstieges.
- Die Konzeption des Geodatabase Datenmodelles erfordert eine intensive Beschäftigung mit den geplanten Anwendungen und den zu überführenden Daten.
- Das Erlernen und Anwenden der UML Notation für die Datenmodellierung erleichtert die Konzeption der Geodatabase.
- Die Geodatabase erfordert in jeder Hinsicht einen hohen Ressourcenaufwand (Schulung, Hardware, Software, Organisation) – dieser ist jedenfalls zu berücksichtigen.
- Es kann derzeit noch nicht auf ein reiches Reservoir an Anwendererfahrungen (mit ähnlichen Ansprüchen) zurückgegriffen werden – daher eine Strategie der kleinen Schritte wählen.
- Die Einbeziehung der Betroffenen (im engeren Sinne die Anwender) entweder in einem Team oder innerhalb einer Projektorganisation ist notwendig.
- Die Beherrschung der Software zur Erstellung einer Geodatabase ist bereits im Vorfeld des Umstieges ratsam.
- Und „last but not least“ der Trost für alle, die trotz Berücksichtigung aller gut gemeinten Ratschläge den harten Weg der Änderung einer Geodatabase beschreiten müssen: Änderungen sind trotz bester Planung die Regel, nicht die Ausnahme.

8 Literaturverzeichnis

- [BOZ04]: www.bozeman.net/gisweb/data/other/www/Geodatabase.htm
- [DFH03]: DISTERER, Georg/FELS, Friedrich/HAUSOTTER, Andreas: Taschenbuch der Wirtschaftsinformatik, Fachbuchverlag Leipzig, 2.A., 2003
- [DL00]: DOPPLER, Klaus/LAUTERBURG, Christoph: Change Management, Campus Verlag Frankfurt/New York, 2000
- [DÖ01]: DÖRNER, Dietrich: Die Logik des Mißlingens, RoRoRo Hamburg, 2001
- [EUR03]: STARTUP EUREGIO MANAGEMENT GMBH: Schulungsunterlagen zum PM Lehrgang der Vorarlberger Landesverwaltung 2003, unveröffentlicht
- [SCHIO3a]: SCHILCHER, M.: Geoinformatik I, Kapitel 4: Modellierung und Datenmodelle, Vorlesungsunterlagen TUM, Sommersemester 2003, S 1ff, München.
- [SCHIO3b]: SCHILCHER, M.: Geoinformatik I, Kapitel 5: Modellierung und Datenmodelle, Vorlesungsunterlagen TUM, Sommersemester 2003, S 1ff, München
- [SW01]: SCHNEIDER, Uwe/WERNER, Dieter: Taschenbuch der Informatik, Fachbuchverlag Leipzig, 4.A., 2001
- [HNBBM94]: HABERFELLNER/NAGEL/BECKER/BÜCHEL/VON MASSOW: Systems Engineering, Verlag Industrielle Organisation Zürich, 8.A., 1994
- [ZEI99]: ZEILER, Michael: Modeling our World. The ESRI Guide to Geodatabase Design, ESRI Press, Redlands, California, 1999
- [ESR94]: ARC/INFO Data Management. Concepts, data models, database design, and storage, ESRI Press, Redlands, California, 1994
- [ESR03]: ESRI: ArcGIS: Working with Geodatabase Topology, ESRI Whitepaper, 2003
- [ESR02bg]: ESRI: Building a Geodatabase, 2002
- [ESR02gw]: ESRI: Geodatabase Workbook, 2002
- [ESR04SDE]: ESRI: Understanding ArcSDE, 2004
- [ESR04CS01]: ESRI: Case Studies
<http://support.esri.com/index.cfm?fa=downloads.dataModels.caseStudies> (04-07-13)
- [ESR04CS02]: ESRI Canada: Case Studies
<http://www.esricanada.com/english/government/case.asp> (04-07-13)
- [RSV02]: RIGAUX Phillipe, SCHOLL Michel, VOISARD Agnes: Spatial Databases with application to GIS, Morgan Kaufmann Publishers, San Francisco, 2002

[SVDW93]: STROHSCHNEIDER Stefan, VON DER WETH Rüdiger: Ja, mach nur einen Plan, Verlag Hans Huber Bern, 1993

[THOM96]: THOMMEN Jean-Paul: Betriebswirtschaftslehre, Band 3, Verlag Versus Zürich, 1996

[UGIS02a]: Modul 3 - Räumliche Daten: Modelle und Strukturen, UNIGIS Schulungsunterlage

[UGIS02b]: Modul 5 – Attributdaten und Datenbanksysteme, UNIGIS Schulungsunterlage, SEEMAYER Ewald, Salzburg 2002.

[UML_QRC01]: GRÉGOIRE Laurent, UML Quick Reference Card, 2001, <http://tnerual.eriogerg.free.fr/umlqrc.pdf>

[UML_Dab00]: DABIC Ivan, Vergleich von UML zu den Vorgängermethoden, Referat an der Goethe-Universität, Frankfurt am Main 2000

[WÖHE90]: Wöhe Günter, Einführung in die Allgemeine Betriebswirtschaftslehre, 17. A., München 1990

[GOR03a]: GORNIK Davor, Entity Relationship Modeling with UML, Rational-Software Whitepaper, IBM 2003

[GOR03b]: GORNIK Davor, Relational Modeling with UML, Rational-Software Whitepaper, IBM 2003

[NN03a]: NN, Mapping Object to Data Models with the UML, Rational-Software Whitepaper, IBM 2003

9 Glossar

DatenmodellDie Gesamtheit der Beschreibung der Entity-Sets mit ihren Attributen sowie die grafische Darstellung der Entity-Sets mit ihren Beziehungen wird als Datenmodell bezeichnet. Ein Datenmodell ist die Grundlage für den Entwurf relationaler Datenbanksysteme.

DBMS.....Ein Datenbankverwaltungssystem (DBMS) ist die Gesamtheit aller Programme (Ressourcen) zur Erzeugung, Verwaltung (einschließlich Daten- und Konsistenzsicherung) und Manipulation einer Datenbank. [SW01]

ER-ModellEntity-Relationship-Modell. Semantische Datenmodelle (SDM) unterstützen die Informationsmodellierung und damit die Formulierung des konzeptuellen Schemas. [SW01]

SDESpatial Database Engine. Eine Middleware, welche geographische Daten aus GIS-Anwendungsprogrammen in Relationalen Datenbanken speichert.

TINTriangular Irregular Network. TIN repräsentieren Oberflächen als kontinuierliche einander nicht überlappende Dreiecke.

UMLUnified Modelling Language.