



geoinformation.net

Projektpartner: Universität Bonn -
Institut für Kartographie
und Geoinformation
Lehrstuhl Kartographie
Datum: 23.12.2003

4. Realweltmodellierung und Datenmodelle

Einleitung

Ziel dieser Lerneinheit ist es zu erläutern, wie raumbezogene Sachverhalte modelliert werden, damit sie in einem GIS genutzt werden können. Dabei wird zunächst auf die Datenmodellierung im Allgemeinen eingegangen. Dort steht vor allem das Vorgehen nach der 3-Schema-Datenmodell-Architektur im Vordergrund. Objekte der realen Welt werden zum einen durch raumbezogene Eigenschaften und zum anderen durch semantische Eigenschaften beschrieben. Die raumbezogenen Eigenschaften werden durch die Topologie und Geometrie beschrieben. Während die Topologie dafür sorgt, dass die Objekte einer Raumlogik (Nachbarschaftsbeziehungen) unterliegen, umfasst die Geometrie eines Objektes die Beschreibung der Form und seiner Lage. Die semantische Modellierung kennt vor allem die Verfahren der Ebenen und der Objektklassen, auf die im Folgenden näher eingegangen wird. Abschließend wird ein Beispiel der Modellierung von amtlichen Geodaten auf Grundlage der AFIS-ALKIS-ATKIS-Datenmodellierung vorgestellt.

Inhalt

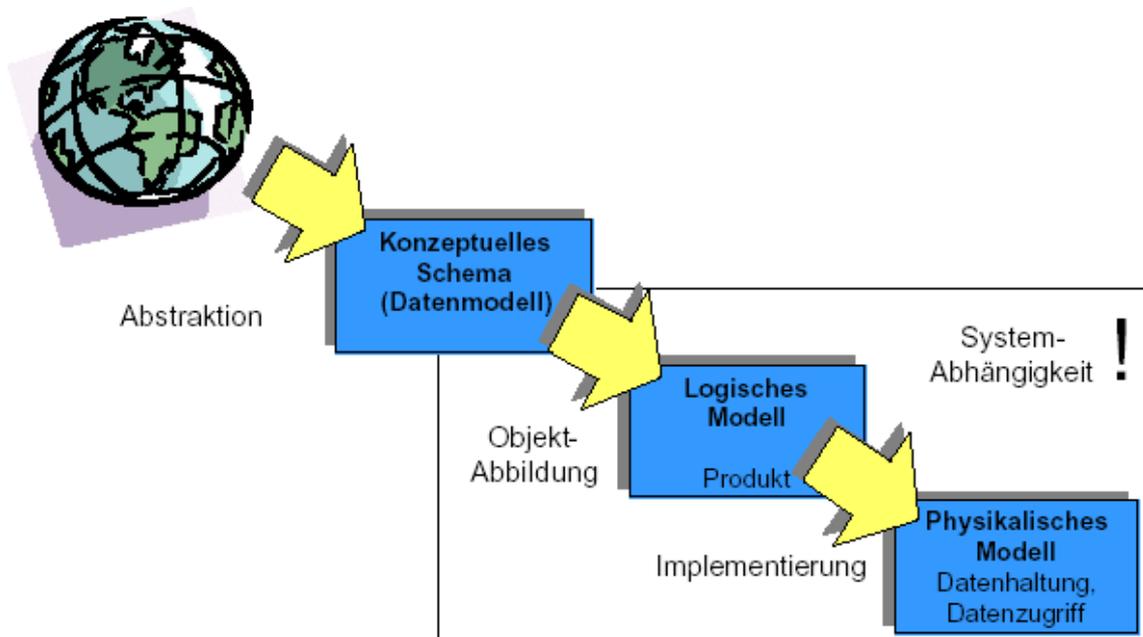
4. Realweltmodellierung und Datenmodelle

4.1 Allgemeine Aspekte der Modellierung	3
4.2 Topologische Modellierung.....	3
4.2.1 Topologische Invariante.....	4
4.2.2 Topologische Beziehungen und Konsistenzbedingungen	4
4.3 Geometrische Modellierung	5
4.3.1 Datentypen.....	5
4.3.1.1 Vektordaten.....	5
4.3.1.2 Rasterdaten	6
4.3.2 Modellierung der geometrischen Daten	6
4.4 Semantische Modellierung	8
4.4.1 Ebenenprinzip	8
4.4.2 Objektklassenprinzip	8
4.4.2.1 Semantischer Baum.....	9
4.4.2.2 Semantisches Netzwerk.....	9
4.5 AFIS-ALKIS-ATKIS-Datenmodellierung	10
4.5.1 Datenmodelle in Informationssystemen	10
4.5.2 Das konzeptuelle Datenmodell.....	10
4.5.2 Das konzeptuelle Datenmodell.....	11

4.1 Allgemeine Aspekte der Modellierung

Ziel der Datenmodellierung ist die Abbildung der Realwelt in ein Modell, um es in einem Computersystem verarbeiten zu können. Das Datenmodell dient der einheitlichen Abbildung von Realweltobjekten unter fachspezifischen Gesichtspunkten. Es bildet sowohl die Realweltobjekte als auch ihre Beziehungen untereinander sowie ihr Verhalten ab. Zur Modellierung der Realweltobjekte existieren Festlegungen wie Definitionen, Beschreibungen und Regeln, die bei der Erfassung angewendet werden.

Die Modellierung der realen Welt erfolgt in 3 Stufen.



Drei-Stufen-Modell der Modellierung [Quelle: Schilcher 2003]

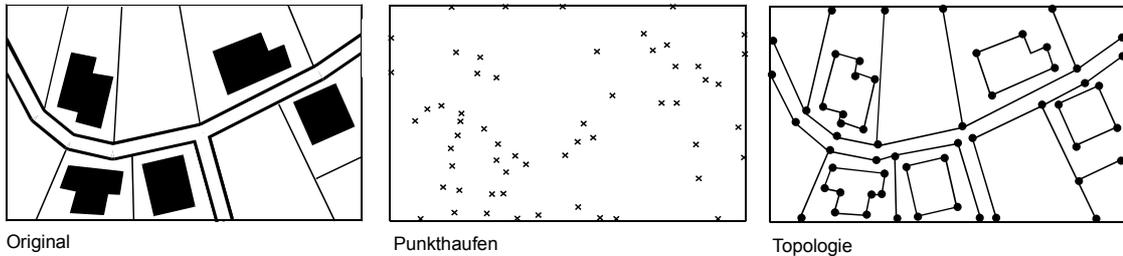
Zunächst werden die Datenstrukturen, wie z.B. die geometrische Modellierung oder die Festlegung der thematischen Hierarchie, in einem **konzeptuellen Modell** festgelegt. Dieses Modell ist im theoretischen Idealfall unabhängig von den verwendeten Applikationen, GIS- oder Datenbanksystemen. Unter der Restriktion der zur Verfügung stehenden Computersysteme wird ein **logisches Modell** entwickelt, das die Umsetzung des konzeptuellen Modells in eine vorher definierte Systemumgebung darstellt.

Im **physikalischen Modell** erfolgt die eigentliche Datenverwaltung in einem Datenverwaltungssystem, welches auf Tabellen-, Baum- und Listenstrukturen als physikalische Organisationsform aufsetzt. Auf dieses Modell hat der Entwickler meist wenig Einflussmöglichkeiten, da die grundlegenden Prinzipien von den Programmen selbst bestimmt werden.

4.2 Topologische Modellierung

Die topologische Modellierung umfasst die Beschreibung, Bearbeitung und Speicherung der Geometrie der Lage von raumbezogenen Objekten, unter Beachtung der topologischen Invarianten und Konsistenzbedingungen. Durch die Vereinigung der metrischen und topologischen Daten können umfangreiche Datenmodelle entstehen. Durch die Berücksichtigung der **Topologie** kann ein Objekt geschlossen geometrisch wiedergegeben werden. Ohne Topologie ist das

geometrisch modellierte Objekt nur als **Punkthaufen** interpretierbar (s. Abb.). Erst durch die Topologie erhalten die Punkte eine Raumlogik, so dass der Zusammenhalt der Figur gewährleistet wird. Die topologischen Beziehungen zwischen den Punkten können als Graphen verstanden werden (s. dazu auch 5.1.4 Topologie).



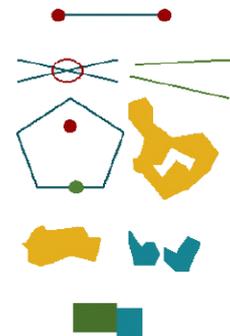
4.2.1 Topologische Invariante

Ziel der Topologie ist es, eine Raumlogik zwischen den Punkten eines Objektes herzustellen. In diesem Zusammenhang ist es vor allem wichtig topologische Invarianten zu bestimmen, so dass bei Transformationen die Metrik der Objekte sich ändern kann, aber die topologischen Eigenschaften erhalten bleiben. Bei der Festlegung der Invarianten kann man sich auf den dreidimensionalen Raum beschränken, so dass folgende Größen invariant sind:

- Geschlossenheit
- Schnittpunkttreue
- Trennung innen /außen
- Randpunkteigenschaften

Beispiele für Invarianten:

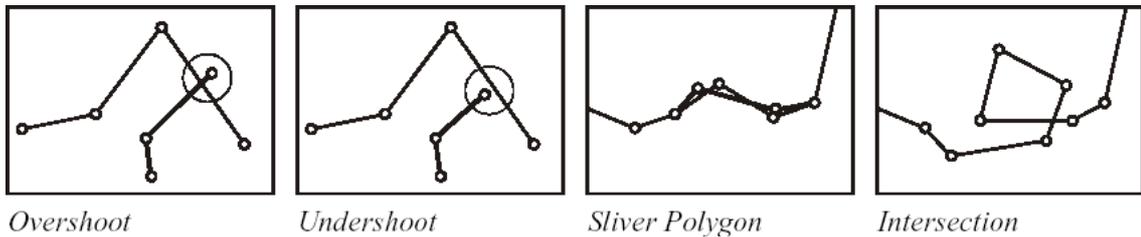
- Ein Knoten ist Endpunkt einer Kante
- Zwei Kanten kreuzen sich / sind kreuzungsfrei
- Ein Punkt liegt im Inneren einer Fläche
- Ein Punkt liegt auf dem Rand einer Fläche
- Eine Fläche hat ein Loch
- Eine Fläche ist / ist nicht zusammenhängend
- Zwei Flächen sind benachbart



Die topologischen Invarianten dienen vor allem der Abfrage topologischer Beziehungen und Konsistenzbedingungen.

4.2.2 Topologische Beziehungen und Konsistenzbedingungen

Unter topologischen Beziehungen und Konsistenzbedingungen versteht man geometrische Beziehungen, um die Daten auf Eindeutigkeit zu prüfen. Gerade bei der Datenfortführung sind die Daten solchen Prüfungen zu unterziehen, da dadurch Doppelspeicherungen vermieden, Nachbarschaftsbeziehungen auf ihre Vollständigkeit überprüft und thematische Daten auf ihre Verknüpfungen untersucht werden. Typische Konstellationen, in denen die Konsistenzbedingungen verletzt wurden, sind in folgender Abbildung dargestellt.



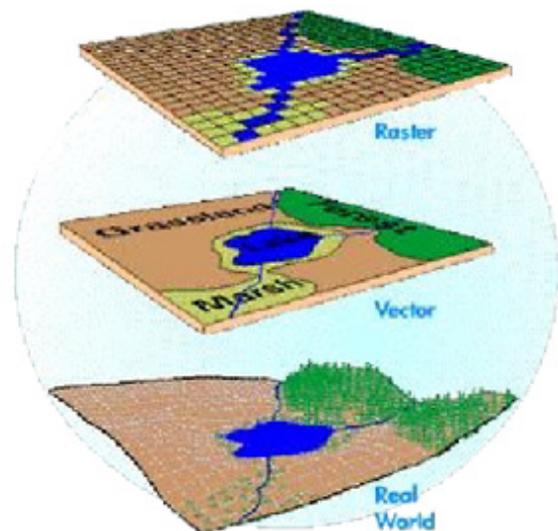
Geometrische Konstellationen bei denen die Konsistenzbedingungen verletzt wurden [Joos 1999]

4.3 Geometrische Modellierung

Die Geometrie räumlicher Objekte wird in der Regel durch die Form und die genaue Lage von Punkten, Linien und Flächen beschrieben. Dabei ist der Punkt Träger der geometrischen Information. Linien bzw. Flächen entstehen aus einer Folge von Punkten. Die Lage der Punkte wird durch Koordinaten festgelegt, wodurch das Bezugssystem und die Metrik der Geometrie feststehen.

4.3.1 Datentypen

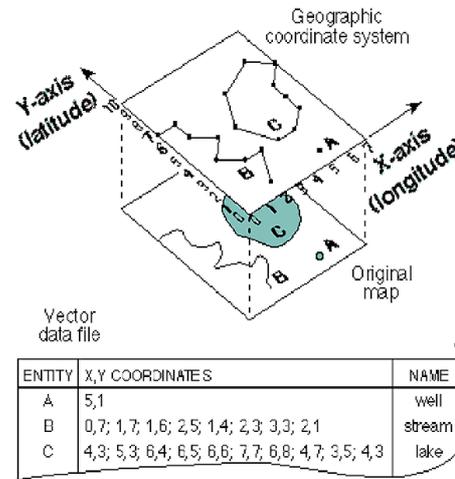
Geometrische Datenelemente können durch Vektor- oder Rasterdaten dargestellt werden. Wie in der folgenden Abbildung zu sehen ist, werden **Vektordaten** durch Punkte und Linien beschrieben, während **Rasterdaten** durch einzelne Bildpunkte (Pixel) abgebildet werden.



Beziehungen zwischen Raster- und Vektordaten [Quelle: <http://www.feut.uni-trier.de:8080/pdf/Kursbegleitung/gis1/einfuehrung-arcview.pdf>, Oktober 2003]

4.3.1.1 Vektordaten

Vektordaten nutzen als graphische Grundstrukturen Punkte und Linien, während die Fläche über einen geschlossenen Linienzug definiert wird. Sie können als Graph dargestellt werden, wobei für die geometrische Darstellung die Lage, der Knoten und die Form der Linien eine Rolle spielt. Die Daten werden nach Objektlinien geordnet. Hierdurch entsteht eine linienhafte Betrachtungsweise der Daten. Die einfachen Grundstrukturen ermöglichen die Herstellung einer logischen Datenstruktur und eines Objektbezugs. Die Strukturierung der Daten gewährleistet kleinere Datenmengen und durch schnellere Zugriffszeiten geringere Rechenleistung. Nachteil der Vektordaten ist die sehr arbeits- und zeitaufwendige punktuelle Datenerfassung.



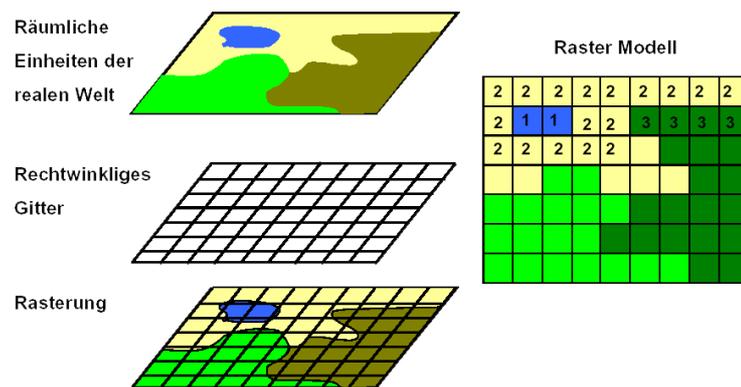
Vektordaten

[Quelle: <http://www.feut.uni-trier.de:8080/pdf/Kursbegleitung/gis1/einfuehrung-arcview.pdf>, Okt. 2003]

4.3.1.2 Rasterdaten

Rasterdaten beziehen sich nur direkt auf Flächen, wobei das geometrische Grundelement der Bildpunkt (Pixel) ist. Das Bild entsteht durch spalten- und zeilenweise angeordnete Pixel mit entsprechenden farblichen Flächenfüllungen. Durch die flächenhafte Betrachtungsweise eignen sich diese Daten eher für Anwendungen in mittleren bis kleinen Maßstäben.

Das Problem von Rasterdaten ist die Strukturierung der Daten und die Erstellung eines Objektbezugs. Dem Vorteil der einfachen Datenerfassung und den kurzen Erfassungszeiten stehen die Nachteile der großen Datenmenge und des daraus resultierenden hohen Rechenaufwands gegenüber.



Rasterdaten

[Quelle: <http://www.feut.uni-trier.de:8080/pdf/Kursbegleitung/gis1/einfuehrung-arcview.pdf>, Okt. 2003]

4.3.2 Modellierung der geometrischen Daten

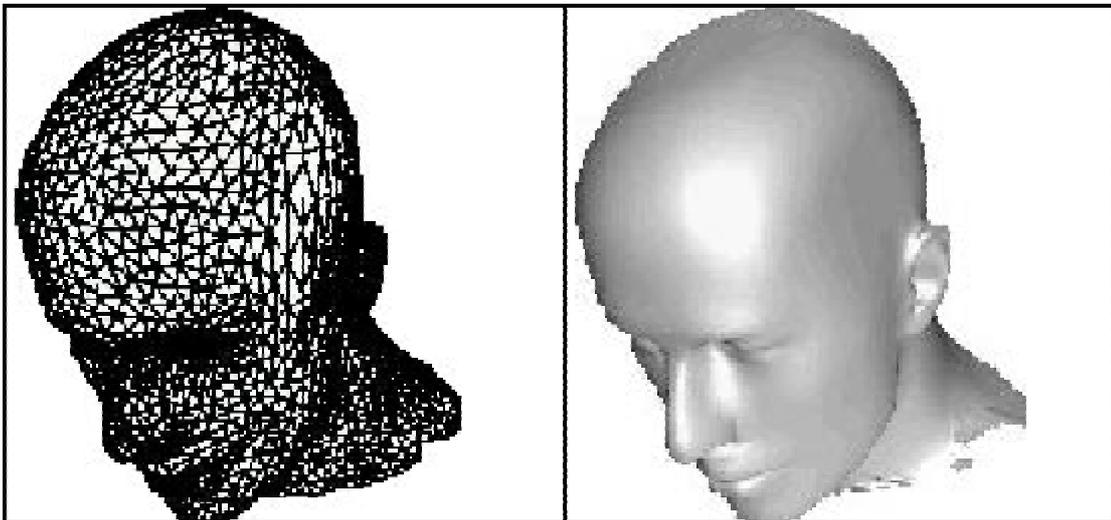
Bei der geometrischen Modellierung von Objekten unterscheidet man zwischen dem analytischen Verfahren, das auf Flächengleichungen oder Standardvolumen beruht und dem approximierenden Verfahren, das auf Interpolationen oder Approximationen infiniter Elemente aufbaut.

Im Bereich der Modellierung von Vektordaten gibt es drei Modellansätze:

- Das **Kanten-** bzw. **Drahtmodell**, bei dem Objekte durch geradlinige oder gekrümmte Verbindungen von Punkten beschrieben werden, wobei keine Informationen über Flächen und Volumina bekannt sind.
- Das **Flächen-** bzw. **Maschenmodell**, das Objekte durch analytische und approximierende Flächen beschreiben.
- Das **Volumen-** bzw. **Körpermodell**, in dem Objekte durch Primitivzellen angegeben werden. Diese können über Standardkörper, Profilkörper oder durch Boolesche Ausdrücke definiert werden.

Auf Grundlage der oben genannten Modellierungsansätze sind einige Darstellungsformen entwickelt worden. Bei Geoinformationssystemen werden vor allem die Darstellungsformen der **Randdarstellung** und das **Enumerationsverfahren** verwendet.

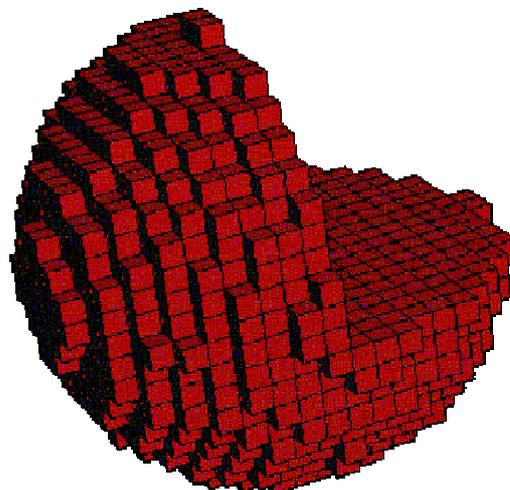
Bei der **Randdarstellung** werden die räumlichen Objekte durch ihre Begrenzungselemente, die Flächen, Linien oder Punkte sein können, beschrieben. Grundlage dieser Darstellung ist das Flächen- oder Maschenmodell.



Randdarstellung eines Kopfes

[Quelle: http://wwwipr.ira.uka.de/~megi/SEMINAR/SS_01/Modelle3D.pdf, Oktober 2003]

Das **Enumerationsverfahren** nutzt gleichförmige Raumzellen zur Definition von Objekten, wobei Grundlage das Volumen- oder Körpermodell ist.



Dreiviertel Kugel dargestellt im Enumerationsverfahren

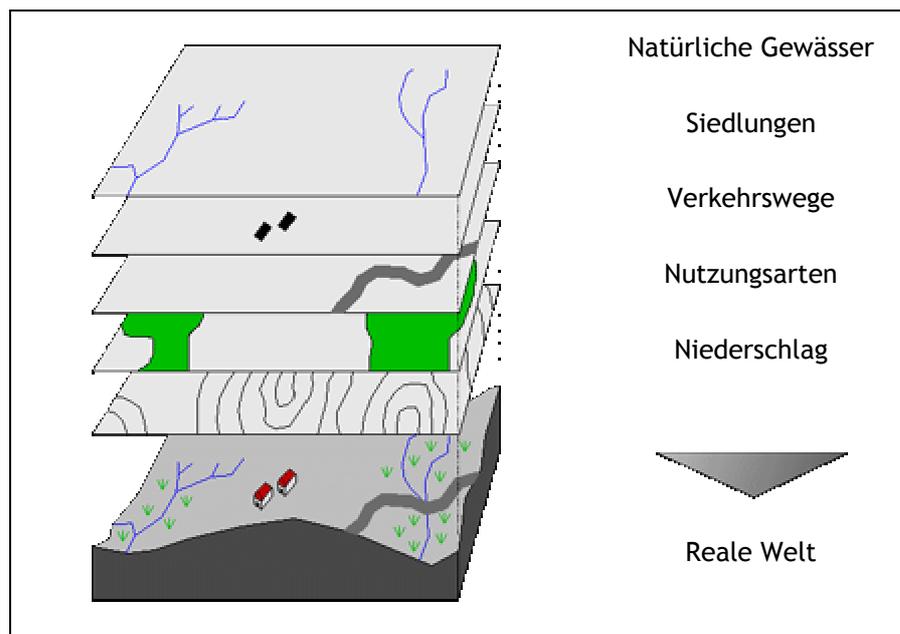
4.4 Semantische Modellierung

Die Modellierung von Daten stützt sich auf zwei Säulen. Eine dieser Säulen besteht aus den Objekten und die andere wird durch die Semantik (inhaltliche Bedeutung) der Objekte gebildet. Es ist also neben der geometrischen und topologischen Modellierung, die die örtliche Zuordnung gewährleistet, die Form berücksichtigt und das räumliche Zueinander widerspiegelt, eine semantische Modellierung der Daten erforderlich.

Im GIS-Sektor haben sich zwei Methoden zur Modellierung der Semantik besonders bewährt. Zum einen die ältere Methode des **Ebenen-** oder **Layerprinzips**, das auf den Methoden der analogen Kartenherstellung beruht. Zum anderen das **Objektklassenprinzip**, bei dem die Semantik, bezogen auf das Objekt, modelliert wird.

4.4.1 Ebenenprinzip

Beim Ebenenprinzip werden die Geometriedaten mit unterschiedlicher semantischer Bedeutung in verschiedenen Ebenen gespeichert. Die Darstellung des gesamten Kontextes erfolgt dann durch das Übereinanderlegen der einzelnen Ebenen, wie es in folgender Abbildung zu sehen ist.



Das Ebenenprinzip

[Quelle: <http://hell.wh8.tu-dresden.de/~anne/old/gis/gisallg.html>, Oktober 2003]

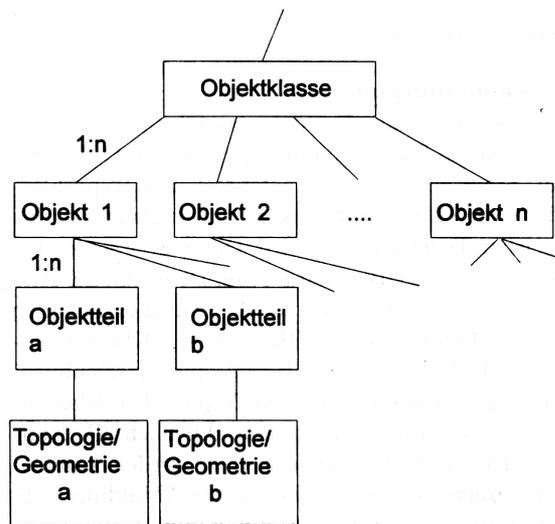
Das Ebenenprinzip ist in vielen GIS realisiert. Es besitzt keine Hierarchie, da alle Ebenen gleichberechtigt sind. Es bietet eine sehr einfache semantische Separation, was besonders dann von Vorteil ist, wenn Vektor- und Rasterdaten miteinander verknüpft werden sollen.

4.4.2 Objektklassenprinzip

Das Objektklassenprinzip ist eine Methode des semantischen Modellierens. Dabei gibt es eine Hierarchie zwischen Objektklasse, Objekt und Objektteil die nach oben offen ist. Diese Hierarchie kann streng eingehalten werden, was zu einem **semantischen Baum** führt oder es werden netzwerkartige Verknüpfungen zugelassen, die zu einem **semantischen Netzwerk** führen.

4.4.2.1 Semantischer Baum

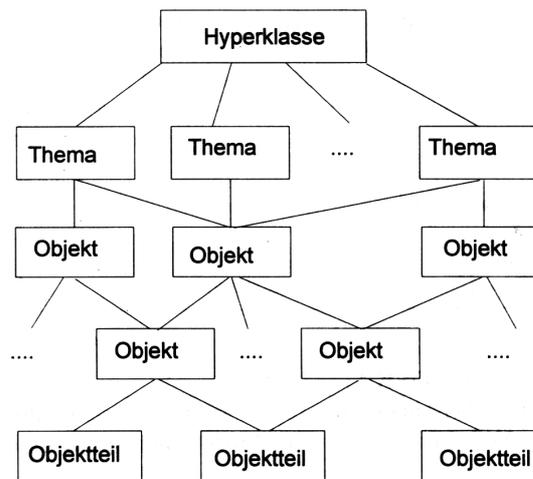
Geht man bei dem Objektklassenprinzip streng hierarchisch vor, so führt dies zu einem semantischen Baum. Innerhalb des Prinzips ergeben sich 1:m-Beziehungen zwischen den einzelnen semantischen Mengen, d.h. eine Objektklasse verzweigt sich in m Individualobjekte.



Semantischer Baum [Quelle: Bill 1999a]

4.4.2.2 Semantisches Netzwerk

Ein semantisches Netzwerk ist dadurch gekennzeichnet, dass alle Wege offen sind, d.h. zwischen den einzelnen semantischen Mengen sind n:m-Beziehungen möglich. Es gibt keine strenge Hierarchie mehr, so dass Objekte zugelassen werden, die auch untereinander Verbindungen haben können.



Semantisches Netzwerk [Quelle: Bill 1999a]

4.5 AFIS-ALKIS-ATKIS-Datenmodellierung

Die Vermessungs- und Katasterverwaltungen der Bundesländer sind für die Bereitstellung von raumbezogenen Basisdaten (Geobasisdaten) für Verwaltung, Wirtschaft und private Nutzer verantwortlich. Für die digitale Erfassung der Daten wurden schon in den 70er und 80er Jahren Konzepte entwickelt, nach denen ALB, ALK und ATKIS aufgebaut werden sollten. Diese Konzepte dienen heute immer noch als Plattform für den Aufbau der entsprechenden Geobasisdatenbestände. In den vergangenen Jahren hat die AdV zunehmend an einem Konzept gearbeitet, um die liegenschaftsbeschreibenden Geoinformationssysteme ALB und ALK in dem System ALKIS zu integrieren.

Zu den Geoinformationssystemen des amtlichen Vermessungswesens gehören auch die Informationen zu den Festpunkten, die aber weder originär zu ATKIS noch zu ALKIS gehören. Deshalb wurde im Mai 2000 deren Modellierung in einem eigenen Informationssystem dem AFIS beschlossen.

Die amtlichen Geobasisdaten lassen sich dementsprechend in drei Produktbereiche einteilen:

- Positionsbestimmende Geobasisdaten (AFIS)
- Liegenschaftsbeschreibende Geobasisdaten (ALKIS)
- Landschaftsbeschreibende Geobasisdaten (ATKIS)

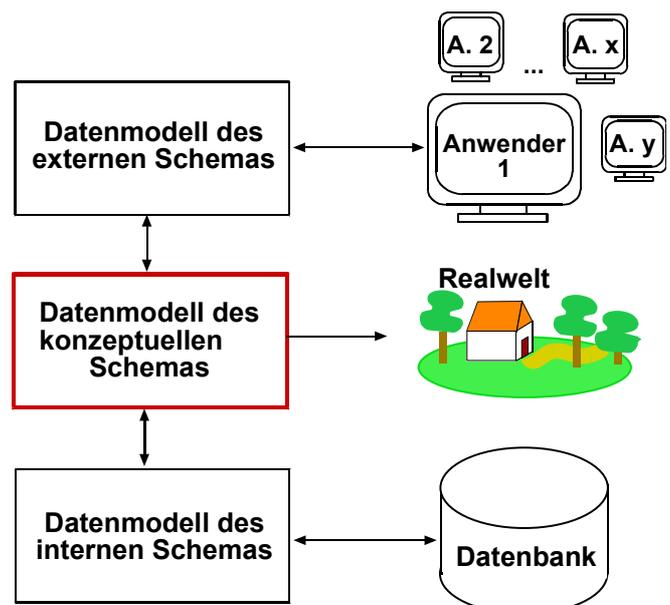
4.5.1 Datenmodelle in Informationssystemen

Der Zusammenhang zwischen der Realität, deren Abbildung in der Datenbank und der Sicht des Anwenders auf die Daten, kann durch die 3-Schema-Datenmodell-Architektur beschrieben werden (s. dazu auch 3.3 Komponententechnologie).

Das Datenmodell des **konzeptuellen Schemas** legt die Voraussetzungen für eine einheitliche Erfassung der Realweltobjekte aus einer bestimmten fachlichen Sicht fest. Die genauen Angaben beispielsweise welche Realweltobjekte, mit ihren selbstbezogenen Eigenschaften (Attribute) und fremdbezogenen Eigenschaften (Relationen), in einem GIS abgebildet werden, sind in einem Objektartenkatalog detailliert dokumentiert.

Das Datenmodell des **internen Schemas** beinhaltet die Umsetzung, der durch das konzeptuelle Schema erfassten Daten, in die logische Struktur der Datenbank.

Das Datenmodell des **externen Schemas** definiert die Datenaustauschnittstellen.



3-Schema-Architektur

4.5.2 Das konzeptuelle Datenmodell

Das konzeptuelle Datenmodell dient der abstrahierten Erfassung der Realwelt. Der Zweck eines solchen Modells ist es, ein gemeinsames und einheitliches Verständnis der Daten zu erreichen und die Dateninhalte für eine bestimmte fachliche Anwendungsumgebung so zu dokumentieren, dass eindeutige Informationen über die Daten erhalten werden.

Die AdV hat für das AFIS-ALKIS-ATKIS-Modell einen einheitlichen und objektorientierten Ansatz gewählt, der möglichst mit den marktüblichen und dem Stand der Technik entsprechenden GIS-Programmen abgebildet und geführt werden soll.

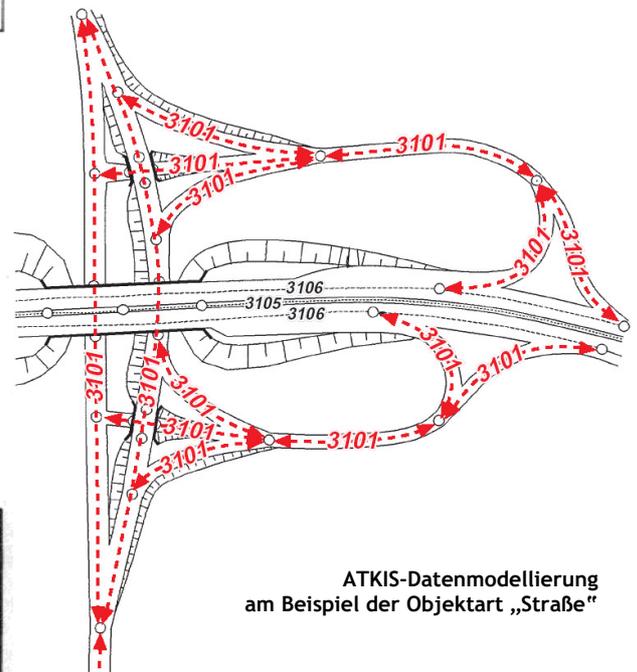
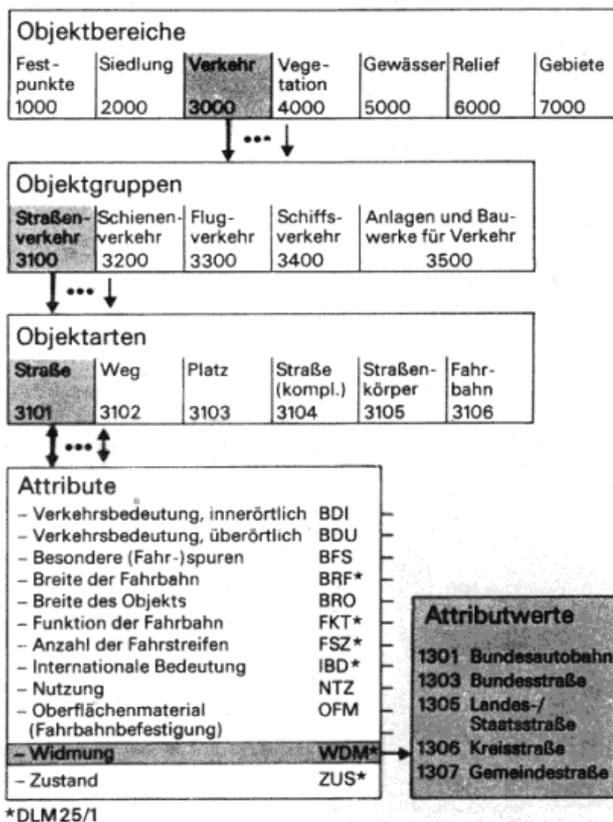
Das konzeptuelle Modell gliedert sich in:

- Basisschema
- Versionsschema
- AFIS-ALKIS-ATKIS-Fachschemata

Beispiel:

Die folgenden Abbildungen zeigen die Datenmodellierung am Beispiel von Realweltobjekten der Objektart „Straße“.

Aufbau des Objektartenkataloges



ATKIS-Datenmodellierung am Beispiel der Objektart „Straße“

ATKIS Objektartenkatalog [Quelle: <http://www.laum.uni-hannover.de/ilr/lehre/lehre.html>, Oktober 2003]

Literaturverzeichnis

Bartelme 2000 Bartelme, N.: **Geoinformatik - Modelle, Strukturen, Funktionen**. Springer-Verlag, Berlin, Heidelberg, New York, 2000 [ISBN 3-540-65988-9]

Bill 1999a Bill, R.: **Grundlagen der Geo-Informationssysteme - Band 1 Hardware, Software und Daten**. Wichmann Verlag, Heidelberg, 1999 [ISBN 3-87907-325-2]

Bill 1999b Bill, R.: **Grundlagen der Geo-Informationssysteme - Band 2 Analyse, Anwendungen und neue Entwicklungen**. Wichmann Verlag, Heidelberg, 1999 [ISBN 3-87907-326-0]

GeoInfoDok 2003 Arbeitsgemeinschaft der Vermessungsverwaltungen (AdV): **Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens**, 2003, online verfügbar (Juni 2003): <http://www.adv-online.de/>

Joos 1999 Joos, G.: **Zur Qualität von objektstrukturierten Geodaten**. Schriftenreihe des Studienganges Geodäsie und Geoinformation der Universität Bundeswehr München, Heft 6/2000, München, 1999

Schilcher 2003 Schilcher, M.: **Skript zur Vorlesung Geoinformatik I der TU München, Kapitel 4: Modellierung und Datenmodelle in GIS**, München, 2003, online verfügbar (Juni 2003): <http://www.gis1.bv.tum.de/Lehre/Vorlesungen/Geoinformatik1/Kapitel4/Dokumente/SkriptGeo1%20Kapitel4.pdf>