

# Modellarten



Ein Dokument der Serie "Model-driven Method Memos" (MM-Serie)

Prof. Stefan F. Keller, [sfkeller@hsr.ch](mailto:sfkeller@hsr.ch)  
Center int>e>gis, Institut ITA, Hochschule Rapperswil

Version 0.9de vom 2003-05-02 (\*\* ENTWURF)

Dokument 'MM\_Modellarten.doc'

Aktuellste Version jeweils auf [www.integis.ch](http://www.integis.ch)

## Inhaltsüberblick

<b>1</b>	<b>Allgemeines.....</b>	<b>2</b>
1.1	Einleitung.....	2
1.2	Der "Geltungsbereich" genormter Datenmodelle.....	2
1.3	Warum konzeptionelle Datenmodelle? Die 3-Schema-Architektur.....	3
<b>2</b>	<b>Konzeptionelle Datenmodelle.....</b>	<b>5</b>
2.1	Allgemeines.....	5
2.2	Ziel und Zweck der Modellierung und des Transfers.....	5
2.3	Modellarten.....	5
2.4	Modelle nach den Prozessen der Geo-Informationverarbeitung.....	6
2.5	Modell- und Transferarten im Detail.....	7
<b>3</b>	<b>Typische Modelle und Beispiele.....</b>	<b>9</b>
	<b>Quellenangaben .....</b>	<b>10</b>

# 1 Allgemeines

## 1.1 Einleitung

Dieses Dokument beschreibt die verschiedenen Modell-Arten, die sich je nach Architektur,

## 1.2 Der "Geltungsbereich" genormter Datenmodelle

Oft stellt sich bei der Modellierung genormter (meist konzeptioneller) Datenmodelle die Frage, ob ein umfassendes, globales Modell (Gesamtmodell) mit Auswahlmöglichkeit oder ein "Kernmodell" erstellt werden soll, das dann ja lokal oder später immer noch erweitert werden kann.

Beide Varianten haben ihre Vor- und Nachteile. Hier ein Auszug der jeweiligen Vorteile:

- Gesamtmodell (Modell mit Auswahlmöglichkeit: Teilmodelle):
  - grössere Abdeckung des Themas; es können mehr Daten gemeinsam genutzt werden
  - Einmal etabliert, ist die Nachführung oft einfacher; gibt es komplexe Datengrundlagen ist eine effiziente Nachführung überhaupt erst möglich
  - vorwiegend eine Art Top-Down-Vorgehen: Man hat oder erhält eine abschliessende Übersicht über das Gebiet
  - (lokale) Erweiterungen sind immer noch möglich
  - ein Gesamtmodell kann man definieren, als eines, das in die "Tiefe" geht, dass also alle Informationen eines Themas von den Basisdaten, bis zum verarbeiteten Endprodukt abdeckt oder eines, das in die "Breite", das also alle mit dem Kernthema verwandten Informationen abdeckt (z.B. das Datenmodell Abwasser, das alle Abwasser-nahen Bereiche, wie Regenwasser auch mit-modelliert, anstelle nur die Kanäle).
- Kernmodell:
  - kleinere Anzahl Regelungen; weniger aufwändig
  - geeignet für komplexe, (ursprünglich) unbekannte, unklare oder nicht einheitliche Themen. Zusätzliche Modellierungen sind immer noch möglich (bis man schliesslich über die Jahre hinweg beim Gesamtmodell landet)
  - Beteiligung von mehr Experten möglich (kürzere Konsensprozesse)
  - vorwiegend eine Art Bottom-up-Vorgehen: Man beginnt dort, wo es am meisten Sinn macht und wo der Bedarf am Grössten ist
  - Daten, die nicht von zentralem Interesse sind, können mittels Erweiterungen abgedeckt werden

Die Wahl für das eine oder andere Modell hängt von diesen Faktoren ab.

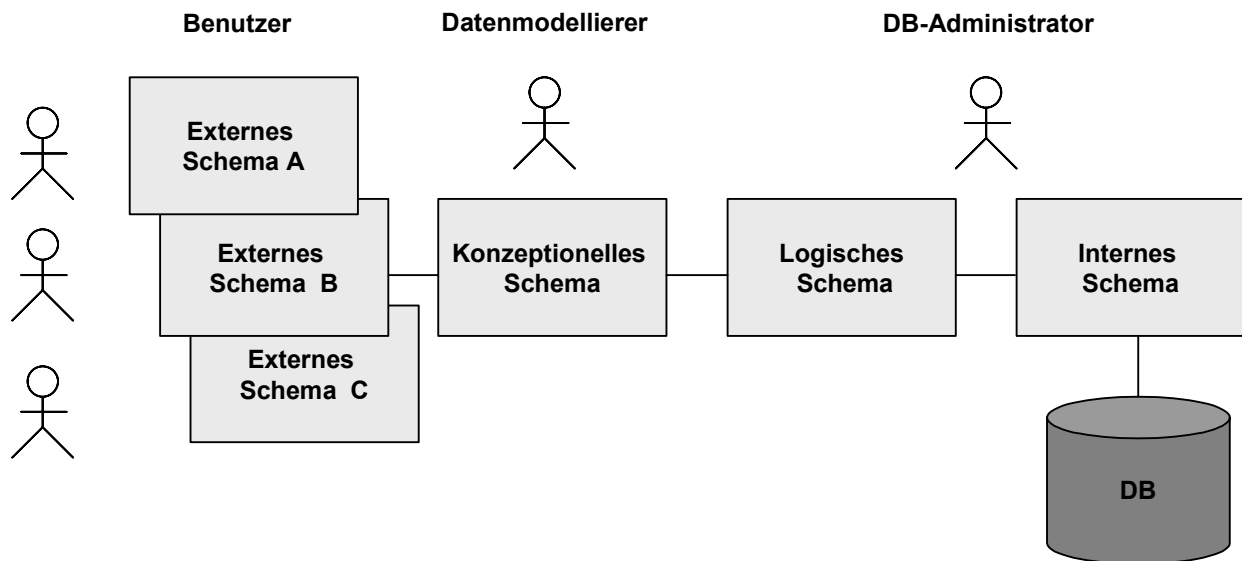


Fig. 1. 3-Schema-Architektur nach ANSI-SPARC

### 1.3 Warum konzeptionelle Datenmodelle? Die 3-Schema-Architektur

Unter dem Namen "3-Schema-Architektur" (oder -Konzept oder -"Modell") wird in der Datenbank-Technologie ein wichtiges Strukturierungsprinzip beschrieben: Während das interne Schema die physische Gruppierung der Daten und die Speicherplatzbelegung beschreibt (also technologie- und produkte-orientiert ist), gibt das konzeptionelle Schema die sach-logischen Zusammenhänge der Daten wieder - es ist also systemneutral und benutzer-orientiert. Mehrere externe Schemata enthalten bestimmte Teilbereiche des konzeptionellen Schemas, die auf die jeweiligen Anwenderanforderungen zugeschnitten sind.

Je nach Sichtweise verstehen wir unter Daten-Beschreibung also etwas anderes: Wir können die Daten aus der Sicht der Speicherungsstrukturen betrachten, dies ist die interne Ebene. Die logisch-konzeptionelle Ebene definiert die Datenelemente und ihre Strukturen. Die externen Sichten beschreiben, wie die verschiedenen Benutzer, die sie interessierende Teilmenge der Daten sehen.

Das 3-Ebenen-Architektur wurde als Strukturierungsmittel in der DB-Welt eingeführt um die Sicht der Benutzer und der Applikationen auf die Daten von der Art der internen Organisation und Speicherung zu separieren. Einige Motive für diese Trennung sind:

- Verschiedene Benutzer haben verschiedene Sichten derselben Datenmenge. Die Art, wie ein Benutzer die Daten sieht, kann sich im Laufe der Zeit ändern.
- Der Benutzer soll von den Problemen der Datenbank-Strukturierung abgeschirmt werden
- Der DB-Administrator sollte die logischen Struktur der Datenbank reorganisieren können, ohne dass die Benutzer davon tangiert werden.
- Der DB-Administrator sollte die Speicherungsstruktur ändern können, ohne dass die logische Struktur davon beeinflusst wird.

Das ANSI-SPARC-Komitee definierte die 3-Schema-Architektur für die Beschreibung der einzelnen Ebenen:

- Die externen Benutzersichten werden durch die externen Schemata definiert. Eine externe Benutzersicht umfasst eine Teilmenge des konzeptionelle Schemas und ist auf die Bedürfnisse der jeweiligen Benutzerklasse abgestimmt.
- Das konzeptionelle Schema spezifiziert die sachliche-logische Ebene und ist Grundlage für den

Entwurf der übrigen Schemata (Schemas). Das konzeptionelle Schema beschreibt die Struktur der gespeicherten Daten.

- Das logische Schema wird je nach Begriffsdefinition mit konzeptionellem Schema gleichgesetzt. Häufiger aber kann es auch als Umsetzung des konzeptionellen in das interne Schema angesehen werden, was mittels einer konkreten Methode (relational) und einer produktspezifischen-konkreten Sprache geschieht (z.B. Oracle-SQL).
- Das interne oder physische Schema beschreibt die Strukturen der physischen Datenorganisation.

Das Prinzip der Datenunabhängigkeit: Die Separation zwischen den einzelnen Schichten erlaubt also, dass eine Schicht (in einem beschränkten Rahmen) reorganisiert werden kann, ohne dass die darüberliegenden Schichten angepasst werden müssen. Diese bedingt aber komplexe Transformationen an den Schnittstellen zwischen den einzelnen Schichten.

## 2 Konzeptionelle Datenmodelle

### 2.1 Allgemeines

Für ein bestimmtes Fachthema gibt es aber nicht nur *ein* konzeptionelles Datenmodell. Es gibt verschiedenste Arten von Datenmodellen, z.B. je nach Zweck oder Verarbeitungsphase. Je nach dem, werden unterschiedliche Datenbeschreibungen notwendig sein - sogar für denselben Zweck und dieselbe Verarbeitungsphase gibt es Varianten. Mit INTERLIS können alle diese Beschreibungen auf den Zweck ausgerichtet und präzise formuliert werden.

Der Zweck von INTERLIS ist die effiziente und gemeinsame Nutzung von Geodaten, sowie deren langfristige Sicherstellung einschliesslich deren Dokumentation.

### 2.2 Ziel und Zweck der Modellierung und des Transfers

Generell gemäss Bartelme (1995, S.349) können die Funktionen von Geoinformationssystemen wie folgt eingeteilt werden (ergänzt durch den ersten Punkt):

1. Modellierung und Dokumentation von Geoinformation
2. Geometriedatenerfassung, Prüfung und Korrektur
3. Sachattributdatenerfassung, Prüfung und Korrektur
4. Strukturierung zu Objekten inkl. geometrische Topologie
5. Transformation: Koordinatentransformation, Randabgleich, Generalisierung, Datentransfer und Datenschnittstellen
6. Berechnung und Konstruktion
7. Abfrage und Analyse
8. Präsentation

INTERLIS soll hauptsächlich (muss aber nicht) für die Modellierung und Dokumentation von Geoinformation sowie der Transfer von Daten zusammen mit der Spezifikation und dem Transfer von Präsentationsdaten vereinheitlichen.

### 2.3 Modellarten

In Abhängigkeit des Verwendungszweckes der ausgetauschten Daten, können folgende Hauptarten von Modellen und damit auch des Transfers oder Austausches unterschieden werden (nach RAV-Dokument, V+D/J. Kaufmann 1991?):

#### 2.3.1 Bearbeitungs-Modell zum Transfer, bzw. Austausch zwischen verteilten Systemen

Darunter verstehen wir den Austausch zwischen Bearbeitern des Fachthemas, z.B. zwischen externen Datenerfassern und/oder Amtsstellen. Der Datentransfer umfasst in diesem Falle sämtliche Daten, inklusive diejenigen Daten, welche fach-'internen' Charakter haben. Ein Bearbeiter muss über vollständige Information verfügen, um seine Aufgabe lösen zu können; er muss die Daten später evtl. nachvollziehen oder nachführen können.

### 2.3.2 Verifikationstransfer-Modell

Datenaustausch zwischen Daten-Erfasser oder Bearbeiter und Aufsichts- oder Auftragsstellen zum Zwecke der Verifikation. Bei dieser Kategorie interessieren vor allem die 'internen' Daten zur Überprüfung, ob Anforderungen eingehalten wurden. Die eigentlichen Daten müssen hier nur stichprobenweise zur Verfügung stehen.

### 2.3.3 Nutzer-Transfer/-Abgabe-Modell

Den Benutzer interessieren vor allem diejenigen Daten die er für die Lösung seiner Aufgaben benötigt. Das können sämtliche verfügbaren Daten sein, nur ein Teil davon (räuml. Ausschnitt oder nur einige Attribute) oder es sind aufbereitete Daten. Er will auf - aus seiner Sicht - überflüssige 'interne' Informationen verzichten, die nur sein System belasten würde, verzichten.

## 2.4 Modelle nach den Prozessen der Geo-Informationverarbeitung

Die Grundfunktionen eines GIS und damit einer Geo-Informationverarbeitung können wie folgt unterschieden werden: Erfassung (meist beim Auftragnehmer), Integration, d.h. Erfassungs- und Ergänzungsphase (z.B. beim Auftraggeber/Veredler), Verwaltung, Analyse, Auswertung inkl. grafische Darstellung in Pläne/Karten und Datenabgabe. Diesen Prozessen können entsprechenden Datenmodell-Arten zugeordnet werden (beim DM.01-AV z.B. handelt es sich mehrheitlich um ein Verwaltungsmodell, ausgenommen z.B. Fixpunkte).

- Datenerfassungs-Modell
- Datenverwaltungsmodell oder Datenaustausch-/Datentransfer-Modell
- Datenanalyse-, Auswertungsmodell oder grafisches Darstellungs-Modell
- Datenabgabe-Modell

### 2.4.1 Erfassungs-Modell/-Schema

Das Erfassungsmodell ist typischerweise projektspezifisch. Für eine effiziente und allenfalls hoch-aktuelle Nachführung ist das Datenerfassungs-Schema z.T. nicht geeignet: Ersterfassung und Nachführung unterscheiden sich, auch wenn es sich um dasselbe System handelt, da im Nachführungsfall ja gewisse Daten vorhanden sind, die bei der Erfassung noch nicht vorhanden gewesen sind.

### 2.4.2 Datenaustausch-Modell/-Schema

Das Datenaustausch-Schema ermöglicht die Bearbeitung von Originärdaten; eine Nachführung wird für den Empfänger nachvollziehbar. Solche Daten erlauben u.a. auch eine Migration ohne Informationsverlust zwischen den Systemen.

### 2.4.3 Datenabgabe-Modell/-Schema

Das Datenabgabe-Schema enthält z.T. Datensichten, d.h. abgeleitete/aggregierte Informationen. Es geht davon aus, dass die Daten Empfänger-gerecht aufgearbeitet sind. Der Empfänger kann (und soll) gewisse Originalinformationen nicht mehr eruieren (Beispiel: als Gebäude bezeichnete Objekte, die im Original in öffentlich/privat unterschieden wurden oder ein statistisches Mittel geben keinen Aufschluss mehr auf die darunterliegenden Daten). Im Empfängersystem kann aufgrund solcher Daten keine Nachführung der Originaldaten erfolgen. Da die Daten in aufbereiteter Form geliefert werden, ist es für das Empfängersystem u.U. einfacher, diese weiterzuverwenden. Müssen die Daten jedoch angepasst werden, so kann das nur mehr bedingt erfolgen (das geht nur mit Informationsverlust oder Verfälschung).

Vgl. auch "Nutzer-Transfer / -Abgabe".

## 2.5 Modell- und Transferarten im Detail

Primär ist Interlis ein Mechanismus für die Modellierung und den Transfer von Daten, insbesondere Geodaten. Allerdings sind recht unterschiedliche Anwendungen denkbar, die jeweils verschiedene Anforderungen nach sich ziehen.

Alle diese Datenmodelle als „Vertrag zwischen Sender und Empfänger“ bzw. als gemeinsam genutzte Information basieren also auf dem Datentransfermodell (bzw. gemeinsames konzeptionelles Datenmodell), das eine konzeptionelle Sicht auf das systemspezifische Datenverwaltungsmodell (internes Datenmodell) darstellt. Sie sind teilweise voneinander abhängig:

- Datentransfer-Modell
- Datenerfassungs-Modell
- Konfigurations-Modell
- Verifikations-Modell
- (Graf.) Darstellungs-Modell
- Analyse-Modell
- Dokumentations-Modell

Die Datenbank-Dokumentation im weiteren Sinne umfasst alle diese Modellarten und dient allen GIS-Funktionskategorien. Man erkennt in diesen Modellen die Grundfunktionen von GIS wieder: Erfassung und Integration, Verwaltung, Darstellung, Analyse und Abgabe (vgl. unten).

Zur Erläuterung dieser Kategorien von Modellen hier ein modifizierter Auszug aus Bächli, Brawer und Dorfschmid 1998):

- *Systemunabhängiger Datentransfer zwischen verteilten Systemen:* INTERLIS wird als eigentlicher Transfermechanismus zum Austausch von Geodaten zwischen Systemen verschiedener Anbieter verwendet. Die Anwender erwarten, dass die übertragene Information möglichst vollständig ist. Auf dieser Stufe sind beispielsweise Konsistenzregeln und Defaultwerte unerheblich. Wie bereits oft darauf hingewiesen, besitzen z.B. optionale Attribute die Bedeutung, dass der Attributwert des modellierten Objekts der realen Welt manchmal auch fehlen kann und nicht etwa, dass dieser nicht zu transferieren sei, obschon er vorhanden ist.
- *Systemunabhängiger Datentransfer für Client-Server-Anwendungen:* Mit den für INTERLIS Version 2 vorgeschlagenen Erweiterungen (Objektstrom?) werden Anwendungen wie INTERLIS-basierte GIS-Server denkbar. Der systemunabhängige Datentransfer zwischen verteilten Systemen und Client-Server-Anwendungen benutzen INTERLIS rein zum Datenaustausch; sie sollten sich daher primär nicht mit grafischen Darstellungen u.ä. befassen müssen ausser es handelt sich dabei (evtl. zusätzlich) um solche Server, die zu bestehenden Datenmodellen ergänzende Modelle und Daten für die grafische Präsentation anbieten.
- *Datenbank-Konfiguration (bzw. -Parametrisierung):* Denkbar ist auch, ein GIS anhand einer INTERLIS-Beschreibungsdatei zu konfigurieren, z.B. für die Datenerfassung. Beispielsweise könnte ein System Angaben für Defaultwerte - die momentan nicht Teil von INTERLIS sind, jedoch oft als Kommentare angegeben werden - direkt aus INTERLIS übernehmen.
- *Datenmodell-Detailspezifikation und Datenbeurteilung:* INTERLIS dient auch der Kommunikation zwischen Datenbesteller und -lieferanten, die in INTERLIS ihre Ansprüche hinsichtlich Inhalt und Qualität raumbezogener Daten in formalisierter Form festhalten können. Somit ist INTERLIS eine weitergehende Datenmodellierungssprache. Zur Detailspezifikation können auch 'harte', d.h. vollautomatisch prüfbare Konsistenzangaben gezählt werden.
- *Datenmodell-Dokumentation:* Es lässt sich jedoch nicht alles mit vernünftigem Aufwand formalisieren. Jene Teile von INTERLIS, die ausschliesslich zu Dokumentationszwecken dienen, müssen nicht unbedingt maschinell verarbeitet werden können. Hier gibt es 'weiche', d.h. nicht

vollautomatisch überprüfbare Konsistenzbedingungen und sogenannte Plausibilitätstests. INTERLIS kennt aus diesem Grund das Konstrukt der Erläuterungen, und nochmals weitergehende Vereinbarungen zwischen Sender und Empfänger. Diese werden wohl auch in Zukunft ausserhalb von INTERLIS beschrieben. Eine solche Dokumentation im weiteren Sinne kann auch als eine Art Grobspezifikation betrachtet werden, während alle vorangegangenen Kategorien selbstverständlich ebenfalls Teil einer vollständigen Dokumentation sind.

Diese können in einer tabellarischen Übersicht charakterisiert und den obengenannten Modellarten zugeordnet werden:

<i>Applikation</i>	<i>Funktion</i>	<i>Defaultwerte</i>	<i>Konsistenzan- gaben</i>	<i>Views</i>	<i>Codierte Daten</i>	<i>Datenmodell und Sprache</i>
Datentransfer zwischen verteilten Systemen und Client-Server	Integration/ Verwaltung, Abgabe	nein	nein, evtl. als Dokumentation	ja, Empfänger als interne Tabelle	ja	Datentransfer-Modell
Datenerfassung	Erfassung, Integration	ja	ja, harte	nein	nein	Datenerfassung-Modell
Datenbank-Konfiguration	Verwaltung	ja	ja, harte	ja	nein	Konfigurations-Modell
Datenbeurteilung (Verifikation)	Verwaltung, Integration	nein	ja, harte und weiche	ja	nein (evtl. Metadaten, wie Qualität etc.)	Verifikations-Modell
Graf. Darstellung (Karten, Pläne, Bildschirmdarstellung und weitere)	Verwaltung, Abgabe	nein (Konstanten!)	ja, harte für Query	ja, für Query	nein (evtl. Signaturen)	(Graf.) Darstellungs-Modell
Daten-Analyse	Analyse	nein (Konstanten!)	ja, harte	ja	nein	Analyse-Modell
Daten-Dokumentation im weiteren Sinne	Erfassung, Integration/ Verwaltung, Abgabe	ja, aus Erfassungs-M.	ja, evtl. harte und weiche	ja	nein	Dokumentation-Modell, alle Modellarten

Tabelle 1: Modellarten.

### 3 Typische Modelle und Beispiele

Mit "konzeptionellem Datenaustausch-Schema" bezeichnet man ein konzeptionelles Schema/Modell, das vor allem denjenigen Teilbereich der Informationen abdeckt, der eine "Datentransfer-Gesellschaft" im Hinblick auf einen Austausch und gemeinsame Nutzung und Nachführung dieser Daten interessiert. Im Rahmen der Datenmodellierung ist dies wohl die häufigste Modellart.

Ein besonderes internes Datenschema ist das "physische Datenschema", welches ein sequentielles Datentransferformat bestimmt, wie z.B. das INTERLIS-Transferformat. Dieses ist systemneutral definiert - also z.B. nicht auf Performance oder Speicherplatzeffizienz, sondern auf den informationsverlustfreien Datentransfer ausgerichtet.

Ein typisch anzutreffender Fehler ist, wenn das interne Schema als das konzeptionelle, gemeinsame Datenaustausch-Modell betrachtet wird.

## Quellenangaben

Siehe auch [www.integis.ch](http://www.integis.ch) > Publikationen.

[INTERLIS1] Eidg. Vermessungsdirektion (1991): INTERLIS - Ein Datenaustauschmechanismus für Land-Informationen-Systeme.

[INTERLIS2] <http://www.interlis.ch>.