# Ein Anfang: Warum, Wie und Womit

Dr.-Ing. Thomas Schwotzer

4. April 2014

**Zusammenfassung**

Wir beschäftigen uns in diesem Modul mit Ortsbasierten Informationssystemen. Ziel ist es, die notwendige Theorie, die wichtigsten Standards und einige praktische Fähigkeiten zu entwickeln. Am Anfang steht eine Einführung und Wiederholung.

## 1 Warum?

Warum beschäftigen wir uns im Schwerpunkt mobile Anwendungen mit Geoinformationssystemen? Die Antwort ist schnell gefunden: Es gibt eine enorme Menge mobiler Anwendungen, die ortsbasiert arbeiten. Die Anzahl dieser Anwendungen wächst enorm. Stichworte: Restaurantfinder, Carsharing, Parkplatzsuche, Geo-Caching etc. pp.

Die typische Architektur einer solche Anwendung ist im Abbildung 1 dargestellt.

Anwender bekommen eine Karte zu Gesicht auf der weiter Informationen vermerkt sind, das kann der Tipp sein wo es die beste Nudelsuppe der Gegend gibt oder wo das nächste Car-Sharing Fahrzeug zu finden ist.

Tatsächlich sind in aller Regel wenigstens zwei Informationsquellen an einem solche System beteiligt: Eines liefert eine Karte, das andere weitere – anwendungsspezifische – Informationen. Die Karte ist in den meisten Anwendungen notwendig. Die zusätzliche Informationen können in vielen Anwendungen sogar ein- und ausgeblendet werden.

Im Rahmen dieser Veranstaltung werden wir uns mit all diese Themen beschäftigen. Themen werden sein:

- Woher bekommt man Kartenmaterial? Wie genau ist das und welche rechtlichen Rahmenbedingungen muss man beachten?
- Wie kann man eigene Karten erstellen? Wie entstanden existierende Karten? Wie genau sind die Karten?
- Wie speichert man die zusätzlichen Informationen?
- Wie kann man mit diesen Informationen rechnen? Wie kann man Routen ermitteln? Wie kann man Abstände errechnen?

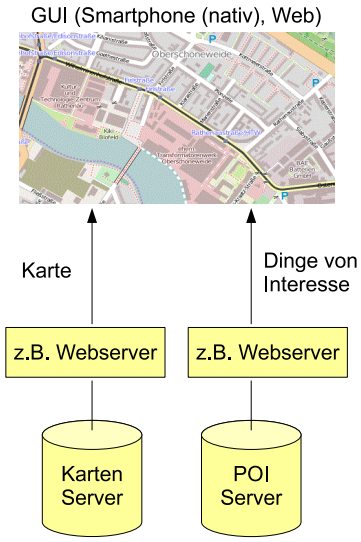


Abbildung 1: Typische Architektur ortsbasierter Systeme

Wir werden uns mit allen Teilen dieser üblichen Architektur im Rahmen dieses Moduls beschäftigen. Für die GUI-Entwicklung werden wir uns Lösungen für das Web anschauen, konkret OpenLayers. Auf Android werden wir die OSM-Variante einführen. Beides ist nicht wirklich schwer.

Die meiste Zeit werden wir aber in die Serverseite investieren, weil dort die Algorithmen und Datenstrukturen realisiert werden. Konkret werden wir mit PostGIS arbeiten, um die ortsbasierten Informationen zu speichern. Wir werden uns damit beschäftigen, wie diese Daten über einen Webserver zur Verfügung gestellt werden.

Wir werden uns auch damit beschäftigen, wie Kartenmaterial aus Datenbanken generiert werden. Und wir werden und die Grundlagen der Kartenerzeugung anschauen müssen. Diese Grundlagen benötigen wir, um mit ortsbasierten Daten korrekt rechnen zu können. Sie werden die üblichen Kartenstandards UTM und Gauß-Krüger kennen lernen und auch die üblichen Austauschformate von Datenbanken, z.B. Shapefiles.

Die Datenbanken, die in diesem Umfeld genutzt werden, sind GIS-Datenbanken. Und mit GIS geht es weiter.

# 2 GIS

Geoinformationssystemen (GIS) sind Systeme, die der Speicherung und der Arbeit mit Geodaten helfen. GIS sind Systemen, d.h. Anwendungen. Technische Basis von GIS sind oft Datenbanken, die in aller Regel über besondere Unterstützung für Geodaten verfügen. Wir beschäftigen uns vor allem mit Post-

GIS. Es gibt aber auch eine Fülle von Programmier-Frameworks, die die Arbeit mit Geodaten erleichtern.

Hier seien nur exemplarisch einige URLs aufgelistet.

- www.geotools.org
- maptools.org
- mapquery.org
- openrouteservice.org

## 2.1 Open Geospatial Consortium Inc. (OGC)

Das OGC ist ein Konsortium aus GIS-Herstellern bzw. Nutzern. Das OGC beschäftigt sich mit der Standardisierung, dessen Wichtigkeit hoffentlich nicht vertieft werden muss. Hier lediglich ein paar wichtige Links.

- OGC Webseite: http://www.opengeospatial.org/
- Die Liste der Standards findet man zwei Klicks weiter:
  http://www.opengeospatial.org/standards/is

Wir werden uns vor allem mit den Simple Feature Model explizit beschäftigen. Andere OGC Standards werden wir implizit nutzen. Doch zunächst müssen wir ein wenig theoretisieren, was ein GIS überhaupt speichert, nämlich Geobjekte.

## 2.2 Das Geoobjekt

Jedes GIS beschreibt Objekte der realen Welt und erlaubt weitere Informationen mit ihm zu assoziieren. Unter einem *Geoobjekt* wird damit ein Objekt (im Sinne von OO verstanden), das eine existierendes reales geografischen Objekte repräsentiert. Dieses reale Objekt kann auf der Erde[1] identitfiziert werden.

Ein Geoobjekt kann ISO 19137 Normen als eine Aggregation folgender Objekte verstanden werden [1, 20f]:

**Geometrie** beschreibt die Form und Lage des Geoobjektes. Beschreibungsformen sind vor allem Vektor- und Rasterdaten.

**Topologie** beschreibt die Nachbarschaftsbeziehungen zu anderen Geoobjekten.

[1] In der Literatur wird in der Regel von der Erde gesprochen, obwohl die gleichen Prinzipien auch für andere Himmelskörper geeignet ist. Diese Einschränkung erklärt sich vermutlich aus der unterdurchschnittlichen Nachfrage z.B. nach Kastasterdaten für Mond oder Mars. Wir bleiben daher auch auf der Erde, obwohl die Ausführungen für alle Himmelskörper mit einer festen Oberfläche gelten.

**Thematik** sind zusätzliche Informationen für das Geoobjekt. Damit sind alle Informationsarten gemeint, die nicht in die anderen drei Klassen gehören. Etwas platt gesagt: Dafür wurden GIS gebaut. Während Geometrie, Topologie und die zeitlichen Eigenschaften das Geoobjekt und dessen Relation zu anderen beschreiben, werden hier die zusätzlichen Eigenschaften hinterlegt, also Dinge wie Einwohnerzahl, durchschnittliche Niederschlagsmenge etc. pp.

**Temporale Eigenschaften** beschreiben, wann anderen drei Eigenschaften gelten. Geoobjekt können sich durchaus verändern oder auch entstehen und vergehen. Man denke nur an Länder und deren Grenzen, die sich über die Jahre durchaus und auch grundlegend verändern können.

Ein GIS enthält eine Menge von Geoobjekten. In einem GIS spricht man von *Dimensionen* des GIS, um damit die einzelnen Eigenschaften der Geoobjekte zu bezeichnen. Man spricht also z.B. von der *geometischen Dimension* eines GIS und meint damit sämtliche geometrische Eigenschaften der Geoobjekte des GIS.

## 2.3 Geometrische Dimension

Es gibt zwei grundsätzliche Methoden, um geometrische Objekte zu erfassen, auf Details wird in anderen Veranstaltungen / Lecture Notes eingegangen.

### 2.3.1 Vektordaten

Geometrische Objekte werden mittels Vektoren beschrieben. Vektoren sind gerichtete Strecken. Sie haben Länge, Position in einem Raum von wenigstens zwei Dimensionen und eine Richtung. Mittels Vektoren lassen sich komplexere geometrische Objekte kontruieren.

Dank analytischer Geometrie kann sehr exakte Rechnungen durchführen. Vektordaten sind in der Regel auch recht kompakt speicherbar.

Der einzige Nachteil besteht in der Gewinnung der Daten. Für exakte Modelle werden exakte Messungen benötigt bzw. Methoden, um Vektordaten zu extrahieren. Vektordaten lassen sich direkt mit Methoden Vermessungskunde, siehe z.B. [2] ermitteln.

Einfachere Methoden der Ermittlung geografischer Gegebenheiten sind Bilder, d.h. Fotos. Diese Bilder recht einfach als Rasterdaten speicherbar. Die Umwandlung ist z.T. möglich. Wir schauen uns Methoden dazu später (grob) an.

### 2.3.2 Rasterdaten

Ein n dimensionaler Raum kann in in Raster eingeteilt werden. Ein Raster ist dabei in der Regel ein n dimensionaler Würfel. Nicht selten kann das Raster an verschiedenen Stellen variieren. Dabei wird in der Regel ein Würfel in weitere Würfel unterteilt.

Die Bezeichnung *Raster* kommt aus der zweidimensionalen Vorstellung und der Idee, dass ein Gitter von Quadraten über eine Karte gelegt wird. Dreidimensionale Räume lassen sich dadurch *rastern*, indem der Raum z.B. durch *Einheitswürfel*, das sind Würfel mit einer definierten und im Raster fixen und einheitlichen Kantenlänge, den Raum aufteilen.

Jedem Würfel können nun Werte zugewiesen werden. Das können einfache Werte sein, wie *belegt, frei, teilweise belegt*, siehe [1] oder auch komplexere Daten, wie Dichte / Volumen eines Gases oder ähnlichem. Ein Raster kann als Skalarfeld verstanden werden.

Die in einem GIS betrachteten Welten sind aber analog, d.h. ein Raster ist ein Modell und natürlich haben die Werte benachbarter Würfel Einfluss aufeinander. Diese Zusammenhänge sind in einem Rastermodell nicht direkt erkennbar und müssten parallel dazu beschrieben werden. So kann man in einem Rastermodell kaum die Bewegung einer Wolke beschreiben. Man kann die Würfel benennen, in der sich die Wolke zu einem Zeitpunkt befindet. Die Wolke in ihrer Gesamtheit existiert in dem Modell aber nicht. Auch die Benennung aller Würfel die die Wolke beinhalten ist kaum hilfreich, um dynamische Modelle zu bilden.

Hier sind Vektordaten deutlich besser. Objekte können in ihrer Gänze und in beliebiger Detailtiefe als ein Modell beschrieben werden. Rasterdaten können mittels Fernerkundung (z.B. Satelliten- oder Flugzeugaufnahmen) direkt produziert werden. Es sind die Methoden der Bilderkennung, die Rasterdaten in Vektordaten umwandeln.

### 2.3.3 Das Simple Feature Model

Es gibt einige einfache geometrische Objekte mit denen bereits die alten Griechen gearbeitet haben und die Sie in dem elementaren Kursen des Geometrieunterrichts kennen gelernt haben: Punkt, Gerade, Flächen.

Diese geometrischen Primitive können auf unterschiedliche Art und Weise in Software modelliert werden. Probieren Sie das einmal aus! Überlegen Sie sich, bevor Sie weiterlesen einmal, wie das Interface `Punkt` aussehen könnte. Das ist nicht schwer! Überlegen Sie.

Fertig? OK, wie wäre es mit der Variante:

```
interface Punkt {
  long getX();
  long getY();
}
```

Haben Sie eine andere Variante? Das wäre normal. Es braucht einfach einen Standard, um Programme schreiben zu können, die Geometrien speichern und welche, die damit arbeiten. Ein Standard ist der Simple Feature Access Standard der OGC, der zu finden ist unter:

`http://portal.opengeospatial.org/files/?artifact_id=25355`

Auf Seite 20 findet man die Spezifikation des Punktes. Schauen Sie sich das einmal an. Liest sich gar nicht so kompliziert, so ein Standard.

Das Modell beschreibt aber auch bereits Referenzsysteme auf Seite 15, die wir uns in der Veranstaltung anhand des Standards etwas genauer ansehen wollen.

**Well-known text format** Das Modell beschreibt ebenfalls wie diese elementaren geometrischen Objekte als Text repräsentiert werden können. Auch diese Spezifikation soll in der Lecture Note nicht abgeschrieben werden, sondern es sei verwiesen auf den Standard, Kapitel 7 oder Wikipedia:

`http://de.wikipedia.org/wiki/Simple_Feature_Access#Well-known_Text`

## 2.4 Topologische Dimension

Die Geometrie beschreibt die Gestalt von Dingen in einem Raum. Die Topologie beschreibt die Position der Dinge untereinander. Diese Betrachtung ist damit vereinfachend im Gegensatz zur Geometrie. Die Ausdehnung der Dinge wird nicht erfasst, es interessiert lediglich, ob sie zum Beispiel benachbart sind.

Die Kombination der Topologie mit der Geometrie ist höchst sinnvoll. Die Geometrie erfasst beispielsweise die Figur eines Hauses, während die Topologie erfasst, welche Häuser, Straßen etc. benachbart sind.

Wie in der Geometrie kann man auch topologische Modell mit beliebig vielen Dimensionen entwerfen. Der grundsätzliche Aufbau ist dabei immer gleich:

Ein dimensionsloser Punkt wird *Knoten (engl. Vertice)* genannt. *Kanten (engl. Edges)* verbinden Knoten. Sie haben eine Dimension. Entsprechend lassen sich mit Kanten zweidimensionale *Maschen* bilden, aus denen wiederum dreidimensionale Zusammenhänge beschrieben werden können. In [1] spricht man auch von Null-, Eins-, Zwei- und Dreizellen je Anzahl der Dimensionen, die zur Beschreibung der Elemente nötig ist.

### 2.4.1 PostGIS

PostGIS[2] ist eine kostenfreie Open Source Geodatenbank. Sie basiert auf Postgres und integriert GEOS[3]. GEOS ist eine Implementierung des Simple Feature Modells der OGC. Daher kann PostGIS direkt mit den genannten Geometrien und WKT arbeiten. Und wir arbeiten damit auch in der Übung.

## 2.5 Thematische Dimension

Die Darstellung geografische Objekte und ihrer Zusammenhänge erfolgt in einem GIS und sind in der Regel aber Mittel zu einem anderen Zweck. Es sollen Daten gesammelt werden. Hier gibt es keine Einschränkungen und wenig Vorgaben und das ist vernünftig, denn ein GIS soll prinzipiell beliebige Daten

[2] http://de.wikipedia.org/wiki/PostGIS
[3] http://de.wikipedia.org/wiki/Geometry_Engine_Open_Source

speichern können. Das können Besichtverhältnisse in den Katasterämtern sein, Verschmutzungswerte der Luft usw. usf.

Es können also Daten ganz unterschiedlicher *Themen* gespeichert werden. Die Gesamtheit der Daten eines Themas bezeichnet man auch als *thematische Dimension*. Bei der Programmierung werden in der Regel Ebenen je Thema erzeugt, die die Daten darstellen. Die Gesamtdarstellung ist dann eine Kombination aller Ebenen.

Die Bezeichnung *Ebene (Layer)* wird teilweise auch für das Datenmodell genutzt. Wird davon gesprochen, dass ein GIS einen Layer der Lufreinheit anbietet, so ist damit eigentlich gemeint, dass es eine thematische Dimension gibt, die diese Daten enthält und an der GUI mittels eines Layers visualisiert wird. Man nutzt Termini des Views, um Dinge des Models zu beschreiben, um es einmal im Vokabular von MVC[4] zu sagen.

Man spricht auch von *thematisch mehrdimensionalen* GIS, wenn mehr als eine thematische Dimension gespeichert werden kann [1].

### 2.6 Zeitliche Dimension

Geografische Objekte haben eine Dynamik. Die Aussage gilt - je nach Zeitskala - für jedes geografische Objekt. Das hat allein mit tektonischem Verschiebungen zu tun. Viele Geo-Objekte verändern sich deutlich schneller verglichen mit der Bewegung der Kontinentalplatten.

Die zeitliche Dimension beschreibt die Zeit, in der die Zuordnung der anderen Dimensionen zum Geo-Objekte gültig ist.

So kann durchaus ein Geo-Objekt Atlantis beschreiben. Nur leider existiert Atlantis nicht mehr. Mehr noch: Noch immer könnte man zu gewisser Berühmtheit gelangen, wenn man die zeitliche Dimension exakt beschreiben könnte und danach natürlich noch die geometrische oder wenigstens die topologische.

Lt. [1] ist die zeitliche Dimension bisher wenig beachtet in GIS, erlangt aber zunehmend an Bedeutung. Das hängt vermutlich auch mit der Geschichte der GIs zusammen. Läuft ein GIS länger Zeit, so wird man Veränderungen erkennen. Wie geht man damit um? Hält man nur die aktuellen Daten im GIS? Lässt man je interessierenden Zeit ein GIS laufen? Oder fügt man eine zeitliche Dimension hinzu? Aus IT Sicht ist die letzte Lösung die eleganteste.

## Literatur


[1] Ralf Bill. *Grundlagen der Geo-Informationssysteme*. Wichmann Verlag, 2010.

[2] Bertold Sparla, Peter; Witte. *Vermessungskunde und Grundlagen der Statistik f.*


[4] Und die Abkürzung schreibe ich hier nicht mehr aus. Die sollten Sie kennen ;)