

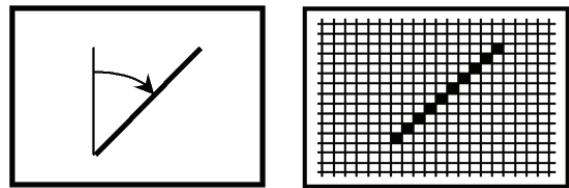
# 37. Geodaten und Geodatenformate

Im GIS werden verschieden Arten von Daten unterschieden. Es gibt die graphischen Daten, die sich in Vektor- und Rasterdaten unterscheiden, und es gibt die alphanumerischen Daten, die Sach- und Metadaten.

## 37.1. Vektordaten

Vektordaten werden durch Koordinaten beschrieben. Je nach Anzahl der Koordinaten ergeben sich die verschiedenen Datentypen.

- Punkt: Eine Punkt wird durch eine Koordinate beschrieben.
- Linie: Eine Linie hat einen Anfangs- und einen Endpunkt.
- Polygon: Ein Polygon ist ein Linienzug mit einem oder mehreren Stützpunkten, an denen sich die Richtung ändert.
- Fläche: Eine Fläche besteht aus einem Polygon, dessen Anfangs- und Endpunkt identisch ist.



Bitmap einer Linie als Vektorgraphik (links: mit Anfangspunkt, Länge und Winkel ) und als Rastergraphik (rechts: Punktraster mit jeweiligem Farbwert)

Abbildung 37.1.1.

Kurven und Kreise sind Sonderfälle von Polygonen und Flächen, auf die aber hier nicht näher eingegangen wird. Da Vektordaten mit ihren Koordinaten bekannt und gespeichert sind, können sie für Berechnungen und Analysen, z. B. Flächenberechnungen oder Verschneidungen, direkt genutzt werden.

## 37.2. Mehrdimensionale Daten

Insbesondere bei den Vektordaten unterscheidet man zweidimensionale Koordinaten (Standard Rechts- und Hochwert) und mehrdimensionale Koordinaten. Die mehrdimensionalen Daten sind im Kapitel 30.4.1.3 genauer beschrieben, hier nur eine kurze Auflistung:

**2D** Standardbezeichnung für "flache" Koordinaten, d. h. nur Rechts- und Hochwert.

**2,5D** Zweidimensionale Koordinaten (Rechts- und Hochwert) und einen Attributwert für die Höhe.

**3D** "Echte" dreidimensionale Koordinaten mit Rechtswert, Hochwert und Höhe.

**4D** Dreidimensionale Koordinaten, ergänzt um eine Zeit, über die eine Veränderung in Abhängigkeit zur Zeit dargestellt werden kann (z. B. die Flugroute von Zugvögeln).

## 37.3. Rasterdaten

Rasterdaten (s. Abbildung 37.1.1 rechts) sind digital gespeicherte Bilder, wie z. B. Luftbilder aus einer Befliegung oder gescannte Karten und Pläne. Rasterdaten liegen in einem eigenen, lokalen (Pixel-)Koordinatensystem vor und müssen vor ihrer Benutzung in einem GIS georeferenziert werden, d. h. jedem Bildpunkt (Pixel) wird durch eine Transformation eine Koordinate im entsprechenden Bezugssystem zugewiesen. In einem Rasterbild hat jedes Pixel einen eindeutigen Farbwert.

### 37.3.1. Farbtiefe

Unter der Farbtiefe eines Rasterbildes versteht man den Wertebereich an möglichen Farben für ein Pixel.

- 1 Bit: Jedes Pixel kann den Wert 0 oder 1 haben, also i. A. schwarz oder weiß. Diese Farbtiefe wird häufig für einfarbige Vorlagen wie alte Lagepläne oder Lichtpausen verwendet. Durch geschickte Veränderungen des Schwellenwertes (wann soll der Punkt aus der Vorlage schwarz oder weiß sein) können so sehr kontrastreiche Vorlagen für die Digitalisierung entstehen.
- 8 Bit: Jedem Pixel können 256 verschiedene Farbwerte zugeordnet werden. Diese Einstellungen werden oftmals für gescannte Pläne mit einer überschaubaren Anzahl verschiedener Farben eingesetzt, wie z. B. Bebauungspläne. Ebenfalls kann man mit 256 Farben eine ausreichende Abstufung für Graustufenbilder erreichen, wenn eine reine schwarz/weiß Darstellung nicht ausreicht.
- 24 Bit: Jedem Pixel werden drei Farbkanäle zu je 256 Farben zugeordnet. So können ca. 16,78 Mio. Farben dargestellt werden. Diese Farbeinstellung ist heute auf den meisten Computersystemen Standard. Bei 24 Bit Farbtiefe wird dem menschlichen Auge ein natürlicher Eindruck vermittelt. So eignet sich diese Farbtiefe besonders für Fotos, wie z. B. Luftbilder.

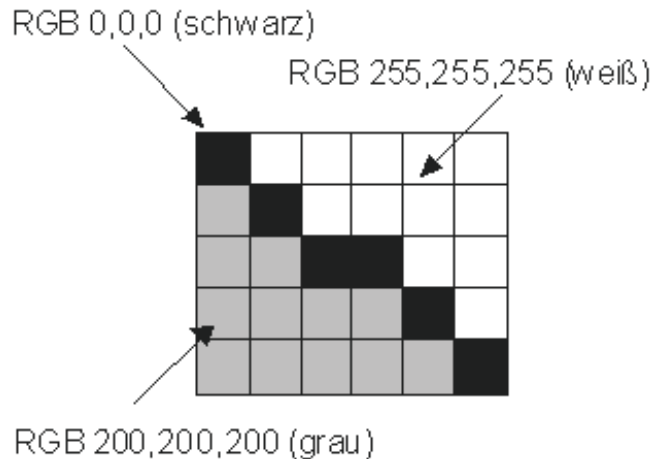


Abbildung 37.3.1.: Farbwerte eines Pixels

Entscheidend für die Wahl der Farbtiefe ist der Speicherbedarf und die damit verbundene Verarbeitungsgeschwindigkeit sowie die Qualität und Art der Weiterverarbeitung des Plans. Ein gescannter Plan von 10 000 x 5000 Pixeln (etwas kleiner als eine Katasterkarte) belegt etwa den folgenden Speicherplatz:

	1 Bit	8 Bit	24 Bit
Kilobyte	6.104	48.828	146.484
Megabyte	ca. 6	ca. 48	ca. 143

### 37.3.2. Komprimierung

Um den Speicherbedarf zu reduzieren existieren verschiedene Komprimierungsverfahren, von denen hier die gängigsten kurz erläutert werden sollen: Man unterscheidet zwischen verlustfreien und verlustbehafteten Verfahren. Bei den verlustfreien Verfahren kann nach der Dekomprimierung, d. h. dem Wiederherstellen des „Originals“, über das gleiche Bild verfügt werden. Bei der verlustbehafteten Komprimierung werden jedoch benachbarte Pixel mit ähnlichen Farbwerten zu gleichen zusammengefasst, d. h. nach dem Dekomprimieren ist das „Original“ nicht wieder herstellbar. Diese Methode ist im GIS-Bereich nur bedingt geeignet, da ja einem einzelnen Pixel Koordinaten und Informationen (sein Farbwert) zugeordnet sind. Werden diese Farbwerte nur durch eine Zusammenfassung geändert, so ändert sich auch die Information. Bei einem Luftbild können so beispielsweise verschiedene Grüntöne, die unterschiedliche Flächen darstellen, nicht mehr unterschieden werden. CCITT 4 (Fax Group) (für 1 Bit), LZW und PackBits sind gängige Vertreter der verlustfreien Komprimierung. Die JPEG-Komprimierung zählt zu den verlustbehafteten Methoden.

### 37.3.3. Datenformat

Für die verschiedenen Farbtiefen und Datenarten haben sich unterschiedliche Dateiformate entwickelt. Einige häufig verwendete sollen an dieser Stelle kurz verglichen werden:

**JPEG/JPG** Das JPEG-Format ist vor allem aus der Digitalfotografie bekannt. Es besitzt unterschiedlichste Komprimierungsstufen und speichert eine Farbtiefe von 24 Bit. Eine stark komprimierte JPEG-Datei kann nicht wieder in den Ursprungszustand zurückversetzt werden. Es unterstützt ohne weiteres keine Transparenz.

**JPEG 2000** Das JPEG 2000-Format ist ein ISO-Standardisiertes Dateiformat mit verschiedenen Möglichkeiten der Bildkompression. Das Format ist in vielen Teilen lizenzpflichtig.

**GIF** Das GIF-Format kann Daten bis zu einer Farbtiefe von 8 Bit speichern. Es eignet sich also eher für Pläne und Grafiken. Das Format ist wegen der guten, verlustfreien Komprimierung im Internet weit verbreitet. Ein Farbwert im GIF-Format kann als Transparent definiert werden. Durch einen langjährigen Patentschutz ist es im GIS-Umfeld jedoch weniger verbreitet.

**PNG** Das PNG-Format wurde wegen des Patentbesitzes auf GIF und der Nachteile, dass GIF nur 256 Farben und JPEG keine Transparenz unterstützt als Nachfolgeformat eingeführt. Es beherrscht 24 Bit Farbtiefe und Transparenz.

**TIFF** Das TIFF-Format ist eines der gebräuchlichsten und universellsten Rasterdatenformate und daher ein beliebtes Austauschformat im GIS-Umfeld. Beim TIFF-Format können die verschiedenen Komprimierungsarten frei gewählt werden. Ebenso können zum TIFF-Format Informationen zur Georeferenzierung abgespeichert werden, entweder als GeoTIFF in der gleichen Datei oder durch Abspeichern in einer separaten Datei. Die separate Datei mit der Endung .tiff wird dann mit dem gleichen Namen im gleichen Verzeichnis abgespeichert.

**ECW** Das ECW-Format (Enhanced Compression Wavelet) ist ein Bildformat, welches für Luft- und Satellitenbilder optimiert entwickelt wurde. Die Kompression ist für sehr große Bilder mit feinen wechselnden Kontrast sehr effizient.

### 37.3.4. Auflösung

Zu jeder Rasterdatei wird die Auflösung in dpi (dots per inch, Punkte pro Zoll) angegeben. Dies ist für einen späteren Ausdruck von Rasterdaten wichtig. So unterscheidet sich die Ausgabe von Rasterdaten am Bildschirm (72 oder 96 dpi) von der auf einem Drucker (z. B. 300 dpi). Ebenso bestimmt man mit der Auflösung die Qualität und den dadurch benötigten Speicherplatz, mit der ein Plan gescannt wird.

Din A4	72 dpi	300 dpi	600 dpi
21,0 cm	596 px	2481 px	4960 px
29,7 cm	842 px	3508 px	7016 px

Grundsätzlich gilt beim Scannen, dass die höchste benötigte Anforderung erfüllt wird. Soll ein Plan z. B. am Bildschirm angezeigt und später ausgedruckt werden, sollte die Auflösung mindestens 300 dpi (kleinste Druckerauflösung) betragen. Der Wert von 300 dpi stellt also in der Regel einen guten Wert für einen Scan dar.

## 37.4. Sachdaten

Unter Sachdaten versteht man die alphanumerischen, also nicht graphischen Daten, die zu einem Objekt gespeichert werden (s. Abbildung 37.4.1). Sie werden in Tabellen in einer Datenbank oder Datei gespeichert und sind einem Objekt eindeutig zuzuordnen. Die einzelnen Datenfelder können, um die Qualität der Eingabe zu verbessern, als Pflichtfelder oder optionale Felder definiert werden. Bei Pflichtfeldern muss eine Eingabe eines Datensatzes erfolgen, bei optionalen Feldern kann eine Eingabe erfolgen, wenn ein Wert vorhanden ist. Um die Qualität der Eingabe weiter zu verbessern und die Daten später auswerten zu können, kann bestimmt werden, welche Art von Daten in einem Feld (Attribut) gespeichert werden dürfen. Dies können z. B. Texte, Zahlenwerte, ein Datum, ja/nein-Felder u.v.a. sein. Die Sachdaten bilden vielfach die Grundlage für spätere Analysen. Nur wenn die Sachdaten vollständig vorhanden sind, können später verwertbare Ergebnisse erzielt werden. Um z. B. die Qualität von Eingaben zu erhöhen wird oftmals mit Auswahllisten und Schlüsselstabellen gearbeitet, da so Eingabe-/Schreibfehler vermieden werden können.



Name	Wert
FOLIE	011
OSCHL	1001
QART	1031
OBJNUM	7ZHL
OBJNAME	HA05122000021210022 001
DATEIKZ	HA
LAND	05
REGBEZIRK	1
KREIS	22
GEMEINDE	000
STRASSE	01234
HAUSNR	22
HAUSNRZUS	
LFDNR	1
FLAECHE	123,984
NUMBEZ	25567569
AKTOBJ	
EDAT	14.10.1997
AINFO	16
KTYP	NN
ROW_ID	17686

Abbildung 37.4.1.: Sachdatenfenster zu einem Objekt

## 37.5. Metadaten

Um mit Daten in einem GIS plausibel arbeiten zu können benötigt man Daten, die Aussagen zur Herkunft, Qualität und Aktualität geben. Dies sind die so genannten Metadaten. Sie beschreiben z. B., wann ein Datensatz erfasst wurde, auf welcher Grundlage und wer dafür zuständig ist, um ggf. Rückfragen zu beantworten. Die Definition von Metadaten ist in einer ISO-Norm (ISO 19115) festgelegt. Die wichtigsten Metadaten sind hier kurz erläutert:

- Aktualität  
Von wann sind die Daten? Ist es der aktuelle Stand oder gibt es verschiedene Versionen?
- Herkunft und Ansprechpartner  
Von wem stammen die Daten und wen kann man über die Daten befragen?
- Bezugssystem  
Mit welchem Bezugssystem wurden die Daten erfasst? Sind es Koordinaten aus dem 2. oder 3. Streifen im Gauß-Krüger-System, ETRS89 / UTM oder andere Koordinaten?
- Erfassungsmaßstab/Maßstabsebene  
Auf welcher Grundlage wurden die Daten erfasst? Wasserschutzgebiete, die z. B. von einer Karte im Maßstab 1 : 50 000 digitalisiert wurden sind nicht unbedingt für einen Plan im Maßstab 1 : 500 geeignet.

## 37.6. Datenaustauschformate

Die in einem GIS zur Verfügung stehenden Daten müssen für die unterschiedlichsten Zwecke aufbereitet werden. Für Planungsverfahren müssen Grundlagendaten an einen Dritten abgegeben werden oder fertige Analysen sollen als Plan oder Karte zu Papier gebracht werden.

Für die Austauschformate gilt ähnliches wie für den Import. Entscheidend für das geeignete Datenformat sind der Verwendungszweck und das Zielsystem. Da in einem GIS die Daten und die Ausgestaltung der Daten getrennt voneinander sind, kommen beim Austausch häufig nur die Geometrien im Zielsystem an. Für die richtige Ausgestaltung werden weitere Informationen benötigt. Dies kann z. B. durch das Erzeugen eines Papierabzuges der Karte oder die Erstellung einer Ansichtdatei als Rasterbild oder im PDF-Format erfolgen. Mit dem Ziel, künftig sowohl die Definitionsgeometrie mit ihrer Ausgestaltung als auch die Attribute zu Objekten automatisiert übergeben zu können, werden zurzeit internationale Standards für die Schnittstellen von Geodaten erarbeitet.

**EDBS/WLDGE** EDBS und WLDGE sind die "alten" Austauschformate für Daten aus dem Liegenschaftskataster (ALK/ALB) oder ATKIS (vgl. Kapitel 27.2.2). Mit dem BZSN-Verfahren (Bezieher Sekundärnachweis) ist eine Methode für eine Differenzdatenabgabe vorhanden.

**NAS** Die Normbasierte Austauschschnittstelle (NAS) ist das neue Standardformat für den Datenaustausch im AAA-Modell (vgl. Kapitel 27.3.1). Mit dem NBA-Verfahren (Nutzerbezogene Bestandsdatenaktualisierung) existiert auch hier eine Definition für eine Differenzdatenabgabe.

**shape** Das ursprünglich von der Firma ESRI entwickelte shape-Datenformat hat sich in der Zwischenzeit als Quasistandard im GIS-Umfeld entwickelt. Es existieren mittlerweile auch freie Bibliotheken, die dieses Datenformat beherrschen.

**dxf/dwg** Das dxf bzw. dwg-Format hat sich als Industriestandard für den Austausch von CAD-Daten etabliert. Beim Austausch von DXF/DWG ist insbesondere auf die entsprechende Version des Formats zu achten, da diese sich vielfach unterscheiden.

- ACAD 14
- ACAD 2000
- ...

**diverse** Daneben existieren noch weitere Austauschformate. Teilweise handelt es sich hierbei um Herstellerformate. Exemplarisch sind die nachfolgenden Formate genannt:

- MapInfo tab
- Microstation dgn
- Google kml
- ...

## 37.7. Datenorganisation im GIS

Beim Aufbau und Einsatz von GIS gibt es einige Grundprinzipien, wie z. B. die Aufteilung von Themen in verschiedene Ebenen und Layern oder die Trennung der Grafik von der Ausgestaltung.

### 37.7.1. Objektklassen und Objekte

Objektklassen bilden die Definition von Objekten gleichen Typs, z. B. Flurstücke. In einer Objektklasse wird beschrieben, was alles zu einem Flurstück gehört und welche Daten dazu gespeichert werden. Für ein Flurstück kann beispielsweise festgelegt sein, dass es sich um eine Fläche handelt und das als Sachdaten die Gemarkung, die Flur und die Flurstücksnummer zu führen sind. Ein Objekt ist dann das „echte“ Flurstück, das in dem System erfasst wird und die Werte „Gemarkung“, „Flur“ und „Flurstücksnummer“ zugewiesen bekommt.

### 37.7.2. Thematische Ebenen

In einer Ebene oder einem Layer werden thematisch zusammengehörende Objektklassen zusammengefasst. Damit ist eine logische Aufteilung im GIS möglich. Durch die Gruppierung von Objektklassen in Ebenen lassen sich ganze Themenbereiche, wie die Liegenschaftskarte in den Vordergrund oder in den Hintergrund legen, je nach dem ob man z. B. das Kataster als Vordergrund auf einem Luftbild oder als Hintergrund für einen Bebauungsplan benötigt. Ebenso wird eine Steuerung der Maßstabsebenen ermöglicht. So kann bei großen Datenbeständen die Datenmenge besser begrenzt werden. Die Liegenschaftskarte wird zum Beispiel nur bis zu einem Maßstab von 1 : 2000 angezeigt, im Bereich 1 : 2000 bis 1 : 8000 wird die DGK 5 bzw. ABK angezeigt und ab 1 : 8000 erfolgt die Darstellung der DTK10, um nicht unnötig viele Objekte laden zu müssen.

### 37.7.3. Ausgestaltung

Unter der Ausgestaltung oder Ausprägung versteht man die graphische Darstellung der Objekte. Diese kann sich je nach Aufgabenstellung unterscheiden. Man kann also in verschiedenen Projekten dem gleichen Objekt ein anderes Aussehen geben. So kann man beispielsweise die Liegenschaftskarte über ein Luftbild legen. Um eine bessere Lesbarkeit zu erhalten, ändert man die Ausgestaltung, in dem man die Linienbreite erhöht und die Farbe