



Vom Shapefile zum GeoPackage

Ein kurzer Überblick über die wichtigsten GIS-
Datenformate

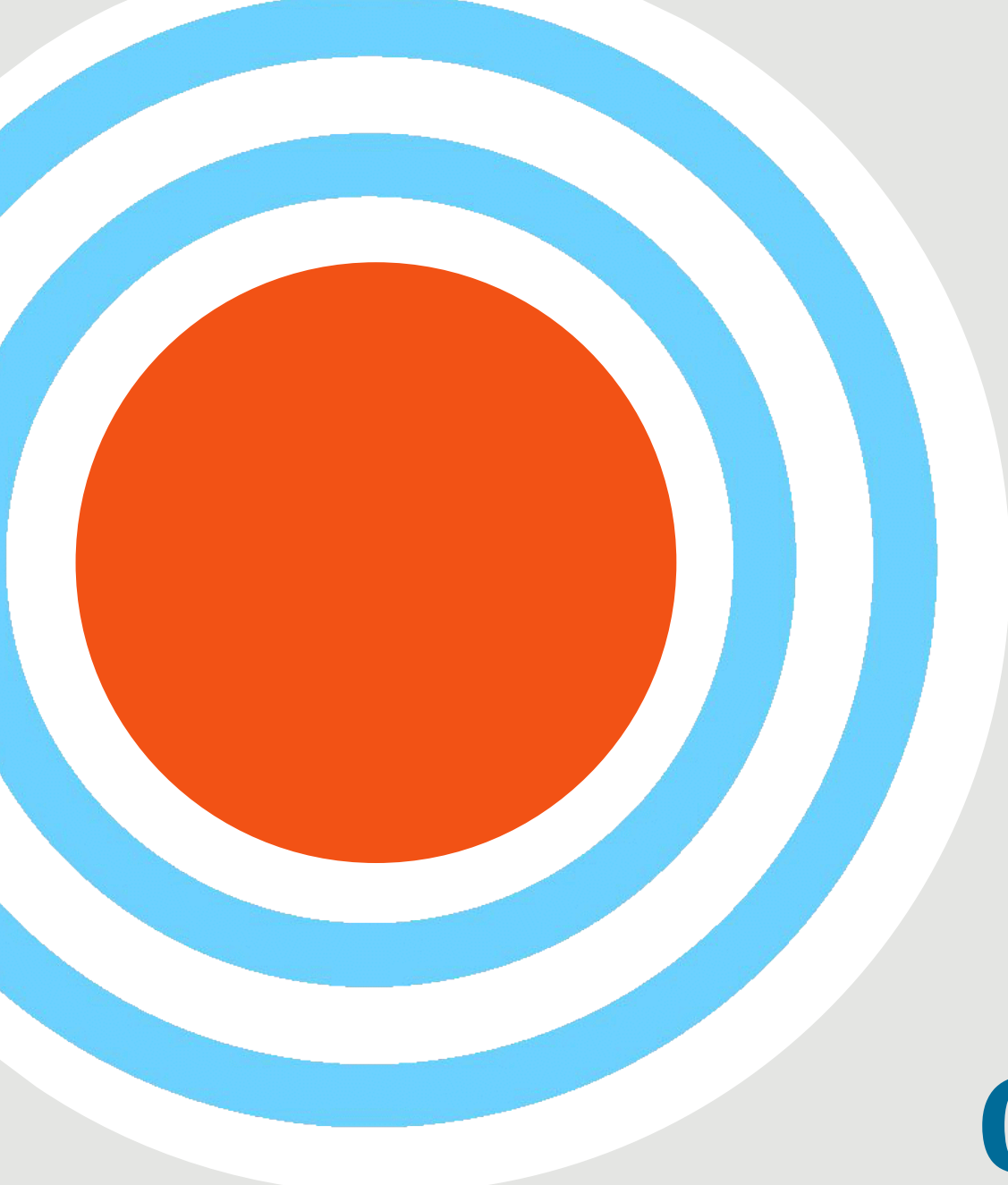


Grundlagen

Entwicklung

GIS-Datenformate

Signierung

A large graphic on the left side of the slide consisting of three concentric circles. The innermost circle is solid orange, the middle ring is light blue, and the outermost ring is white.

Was sind GIS-Datenformate



Definition

GIS-Datenformate sind spezielle Dateitypen, die geografische Informationen und deren Attribute speichern.

Sie bilden die Grundlage für die Arbeit in GIS-Programmen und ermöglichen die Analyse, Darstellung und Verwaltung von räumlichen Daten.

👉 **GIS-Datenformate dienen dazu, räumliche Informationen strukturiert zu speichern und auszutauschen, sodass sie in Geoinformationssystemen visualisiert und analysiert werden können.**



Warum braucht man Datenformate?

WKT ist eine OGC-konforme Textdarstellung von Geometrien. Es beschreibt einfache geometrische Objekte in einer einfacher Textform:

```
POINT (8.6821 50.1109)  
LINESTRING (30 10, 10 30, 40 40)  
POLYGON ((30 10, 40 40, 20 40, 10 20, 30 10))
```

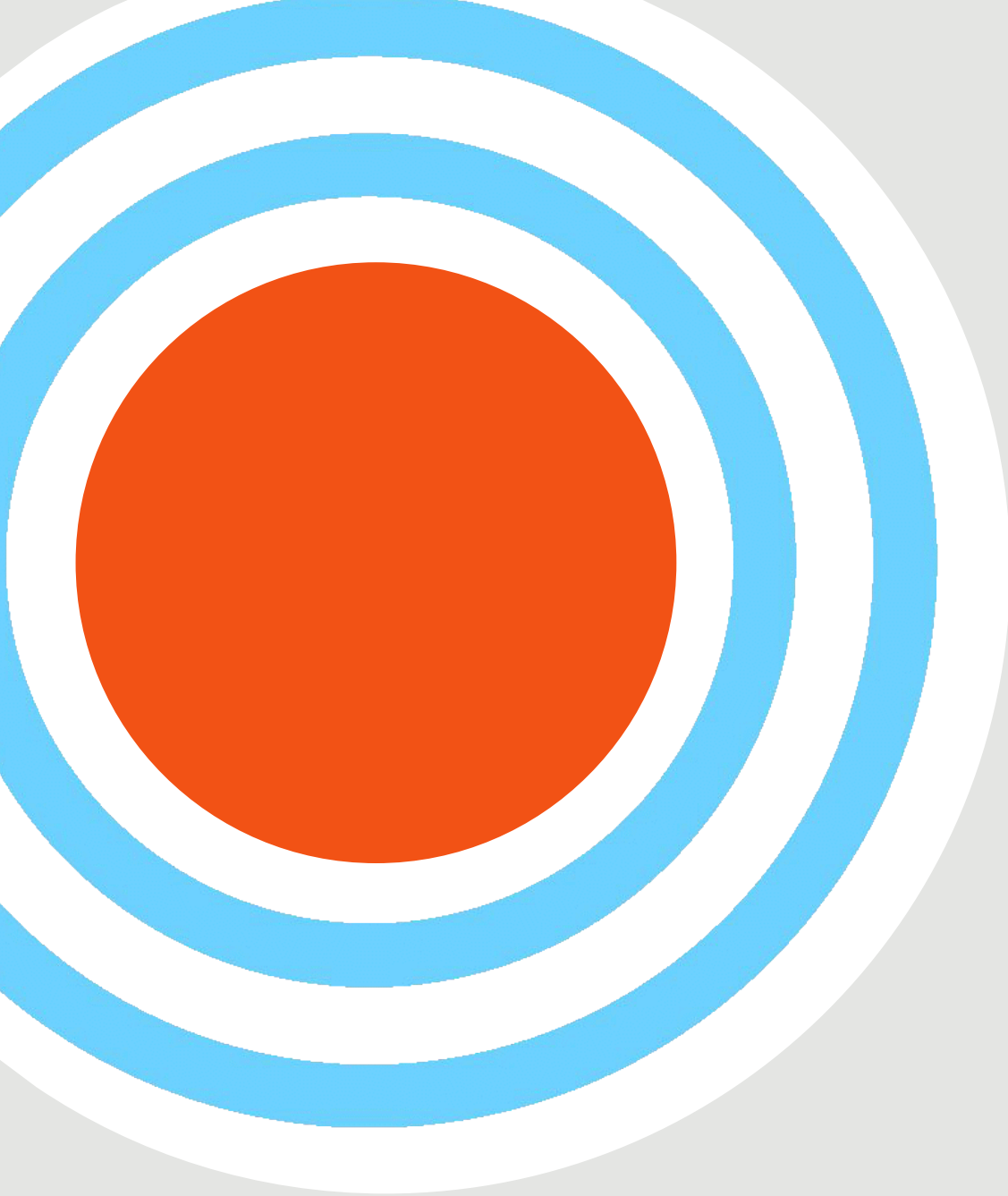
WKT dient jedoch nicht der Speicherung kompletter GIS-Datensätze, sondern der Beschreibung einzelner Geometrien. Es ist eine Basistechnologie, kein Ersatz für GIS-Datenformate.

Es eignet sich hervorragend zum Darstellen, Debuggen oder Übertragen einzelner Geometrien, nicht jedoch zur Speicherung oder Analyse kompletter Geodatenprojekte.



Kennzeichen von GIS-Datenformaten

- Beinhalten räumliche Objekte als Geometrien im Raum (Punkt, Linie, Fläche)
- Verknüpfung von Sachinformationen (Attributen) mit räumlichen Objekten
- Raumbezugssystem (Koordinatensystem)
- Unterstützung von Topologien (Beziehung von Objekten zueinander)
- Metadaten (Nachvollziehbarkeit und Qualitätssicherung)
- Unterstützung räumlicher Abfragen und Indizes
- Interoperabilität (Austauschfähigkeit mit unterschiedlichen Systemen)
- Strukturierung in Layer und Feature-Klassen
- Informationen zur Darstellung der Geometrien (Symbole, Schriften, Schraffuren etc.)



**Welche Arten von GIS-
Formaten gibt es?**



Datenarten

Vektordatenformate:

Diese speichern geografische Objekte in Form von Punkten, Linien oder Polygonen. Beispiele:

- Shapefile: Eines der bekanntesten Formate, bestehend aus mehreren Dateien.
- GeoJSON: Textbasiertes Format auf JSON-Basis, gut für Webanwendungen geeignet.
- GeoPackage: Modernes, kompaktes SQLite-basiertes Format für Vektor- und Rasterdaten.

Rasterdatenformate:

Diese speichern Daten als Pixelraster, ähnlich wie ein digitales Bild. Jedes Pixel enthält einen Wert (z. B. Höhe, Temperatur, Satellitenreflexion). Beispiele:

- GeoTIFF: Häufig für Fernerkundungsdaten und Höhenmodelle verwendet.
- ECW: Format zur verlustbehafteten Speicherung sehr großer Rasterdaten.

Zusätzlich gibt es Datenbank- und Webformate wie PostGIS (räumliche Erweiterung von PostgreSQL) oder WMS/WFS (Webdienste für Karten und Features).



OpenSource vs. Proprietär

Vorteile von Open-Source-Formaten

Open-Source-Formate bieten Transparenz, Flexibilität und starke Community-Unterstützung, was eine einfache Integration und Innovation ermöglicht.

Stärken von proprietären Formaten

Proprietäre Formate sind oft leistungsfähiger und besser in bestimmte Softwarelösungen integriert, können jedoch das Teilen und Bearbeiten einschränken.

Beispiele: **ESRI File Geodatabase, MapInfo TAB/MIF, GeoMedia SmartStore, , Bentley MicroStation Design File**

Die richtige Formatwahl

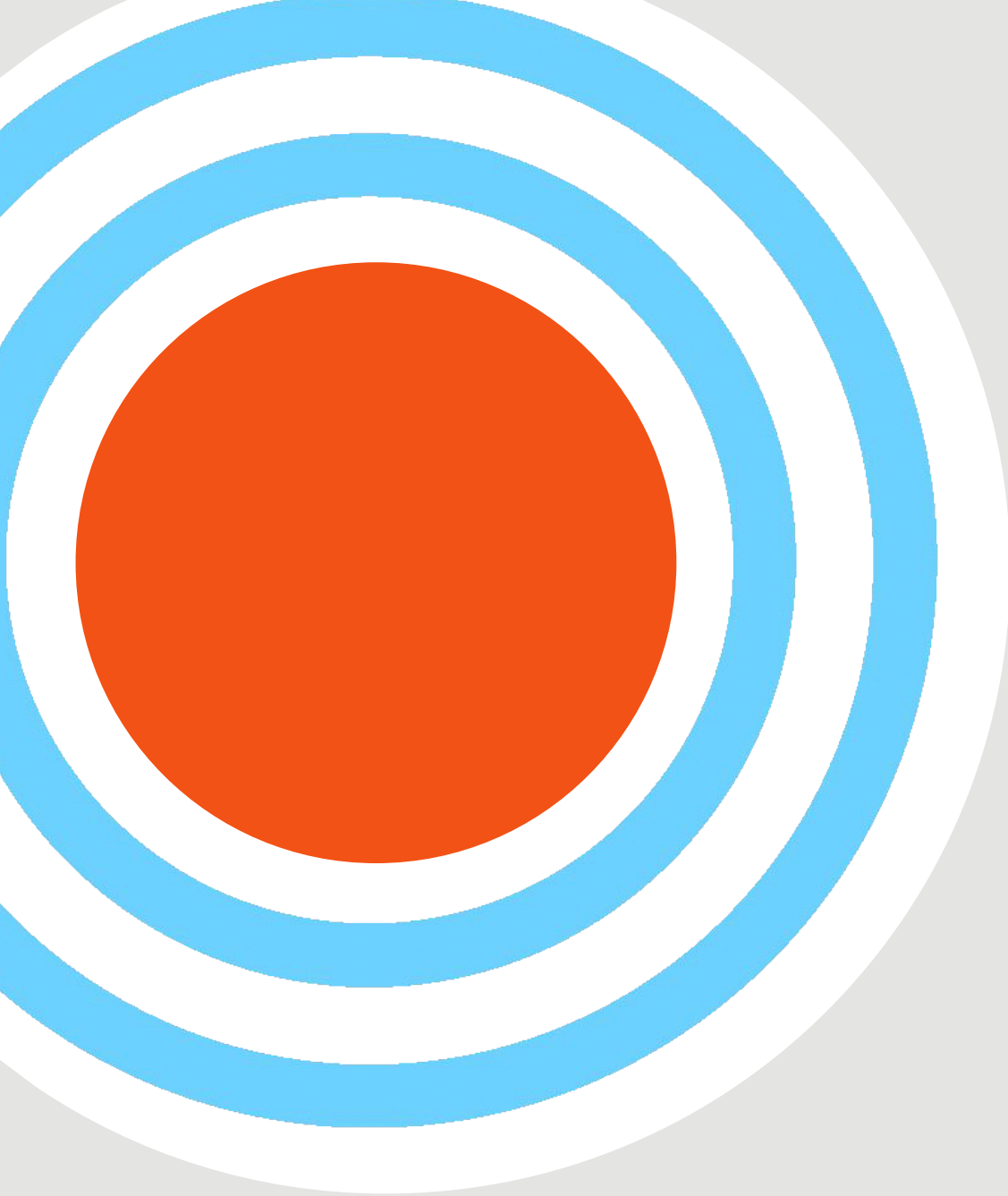
Die Entscheidung zwischen den Formaten hängt von den Projektanforderungen, der benötigten Funktionalität und langfristigen Überlegungen zur Datenstrategie ab.



Wieviele Datenformate gibt es?

- 15 – 20 Kerndatenformate (etabliert und weit verbreitet)
z.B. Shapefile, GeoPackage, GeoJSON, GeoTIFF, FileGDB, GML
- 20 – 80 Erweiterte Formate (Spezialformate, proprietär)
z.B. LAS, NetCDF, CityGML
- ca. 30 OGC-Standardisierte Formate
z.B. GML, KML, GeoPackage, GeoTIFF, 3D Tiles

ca. 250 Formate insgesamt (inkl. historischer Formate)



Entwicklung



1970er–1980er

📄 Erste digitale Kartenformate
E00, Coverage, GRID – einfache
Raster- und Vektorstrukturen

2000er

🌐 XML & Web-GIS
GML (OGC-Standard), KML (Google Earth),
WMS/WFS-Webstandards

2020er–heute

🔒 Integration, 3D & Signierung
3D Tiles, CityGML, GeoParquet,
Authentifizierte GeoPackages

1990er

📁 Shapefile & GeoTIFF
ESRI führt Shapefile ein
GeoTIFF wird Standard für Rasterdaten

2010er

📦 Moderne, offene Formate
GeoPackage (OGC 2014),
GeoJSON, Spatialite



Aktuelle Trends

Echtzeit-Datenstreaming

Das Streaming von Geodaten ermöglicht die Echtzeitverarbeitung räumlicher Informationen für sofortige Erkenntnisse und Entscheidungsfindung.

Semantische Datenformate

Semantische Formate bereichern GIS-Daten mit Bedeutung, was die maschinelle Interpretation und Interoperabilität verbessert. z.B. RDF

KI-Integration in GIS

Die Integration von KI in GIS erfordert Datenformate, die große Datensätze effizient verarbeiten, um fortgeschrittene Analysen zu unterstützen.

Cloubasierte GIS-Lösungen

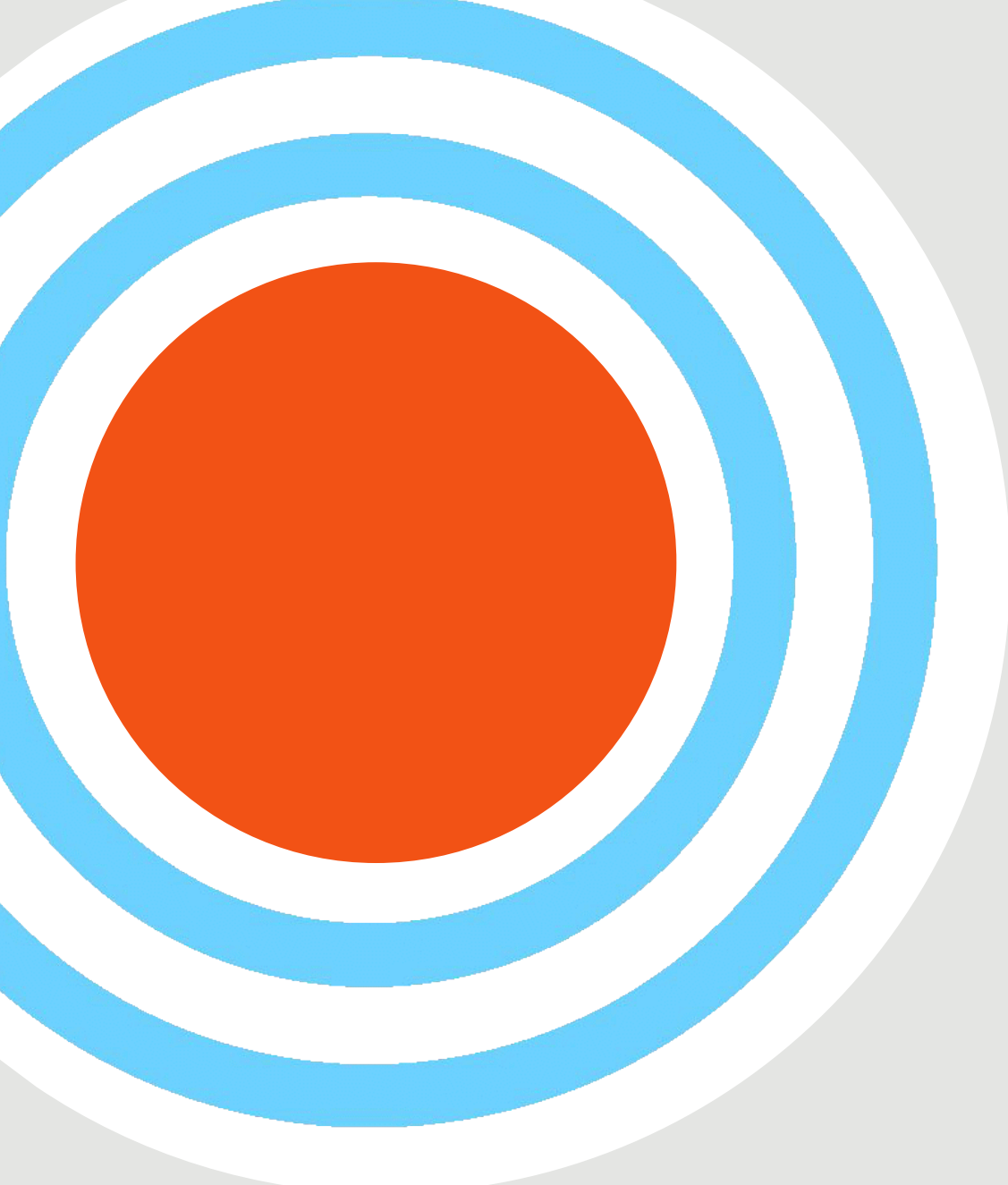
Cloud-GIS-Lösungen setzen auf Datenformate, die für verteilte Systeme und skalierbare Rechenumgebungen optimiert sind.

NoSQL, Datalakes, BIGData

NoSQL, Data Lakes und GeoParquet markieren den Übergang von klassischen GIS-Datenformaten zu datenbank- und cloudzentrierten Geodaten-Ökosystemen.

Datenaustausch über Schnittstellen (API)

Die GIS-Welt bewegt sich zunehmend weg von Datei-basierten Formaten hin zu API-zentrierten Ökosystemen, in denen Daten nicht mehr exportiert, sondern nur noch abgefragt werden und Analysen in der Cloud stattfinden.



Überblick über die wichtigsten GIS-Datenformate



Shapefile

Das Shapefile ist eines der ältesten und am weitesten verbreiteten GIS-Vektorformate, ursprünglich von **ESRI** entwickelt.

Aufbau:

Ein Shapefile besteht aus **mehreren Dateien**, die zusammengehören (z. B. `.shp`, `.shx`, `.dbf`, `.prj`)

- `.shp` → Geometrien (Punkte, Linien, Polygone)
- `.dbf` → Attributtabelle (Datenbank im DBase-Format)
- `.shx` → Index der Shapes (Offset + Länge)

Vorteile:

- Mit Abstand am weitesten verbreitetes Geo-Datenformat
- Kompatibel mit nahezu allen GIS-Programmen
- Geringer Speicherplatzbedarf, relativ effizient in Bezug auf die Dateigröße. Die resultierende Datei ist im Vergleich zu anderen (meist textbasierten) Formaten relativ klein.
- Proprietär aber Offene Spezifikation → hohe Interoperabilität

Shapefile must die!



**NO
I PREFER
GeoPackage!**



- Kein einheitliches Dateiojekt (mehrere Dateien nötig)
- Begrenzte Feldnamenlänge (max. 10 Zeichen)
- Keine Unterstützung für Unicode oder 3D-Geometrien



<http://switchfromshapefile.org> ;

Shapefile must die!

08.10.2017 ... Alternatives · OGC GeoPackage · FlatGeobuf · GeoJSON · OGC GML · SpatiaLite · CSV · OGC KML · ESRI GeoDatabase.

Shapefile is a bad format

Why is Shapefile so bad? Here are several reasons why the Shapefile is a bad format and you should avoid its usage:

- No coordinate reference system definition.
- It's a multiformat.
- Attribute names are limited to 10 characters.
- Only 255 attributes. The DBF file does not allow you to store more than 255 attribute fields.
- Limited data types. Data types are limited to float, integer, date and text with a maximum 254 characters.
- Unknown character set. There is no way to specify the character set used in the database.
- It's limited to 2GB of file size. Although some tools are able to surpass this limit, they can never exceed it.
- No topology in the data. There is no way to describe topological relations in the format.
- Single geometry type per file. There is no way to save mixed geometry features.
- More complicated data structures are impossible to save. It's a "flat table" format.
- There is no way to store 3D data with textures or appearances such as material definitions. There are no objects.
- Projections definition. They are incompatible or missing.
- Line and polygon geometry type, single or multipart, cannot be reliably determined at the layer level.
- There is no NULL value, it is painful for numeric values
- Add more ...





OGC GeoPackage

Das **GeoPackage** ist ein modernes, offenes Format, basierend auf einer **SQLite-Datenbank**. Es kann sowohl Vektor- als auch Rasterdaten speichern.

Aufbau:

Eine einzige Datei (.gpkg), die Tabellen mit Geometrien, Attributen und Metadaten enthält.

Vorteile:

- Sehr kompakt (eine Datei .gpkg)
- Unterstützung für mehrere Layer in einer Datei
- Standardisiert durch **OGC (Open Geospatial Consortium)**
- Unterstützt 3D-Geometrien und komplexe Datentypen

Nachteile:

- Weniger verbreitet als Shapefile, aber zunehmend Standard
- Etwas komplexer in der internen Struktur



	<u>GeoPackage</u>	<u>Shapefile</u>
Dateiobjekt	Alles in einer einzigen Datei (.gpkg)	Besteht aus mehreren Dateien (.shp, .shx, .dbf, .prj, ...).
Dateigröße	Unterstützt große Datenmengen (>4 GB problemlos)	Begrenzung auf 4 GB und max. 2 GB pro Shape-Datei
Attribute	Keine relevante Begrenzung bei Feldnamen oder -typen	Max. 10 Zeichen pro Feldname, nur einfache Datentypen (z. B. kein Text > 254 Zeichen).
Encoding	Unterstützt utf-8 → Umlaute, Sonderzeichen, internationale Zeichen	Nur ANSI-Kodierung, was zu Zeichenproblemen führen kann
Projektion / Metadaten	Speichert Koordinatensysteme, Metadaten und Stile in der Datei	Benötigt eine separate .prj-Datei
Featureklassen / Layer	Kann mehrere Layer (Vektor + Raster) in einer Datei speichern	Nur ein Layer pro Datei
Standard	Offener OGC-Standard	Proprietäres ESRI-Format (zwar dokumentiert, aber alt)
Erweiterungen	Unterstützt Erweiterungen (z. B. 3D, Raster, Styles, SQL-Abfragen)	Kein Support für moderne GIS-Funktionen

Shape vs. GeoPackage



Vorteile von GeoPackage

- Weniger Chaos, einfacher zu kopieren, verschicken und versionieren.
- Ideal für umfangreiche Projekte und große Datensätze.
- Mehr Flexibilität bei Datenstrukturen und längeren Attributnamen.
- Sprach- und systemunabhängig.
- Selbstbeschreibend und vollständiger.
- Kompakter, strukturierter und besser für Projektorganisation.
- Zukunftssicher und interoperabel.
- Wächst mit neuen Anforderungen.

Das GeoPackage ist das zeitgemäße, offene und leistungsfähige Nachfolgeformat des Shapefiles – kompakter, flexibler und besser für moderne GIS-Projekte geeignet.

Nur wer mit sehr alten GIS-Systemen arbeitet, hat heute noch echte Gründe, Shapefiles zu verwenden!



```
{  
  "type": "Feature",  
  "geometry": {  
    "type": "Point",  
    "coordinates": [125.6, 10.1]  
  },  
  "properties": {  
    "name": "Dinagat Islands"  
  }  
}
```

GeoJSON

Ein modernes, textbasiertes Format auf **JSON-Basis** (JavaScript Object Notation). Es wird häufig in **Web-GIS-Anwendungen** verwendet.

Aufbau:

Alle Daten (Geometrien + Attribute + Koordinatensystem) stehen in einer einzigen, lesbaren Textdatei.

Vorteile:

- Einfach zu lesen und zu bearbeiten (auch mit Texteditoren)
- Ideal für den Datenaustausch im Web
- Unterstützt verschiedene Geometrie-Typen (Point, LineString, Polygon usw.)

Nachteile:

- Größere Dateigröße im Vergleich zu binären Formaten
- Weniger effizient bei sehr großen Datensätzen



CSV

Einfaches Text-Tabellenformat (Comma-Separated Values), das Spalten durch ein Trennzeichen voneinander abgrenzt.
Kein „echtes“ GIS-Format, wird aber häufig genutzt, um Punktdaten mit Koordinaten zu speichern.

Aufbau:

Jede Zeile entspricht einem Datensatz; Spalten enthalten Attribute.
Wenn Spalten für Koordinaten (x, y, z) vorhanden sind, können GIS-Programme daraus Punkte erzeugen.

Vorteile:

- Sehr einfach, klein und universell lesbar
- Kompatibel mit Excel, QGIS, Python etc.
- Ideal für Datenaustausch und Attributtabellen

Nachteile:

- Keine native Geometrie- oder Projektionseigenschaften
- Nur für Punktdaten geeignet (keine Linien/Polygone)
- Fehleranfällig bei falscher Trennung oder Kodierung



	<u>Typ</u>	<u>Struktur</u>	<u>Hauptvorteil</u>	<u>Hauptnachteil</u>	<u>Typische Anwendung</u>
Shapefile	Binär	Mehrere Dateien	Sehr verbreitet	Sehr viele Einschränkungen, Nicht zeitgemäß	Klassische GIS-Projekte
GeoJSON	Text	Eine Datei	Lesbar & Web-freundlich	Dateigröße	Web-GIS, Datenaustausch
GeoPackage	SQLite DB	Eine Datei	Modern & vielseitig	Komplexer Aufbau	Moderne GIS-Systeme
CSV	Text	Tabellarisch	Einfach & universell	Keine echte Geometrie	Punktdaten, Tabellenimport

Übersicht der häufigsten GIS-Datenformate



GML

Offener XML-basierter OGC-Standard zur Beschreibung und Austausch von geografischen Informationen. Es wird vor allem in **Verwaltungen, Kataster- und INSPIRE-Projekten** verwendet.

Aufbau:

GML-Dateien bestehen aus **XML-Elementen**, die Geometrien (z. B. `<gml:Point>`, `<gml:Polygon>`) und zugehörige Attribute beschreiben. Sie sind **textbasiert**, menschenlesbar und maschinenverarbeitbar.

Vorteile:

- Offener und standardisierter OGC-Standard
- Unterstützt komplexe Geometrien, 3D und Metadaten
- Ideal für Datenaustausch in Behördenumgebungen (INSPIRE)

Nachteile:

- Sehr **umfangreich und komplex** in der Struktur
- Große Dateigrößen
- Nicht ideal für schnelle Visualisierung oder große Datenmengen
- **Langsam in Verarbeitung** gegenüber binären Formaten

<GML>

```
<gml:Point gml:id="p21" srsName="http://www.ageo.at">
  <gml:coordinates>45.67, 88.56</gml:coordinates>
</gml:Point>
```



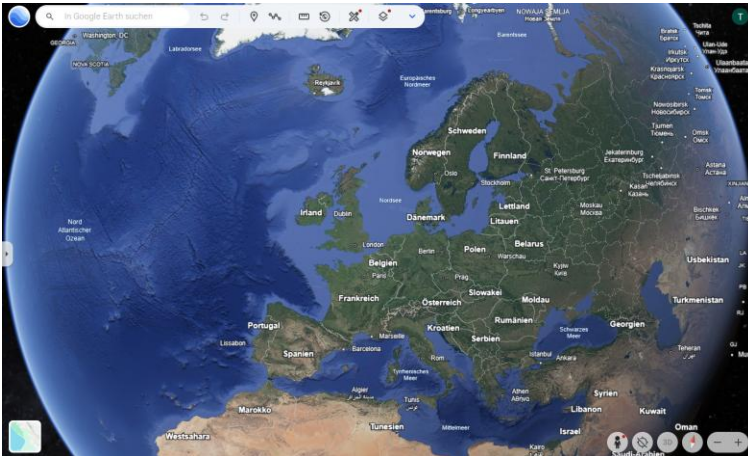
	<u>GML</u>	<u>GeoJSON</u>
Entwicklung / Standard	Vom OGC entwickelt, offizieller Standard für Datenaustausch in Behörden und INSPIRE	Entwickelt aus JSON , weit verbreitet in Web- und Open-Data-Anwendungen
Datenstruktur	XML-basiert (verschachtelte Tags, umfangreich)	JSON-basiert (einfach, klar strukturiert)
Lesbarkeit / Komplexität	Komplex, schwer lesbar für Menschen	Leicht verständlich und manuell lesbar
Dateninhalt	Sehr detailliert. Unterstützt komplexe Geometrien, Metadaten, Relationen, 3D-Modelle	Einfacher – Fokus auf Geometrien (Punkte, Linien, Polygone) und Attribute
Dateigröße / Effizienz	Groß, da XML-Overhead	Kleiner und effizienter
Kompatibilität	Gut in Behörden- und Desktop-GIS (z. B. ArcGIS, QGIS)	Ideal für WebGIS, APIs, mobile Anwendungen
Kodierung	UTF-8, stark strukturiert (XML-Schema)	UTF-8, flexibel und leicht validierbar
Anwendung	Austausch und Speicherung komplexer, normierter Geodaten	Schneller, einfacher Datenaustausch im Web

GML vs. GeoJSON



	Typische Nutzung
GML	Amtliche Geodaten (z. B. Kataster, INSPIRE-Daten, Verwaltungsportale) - Austausch zwischen Behörden und GIS-Systemen - Langfristige, standardisierte Datenspeicherung stark, normiert und präzise, aber komplex → ideal für Behörden, Kataster, INSPIRE
GeoJSON	Webkarten & interaktive Anwendungen (Leaflet, Mapbox, OpenLayers) - Mobile GIS-Apps - Open-Data-Portale - Datenaustausch zwischen Webdiensten leicht, schnell und modern → ideal für Webkarten, Open Data, interaktive Anwendungen

GML vs. GeoJSON



KML

Ein **XML-basiertes Format**, ursprünglich von **Keyhole Inc.** entwickelt (später von Google übernommen). Es ist das **Standardformat von Google Earth** zur Visualisierung räumlicher Daten.

Aufbau:

- .kml → unkomprimierte Textdatei.
- .kmz → komprimierte eingebetteten Bildern oder Symbolen

Vorteile:

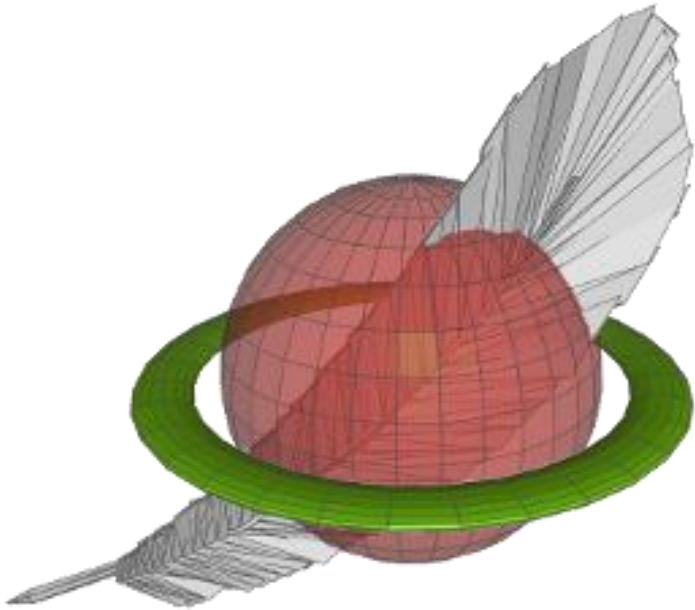
- Einfach zu visualisieren (z. B. in Google Earth, Google Maps)
- Lesbar und leicht zu verstehen
- Kann Multimedia-Elemente enthalten (Icons, Fotos, Hyperlinks)

Nachteile:

- Wenig geeignet für GIS-Analysen (mehr Präsentationsformat)
- Genauigkeit bei großen Datensätzen
- Teilweise proprietäre Erweiterungen von Google

KML = Anzeige & Visualisierung

GML = Austausch & semantische Genauigkeit



SpatialLite

Räumliche Erweiterung von SQLite, einer kompakten, dateibasierten Datenbank. Ermöglicht, **räumliche Datenbankfunktionen** (ähnlich wie PostGIS, aber ohne Server) in einer einzigen Datei zu nutzen.

Aufbau:

- Eine .sqlite-Datei enthält Tabellen für Geometrien, Attribute und räumliche Indizes.
- Unterstützt SQL-Abfragen, Layerverwaltung und OGC-konforme Geometrien.

Vorteile:

- Alles in einer Datei (ähnlich wie GeoPackage)
- Unterstützt SQL, Indizes, räumliche Abfragen
- Kompakt und effizient
- Offenes Format, leicht mit GIS- und Datenbanktools nutzbar

Nachteile:

- Weniger verbreitet als GeoPackage
- Etwas komplexer für Einsteiger
- Teilweise Inkompatibilitäten zwischen GIS-Programmen
- Keine native Rasterunterstützung



OGC GeoPackage	• Offener OGC-Standard – weltweit anerkannt und dokumentiert • Ist heute das Standardformat für Austausch, Speicherung und Nutzung in praktisch allen GIS-Umgebungen. • Ideal für Datenaustausch und Archivierung • Eingeschränkte SQL-Funktionalität
SpatialLite	• Nach wie vor im Einsatz für lokale, SQL-basierte Analysen • Räumliche Indizes und komplexe SQL-Abfragen möglich • wenn man volle Kontrolle über die Datenbank braucht (PostGIS light) • Entwicklung und Prototyping • Keine Rasterdaten-Unterstützung

GeoPackage vs. SpatialLite



Signierung von GIS-Datenformaten



<u>Format</u>	<u>Eignung</u>	<u>Begründung</u>
Shapefile	✗ Sehr schlecht geeignet	Besteht aus mehreren Dateien (.shp, .dbf, .shx etc.). Keine Möglichkeit, das Gesamtdataset konsistent zu signieren. Hohe Fehleranfälligkeit bei Einzeldateien.
GeoPackage	✓ Gut geeignet	Da auf SQLite basierend, kann die gesamte Datei über digitale Signaturverfahren (z. B. CMS, PGP oder Zertifikats-Hash) gesichert werden. OGC arbeitet zudem an Spezifikationen für authentifizierte GeoPackages.
GeoJSON	⚠ Teilweise geeignet	JSON kann über JSON Web Signature (JWS) signiert werden, aber dafür braucht es separate Standards. Noch keine weit verbreitete Implementierung im GIS-Kontext.
GML	✓ Sehr gut geeignet	XML-Struktur ermöglicht digitale Signaturen nach XAdES-Standard (XML Digital Signature). Wird bereits in INSPIRE, AFIS-ALKIS-ATKIS, E-Government-Diensten
CSV	⚠ Begrenzt geeignet	Kann signiert werden (z. B. über Prüfsummen oder digitale Signaturen), aber keine native GIS-Struktur oder räumliche Integrität – also kein echtes GIS-Format.

GIS-Formate und Integritätsprüfung (Signierung)



<u>Format</u>	<u>Rang</u>	<u>Begründung</u>
GML		<u>Sehr Hoch:</u> XML-basierte Standards unterstützen eingebettete digitale Signaturen (XAdES), bereits in amtlichen Prozessen etabliert. <u>Am besten geeignet, insb. für amtliche, standardisierte Anwendungen (mit XML-Signaturen).</u> <u>„eGovernment-konform, behördlich, standardisiert“</u>
GeoPackage		<u>Hoch:</u> Ganze Datei kann sicher signiert werden; OGC plant Signatur- und Hash-Erweiterungen. Ideal für moderne, portable Workflows. <u>Geeignet für für moderne, portable, dateibasierte GIS-Workflows (Signatur der gesamten Datei).</u> <u>„modern, effizient, universell portabel“</u>
GeoJSON		<u>Mittel:</u> Technisch möglich (JWS), aber noch kein Standard in GIS-Systemen. <u>Zukunftspotenzial für WebGIS-Signaturen.</u>
Shapefile		<u>Ungeeignet:</u> Keine integrierten oder etablierten Mechanismen für kryptografische Signaturen. Multifile-Format.

GIS-Formate und Integritätsprüfung (Signierung)



Thomas Galka
Technologie Manager | Senior Consultant

thomas.galka@axmann.at

axmann geoinformation

Modecenterstraße 22 / Top D59-D61
1030 Wien

Tel.: +43 1 203 91 47
office@axmann.at
www.axmann.at