FOSSGIS MÜNSTER 2025

26.-29. März 2025

TAGUNGSBAND

Anwenderkonferenz für Freie und Open Source Software für Geoinformationssysteme, Open Data und OpenStreetMap



Platinsponsoren

Goldsponsoren

Silbersponsoren

Bronzesponsoren

Medienpartner

ISBN 978-3-00-081664-2
DOI: 10.5281/zenodo.14711640



Anwenderkonferenz für Freie und Open Source Software für Geoinformationssysteme, Open Data und OpenStreetMap

Herausgeber:

FOSSGIS e.V.
Bundesallee 23
10717 Berlin

http://www.fossgis.de

ISBN 978-3-00-081664-2

DOI: 10.5281/zenodo.14711640

Bild Titelseite: DOP Münster created by mundialis

Gestaltung Titelseite: André De Wit

Inhaltsverzeichnis

Eine Reise durch die Geoportale Deutschlands 9
spatial.IO - Cloud-basierte Open-Source-Lösung zur Verwaltung räumlicher Daten 10
GeoPandas - als Tool zur Basiskartenaktualisierung 11
Keine Angst vor der GeoInfoDok 7 - 3A-Datenverarbeitung mit PostNAS 12
GIS-Schulprojekte in Zusammenarbeit mit kommunalen Gebietskörperschaften 19
eoAPI - eine skalierbare Geodateninfrastruktur 22
aviary - ein generisches Python-Framework zur KI-Inferenz für Fernerkundungsdaten 23
GDI per Knopfdruck: Automatisierung mit DevOps und Infrastruktur als Code 25
basemap.de als Open Data - Neue Stile und Anwendungsbeispiele 26
Automatischer Import und Veröffentlichung von Betriebsmittelgeometrien mittels PyQGIS 27
Superset - Business Intelligence meets Cartography 32
FOSS-basierte Schnittstelle zum Management von Heritage BIM Modellen 32
Amtliche Orthofotos zentral verfügbar: Eine Open-Source-Lösung für Deutschland 33
Open Data in D: Perfekte Idee, halbherzige Umsetzung? Ein Erfahrungsbericht. 34
QGIS-Werkzeuge und Python 35
Visualisierung von historischen Schiffsrouten mit unscharfer Datengrundlage 36
Wie MapLibre und Vektorkarten die Welt übernehmen 39
Open Data des BKG (II) 39
Schnupperkurs: Das Potential von QGIS mit der Python-Konsole freischalten 39
Automatisierte Verarbeitung von Daten der Meeresbodenkartografie mit QGIS 40
Vektor Tiles für Karten mit Echtzeitdaten 45
Small seeds - FOSS Communities stärken! 46
Von Proprietär zu QGIS 47
Open Source and Commercial Large Language Models comparison: 48
Open ALKIS? – Oder was passiert, wenn der deutsche Föderalismus auf EU-Recht trifft 55
Governance von Open-Source-Software im öffentlichen Sektor: Make, Buy or Contribute? 66
Anhang 76
Von proprietär zu Open-Source - Umstellung der kommunalen GDI bei der Stadt Reutlingen 78
Entwicklung eines LLM-basierten Assistenten für die Suche nach Geodaten 81
Was wäre, wenn wir Algorithmen demokratisieren? Neue kollaborative Infrastrukturen für Modelle und Simulationen in digitalen Stadtzwillingen 85
Migration eines Auskunftssystems zu einer Open-Source Lösung mit QGIS 93
Künstliche Intelligenz als Unterstützung in geografische Applikationen 94
Metadaten für eine verantwortungsvolle und kritische Geodatenpraxis 95
Das Projekt GEOrg bei den SWM: Mit Freier Software zur konzernweiten Geodatenplattform 110
Open Data und KI im Einsatz: Geodaten für alle, nicht nur für Profis? 112
MOSIDI - Homogenisierung von offenen Daten für die kommunale Planung 113
Der Digitale Zwilling - so wertvoll wie eine Karte im Maßstab 1:1 119
Routing, aber mehr explorativ statt automatisch 120
Anreicherung von Straßendaten mittels Deep-Learning-Methoden und Mapillary Bildern 121
G3W-SUITE: das OS-Framework für die Veröffentlichung und Verwaltung von QGIS-Projekte 123

Der Wuppertaler Weg vom Geoportal zum Digitalen Zwilling....124
Flexibles Open Source Routing mit Valhalla....124
Versiegelungsanalyse zur bioklimatischen Bewertung von Stadtgebieten....125
Historische Karten mit QGIS georeferenzieren....126
Masterportal - Liegenschaftsauskunft ONLIKA 4.0 mit Keycloak und BundID....127
Transitous - Freies Public Transport Routing....130
Ermittlung von Flächenversiegelung als komplexe Aufgabe (Projekt SEAL)....131
FOSSGIS bei OpenCode.de....136
Barrierefreies Routing mit MOTIS....137
Verarbeitung offener Satellitendaten mit freier Software für die visuelle Anwendung....137
10 Jahre openSenseMap: Neue Features für die Zukunft der Umweltdaten....138
Datenklassifikation: Mehrwert oder Manipulation?....139
Verkehrliche Maßnahmen aus historischen OSM-Daten identifizieren....147
Erfahrungsbericht: MerginMaps für die Inventarisierung von Vogelnestern....149
OpenStreetMap FeatureInfo....150
Pünktlich zur Uni? Bus-Tracking mit der EFA JSON API....150
2.5D Indoor-Karten auf Basis von OpenStreetMap-Daten....151
Sheepsmeadow: Agentenbasierte Modellierung einfach erlernen....157
XPlanung mit Open Source Software....159
Online-Karten für die Verkehrswende mit OpenData und FOSS....168
Mit EnMAP-Box und OSM durch Namibia: ein hyperspektraler Praxisbericht....168
Digitaler Zwilling Niedersachsen auf Basis der Unreal Engine und LGLN Open Data....169
Artificial Ground Truth Data Generation for Map Matching with Open Source Software....170
Efficient processing of high-volume spatial data with Spark....187
Scalable big data processing – Postgres' drawbacks eliminated with Spark....187
Modellgetriebene XPlanung: von UML zur OGC API for Features und mehr....188
Erreichbarkeitsanalyse in ländlichen Räumen: Potenziale des Radverkehrs....192
Klimaanpassungsrelevante Vegetationsindikatoren mit Hilfe von Sentinel-2 Zeitreihen....194
SidescanTools - open source Sidescan Processing Software....195
QFieldCloud Erweitern - Ideen und Praxisbeispiele....195
Eine automatisierte FOSS-GDI zur Exploration von Erdsystem-Forschungsdaten....196
An Inventory of Spatial Machine Learning Packages in R....197
OpenStreetMap-Daten zur Geolokalisierung unbemannter Luftfahrzeuge....198
(k)NNDM - Kreuzvalidierung im geographischen und Prädiktorraum....199
UmweltNAVI: Eine interaktive und kooperative Plattform für Umweltinformationen....200
Offene Daten für barrierefreie Mobilität und inklusive Reiseketten im ÖPV....202
Einsatz für die Energiewende....203
Einbindung von OGC APIs in Webtools für Wärmestromdaten....203
Offene Daten für barrierefreie Mobilität und inklusive Reiseketten im ÖPV....204
Kinder, Karten, Open Source....207
Fahrzeugortung DB Regio – mehr als GPS....209
Landesweite Datenerfassungen organisieren und effizient gestalten mit QGIS....214

QGIS fernsteuern: Kickstart mit PyQGIS zur Automatisierung....216
QGIS Server per REST API konfigurieren....220
OpenStreetMap ist doch vollständig ….....220
DFS Deutsche Flugsicherung: Open Source und Sicherheit bei UAS? Ein Erfahrungsbericht....221
QGIS im Glasfaser-Ausbau der Deutschen Telekom....222
Zaubern mit dem Geopackage....223
Overpass Turbo goes PostGIS....223
Geodatenmanagement in einer Baubehörde....224
Web trifft Desktop....231
Automatische Veredelung von offenen Nahe-Echtzeit-Daten....232
QField: Neue Strategie und Anwendungspotentiale....233
Wie können OpenStreetMap und QGIS einen Wegewart unterstützen?....234
Leerstandsmelder & A Thousand Channels - Countermappings aus der Zivilgesellschaft....236
Amtliche Satellitenpositionierung 2.0....237
SMASH, Stand der Technik der digitalen Feldkartierungs-app....240
Parkraumdaten aus OSM mit der Verwaltung pflegen (Praxisbericht)....241
Wie mache ich eine gute Karte mit QGIS?....242
Neue Methoden zur autarken Indoor-Positionsbestimmung....242
OGIS in der Öffentlichen Verwaltung....243
MerginMaps - Mobile Geodatenerfassung beim RVR....244
BRouter und BRouter-Web für Anfänger und Fortgeschrittene....248
FOSS-GIS+FOSS in der Wind- und Solarbranche (Praxisbericht)....249
QWC2 als WebGIS in der kommunalen Verwaltung und Einblick in die 3D Funktionalität....250
Nicht von neuen Algorithmen überflutet: OSM-Daten für WaterwayMap verarbeiten....251
Unterstützung des regionalen Naturschutzes in Luxemburg mit Open Source GIS-Software....251
State of SHOGun: Ein flexibles Web-GIS Framework....252
OSM2World: Updates für den 3D-Pionier....252
Kartografie-Verbesserungen, Tips und Tricks in QGIS....253
Mapbender - Neuigkeiten aus dem Projekt....254
OpenLayers - Neues und Nützliches....254
Mit QGIS zum digitalen Prozess 1 - Gebäudeeigenschaften....255
GeoStyler - ArcGIS Unterstützung und weitere neue Features....258
Text und Data Mining in der OpenStreetMap-Datenbank aus rechtlicher Sicht Text....259
State of GeoServer Cloud....265
Neues von actinia - mit Knative weiter wolkenwärts....265
Stand des GRASS GIS Projekts: Neuigkeiten und Perspektiven....266
Fehlende Bürgersteige....267
Ableitung von Rasterkarten aus Vector Tiles für basemap.de....267
QGIS und Blender für beeindruckende 3D-Darstellungen....268
Der Elefant kann's auch allein: Graph-Erstellung aus OSM in der PostGIS-Datenbank....268
radinfra.de – gute Daten und Kampagnen zur Radinfrastruktur für ganz Deutschland....269
PMTiles - das cloud-native Format für Kacheln....270

Airborne-Laserscanning in Deutschland: Verfügbarkeiten, Herausforderungen, Potenziale....270
OpenSource sicher entwickeln und betreiben – Prozesse, Anforderung und Tools im Fokus....271
PTNA: Qualitätssicherung für ÖPNV-Linien in OpenStreetMap....272
RoutingPlus goes Masterportal: Unser Weg zu einer modernen Benutzeroberfläche....289
Römische Grabdenkmäler im 3D-Webviewer....290
fiboa - Standardisierte Feldgrenzen....293
Absicherung von Diensten mit Keycloak....294
Datenbankschema mit Mermaid visualisieren....294
Koordinatenreferenzsysteme für D-A-CH....294
Lebewohl Web Merkator....295

Editorial

Was ist die FOSSGIS-Konferenz?

Die FOSSGIS-Konferenz ist im D-A-CH-Raum die führende Konferenz für Freie und Open Source Software für Geoinformationssysteme sowie für die Themen Open Data und OpenStreetMap. Die FOSSGIS-Konferenz 2025 wird vom gemeinnützigen FOSSGIS e.V und der OpenStreetMap Community in Kooperation mit dem Institut für Geoinformatik der Universität Münster organisiert und veranstaltet.

Ziel der jährlich stattfindenden Konferenz ist die Verbreitung von Freier und Open Source Software für Geoinformationssysteme. Hier treffen sich Anwender:innen, Entwickler:innen und Interessierte zum gemeinsamen Austausch über Anwendungs- und Arbeitsmöglichkeiten sowie neuste Entwicklungen in diesem Bereich. Es geht um Webmapping, Geodateninfrastrukturen, Geodatenmanagement, Geoprocessing, Geodatenanalysen und -modellierungen und Desktop-GIS und um freie Geodaten, wie OpenStreetMap. Themen wie Digitale Souveränität im Zusammenhang mit Open Source werden ebenso behandelt, wie Open Science und die FAIR Prinzipien.

Die Konferenz wird vorwiegend durch ehrenamtliches Engagement getragen, wobei neben den Teilnehmergebühren, die Spenden- und vor allem die Sponsoren- und Ausstellerstandeinnahmen eine wichtige finanzielle Einnahmequelle darstellen, um die Kosten für die Konferenz zu decken und darüber hinaus Open Source Projekte zu fördern.

Was finden Sie in diesem Tagungsband?

Sie finden die Beiträge zum Programm der FOSSGIS 2025. Einreichungen zum Tagungsband verfügen über die Angabe Verfasser:innen-Angabe unter dem Titel und die Kontaktdaten am Ende des Beitrags. Die Beiträge, die einen Beschreibungstext enthalten und der Name der Verfasser:in unter dem Beitrag steht, sind am 24.02.2025 aus dem Vortragsverwaltungsprogramm Pretalx exportiert worden.

Unter welcher Lizenz stehen die Beiträge?

Die Konferenzbeiträge stehen unter der Lizenz CC BY 4.0 zur Verfügung. Der Konferenzbeitrag umfasst die Kurzbeschreibung des Beitrags, den Aufsatz im Konferenzband, die Vortragsfolien sowie die Videoaufzeichnung und Liveübertragung des Vortrags.

Eine Reise durch die Geoportale Deutschlands

Matthias Mohr

Im letzten Jahr musste ich Datensätze für Feldgrenzen in Deutschland finden. Was erst einmal einfach klingt endete mit einer Odyssee durch die verschiedensten Geoportale der Länder. Ich nehme euch auf dieser Reise mit und beschreibe, was mir bei der Reise aufgefallen ist und wie man diese Erfahrung in der Praxis verbessern könnte.

Wir schauen, welche Länder die Daten verfügbar haben, wie man diese Daten erkunden kann, und vor welchen Problemen man im Endeffekt als Nutzer steht. Haltet euch fest und macht euch bereit für einen wilden Ritt…

spatial.IO - Cloud-basierte Open-Source-Lösung zur Verwaltung räumlicher Daten

Rebekka Lange[1]

[1] Helmholtz-Zentrum für Umweltforschung (UFZ)

Zusammenfassung: spatial.IO ist eine cloud-basierte Open-Source-Lösung zum Speichern, Verarbeiten, Verwalten, Veröffentlichen und Visualisieren von räumlichen Daten, die OGC Schnittstellen und ein individualisierbares WebGIS bietet.

Schlüsselwörter: Geodateninfrastruktur, Open-Source, WebGIS

Mit zunehmenden Datenmengen im Bereich der Umweltsystemforschung steigt die Notwendigkeit für geeignete Dateninfrastrukturen zur automatisierten Verwaltung und Repräsentation der Daten. Damit Wissenschaffende weiter Forschung betreiben können ohne sich mit dem Datenchaos herumschlagen zu müssen, sind standardisierte Lösungen zur Speicherung, Prozessierung, Veröffentlichung und Darstellung von räumlichen Daten notwendig.

Abb 1: spatial.IO Logo

Das Helmholtz-Zentrum für Umweltforschung (UFZ) hat spatial.IO entwickelt, eine modulare Anwendung zur Verwaltung, Speicherung, Verfügbarstellung und Visualisierung von räumlichen Daten. spatial.IO kann unabhängig von bestehenden IT-Infrastrukturen überall betrieben werden. Die Anwendung beinhaltet einen integrierten Cloud-Speicher, GeoServer, GeoNetwork, FROST-Server und ein individuell konfigurierbares WebGIS. Alle Komponenten sind über automatisierte Prozesse verknüpft und eine Web-Anwendung bietet eine nutzenden-freundliche Verwaltung.

spatial.IO Key Features:

- modularer Aufbau mit Docker
- zentraler Cloud-Speicher (S3)
- Admin-Frontend zur einfachen Verwaltung aller Daten
- FROST-Server zur Speicherung von Sensor-Daten und der Schnittstelle SensorThings API (STA)
- GeoServer zur Bereistellung von OGC Web Services
- GeoNetwork als Metadatenkatalog
- WebGIS-Template zur einfachen Erstellung von eigenen Viewern
- automatisierte Prozesse zwischen allen Komponenten

Kontakt zum Autor:

Rebekka Lange
Helmholtz-Zentrum für Umweltforschung UFZ
Permoserstraße 15, 04318 Leipzig
rebekka.lange@ufz.de

Abb 2: spatial.IO Veröffentlichung

GeoPandas - als Tool zur Basiskartenaktualisierung

Markus Gruber, Markus Albrecht[1]

[1] Stadtwerke München GmbH (SWM)

Zusammenfassung:

Die Basiskarte (GEOBASISDATEN) der Stadtwerke München ist eine Vermengung von verschiedensten Datenquellen, welche den Bedürfnissen des Netzinformationssystems entsprechen. Hauptsächlich besteht die Basiskarte aus amtlichen Daten (ALKIS, DFK, …) vom Staatlichen Vermessungsamt, aus Zulieferer Daten von Fremdfirmen, OpenData oder selbst erfassten Daten der Fortführung und Planung der Stadtwerke München.

Die Daten werden in unterschiedlichsten Formaten geliefert (XML, SHAPE, CSV) und werden erst in vorgelagerten PostgreSQL Schema und von diesen in das PostgreSQL Ziel Schema BK importiert.

Die Basiskarte der SWM beinhaltet eine Fläche von ca. 17.000 qkm, was einer Fläche von Thüringen entspricht. Das Datenmodell wurde für den Anforderungen der SWM entwickelt, es enthält 67 Tabellen und ca. 30 Mio. Objekte. Sie ist die Datenbasis im Netzinformationssystem, in QGIS-Anwendungen oder in Java REST Schnittstellen wie der Adress-Service.

Der Vollimport wurde in der Vergangenheit mit verschiedensten ETL-Tools aus dem Open Source oder kommerziellen Bereich mehr oder weniger zufriedenstellend durchgeführt. Anfang 2024 erfolgte in nur drei Monaten die Migration der bestehenden 180 ETL-Transformationen vom Kaufprodukt hin zur Open Source Python / GeoPandas Lösung, mit dem Ziel, wesentlich mehr Flexibilität und Performance zu gewinnen.

Die Transformationen müssen in einer komplexen Prozesskette seriell, oder wenn möglich, parallel ausgeführt werden, so dass ein performanter Import mit allen Abhängigkeiten und Prioritäten erfolgreich durchgeführt werden kann. Mittels der der Open Source Python-Bibliothek Dask werden diese Aufgaben für paralleles und verteiltes Rechnen sichergestellt. Dask ist einfach zu verwenden und im Python Code einzubinden.

Die Ausführung der Transformationen erfolgt auf dem Open-Source Kubernetes Cluster, welcher on premise bei den SWM betrieben wird. Aufgrund der Skalierbarkeit ist es uns gelungen den bisherigen Import Prozess der Basiskarte von mehreren Tagen auf 6 Stunden zu reduzieren. Aufgrund der hohen Flexibilität von GeoPandas können jetzt neue Anforderungen im Bereich von Geo-Daten Schnittstellen oder Auswertungen problemlos implementiert werden.

Schlüsselwörter: Basiskarte, Geobasisdaten, ALKIS, Python, Jupyter Notebooks, GeoPandas, Dask, Kubernetes

Keine Angst vor der GeoInfoDok 7 - 3A-Datenverarbeitung mit PostNAS

Oliver Schmidt

Zusammenfassung: Für viele Anwendungen müssen 3A-Daten im Format der GeoInfoDok 7.1.2 verarbeitet werden. Vor dem Hintergrund der OpenData-Bereitstellung des LVermGeo Rheinland-Pfalz steigt nun auch das Interesse an originären 3A-Daten und an AdV-konformen WFS und WMS. Import- und Verarbeitungsschritte, mit denen 3A-Daten in eine PostGIS-Datenbank überführt werden, sind Teil dieses Vortrages. Ebenso werden eigene Skripte und hieraus erzeugte Geowebdienste vorgestellt.

Schlüsselwörter: PostNAS, 3A-Daten, ALKIS, Liegenschaftskarte, AdV

Ausgangspunkt

Das Landesamt für Vermessung und Geobasisinformation Rheinland-Pfalz (LVermGeo RP) stellt eine Vielzahl seiner Geobasisdaten mittels Web Map Services (WMS), Web Feature Services (WFS) und Web Coverage Services (WCS) bereit. [1] Zur Erstellung dieser Dienste werden AAA- und weitere Quelldaten verwendet, die entweder in einem verwaltungsinternen oder einem bundeseinheitlichen Datenformat vorliegen. Die Bereitstellung erfolgt mittels MapServer 8 und deegree 3. Der Gesamtaufbau des komplett auf Open-Source-basierten Serververbundes wurde im Rahmen der FOSSGIS-Konferenz 2021 vorgestellt [2].

3A-Daten der Vermessungsverwaltungen lassen sich entweder mit proprietärer Software verarbeiten oder man entscheidet sich für die Opensource-Suite „PostNAS“. Mit PostNAS werden die 3A-Daten in der Version 7.1.2 der GeoInfoDok [3] mittels GDAL/OGR und einem hierfür entwickelten OGR-Treiber in eine PostGIS-Datenbank überführt. Zur weiteren Verarbeitung und anschließenden Präsentation als AdV-konformer ALKIS-WFS oder zur Darstellung einer Liegenschaftskarte sind zahlreiche weitere Schritte nötig. Die hier gezeigten Skripte und Bespiele basieren allesamt auf der Verwendung von Debian 12, PostgreSQL 15, PostGIS 3.3, GDAL/OGR 3.8 und MapServer 8.0.

Einrichtung des ALKIS-Imports

„PostNAS“ beschreibt kein eigenständiges Programm – vielmehr handelt es sich um eine geschickte Verkettung von GDAL/OGR, Import- und Verarbeitungsskripten und PostGIS-Funktionen. Ferner gibt es ein QGIS-Plugin und eine ALKIS-Auskunft, die mittels Mapbender realisiert wird. All diese Komponenten werden unter dem Namen „PostNAS-Suite“ [4] zusammengefasst. Der Import und die Verarbeitung wird mit dem „norBIT ALKIS-Import“ durchgeführt. In diesem Vortrag liegt der Fokus auf der Verarbeitung von Daten im Format der GeoInfoDok 7.1.2 und von differenziellen NAS-Daten. Daher findet der Branch „gid7i“ [5] des ALKIS-Import Verwendung, um diese Möglichkeiten zu nutzen.

Mit der Umstellung von GeoInfoDok 6.0.1 auf die Version 7.1.2 ergaben sich Änderungen, die bei einigen Objektarten gravierende Auswirkungen hatten. Es gab auch Änderungen bei Katalogobjekten oder nur minimale Anpassungen an der Benennung von Attributen. Dennoch sind die Änderungen in Summe ausreichend vielfältig, um eine entsprechende Anpassung durchzuführen. Hierzu mussten nicht nur Import- und Verarbeitungsskripte angepasst werden, sondern auch der entsprechende GDAL/OGR-Treiber. Erst ab der GDAL-Version 3.8.0 wird der Import der NAS-Daten in der GeoInfoDok-Version 7.1 unterstützt. Bevor der Import erstmalig durchgeführt werden kann, müssen im Stammverzeichnis des geklonten Repositories zwei Dateien noch erstellt werden: `make tables.lst alkis-functions.sql`

Es ist empfehlenswert, ein kleines Import-Skript zu schreiben, welches die folgenden Vorbereitungen für einen Voll- oder Differenzimport import durchführt:

- Vorbereiten der Daten (z.B. Ordner erstellen, Kopieren, Entpacken)
- Erstellung der Steuerdatei `import.lst`
- Ggf. Erstellung einer leeren PostGIS-Datenbank
- Aufruf des norBIT ALKIS-Import
- Ggf. Aufruf weiterer Verarbeitungsskripte im Anschluss

Die Steuerdatei import.lst sorgt für die korrekte Verarbeitung der Daten. Diese unterscheidet sich je nach Anwendungsfall und zu importierenden Daten [6]. In Listing 1 wird ein Beispiel für einen Erstimport gezeigt, Listing 2 ermöglicht einen Differenzimport.

Listing 1: import.lst für den Erstimport

```
PG:dbname=alkis user=import password=xxx
clean
create
log import.log
options -skipfailures -gt 65000
/daten/import/alkis/ALKIS_20250111T000000Z_0000von1240_32290000_5420000.xml.gz
/daten/import/alkis/ALKIS_20250111T000000Z_0001von1240_32435000_5420000.xml.gz
/daten/import/alkis/ALKIS_20250111T000000Z_0002von1240_32440000_5420000.xml.gz
/daten/import/alkis/ALKIS_20250111T000000Z_0003von1240_32420000_5425000.xml.gz
[...]
/daten/import/alkis/ALKIS_20250111T000000Z_1239von1240_32290000_5420000.xml.gz
/daten/import/alkis/ALKIS_20250111T000000Z_1240von1240_32290000_5420000.xml.gz
```

Listing 2: import.lst für den Import von Differenzdaten ohne Historie

```
PG:dbname=alkis user=import password=xxx
update
historie off
log import.log
options -skipfailures -gt 65000
/daten/import/alkis/ALKIS_folge_20250117T000000Z_000von666_32295000_5425000.xml.gz
/daten/import/alkis/ALKIS_folge_20250117T000000Z_001von666_32440000_5425000.xml.gz
/daten/import/alkis/ALKIS_folge_20250117T000000Z_002von666_32445000_5425000.xml.gz
[...]
/daten/import/alkis/ALKIS_folge_20250117T000000Z_665von666_32410000_5635000.xml.gz
/daten/import/alkis/ALKIS_folge_20250117T000000Z_666von666_32295000_5425000.xml.gz
```

Verarbeitungsskripte des ALKIS-Imports

Nach dem Import laufen zahlreiche Verarbeitungsskripte (Postprocessing), die die Rohdaten für eine Präsentation in einer Liegenschaftskarte oder zur Beauskunftung aufbereiten. Ein Großteil dieser Skripte hatte für die Verarbeitung der rheinland-pfälzischen Daten keinen Anpassungsbedarf. Landesspezifische Ausprägungen wurden in den Skripten nachträglich hinzugefügt, da sie bisher nicht Bestandteil des norBIT ALKIS-Imports sind. Zudem wurde ein Skript (postprocessing.d/1_ableitungsregeln/11001.sql) in 6 einzelne Skripte aufgeteilt. Durch eine parallele Abarbeitung der Skripte kann eine schnelle Verarbeitung erreicht werden. Das erwähnte Skript weist in seiner ursprünglichen Form die längste Laufzeit auf, sodass durch die parallele Abarbeitung die Laufzeit insgesamt deutlich verkürzt werden konnte.

Hinzugefügt wurden Postprocessingskripte für die Objektarten „AX_AndereFestlegungNachWasserrecht“ (postprocessing.d/1_ableitungsregeln/71004.sql) und „AX_Schutzgebiet_NachWasserrecht“ (postprocessing.d/1_ableitungsregeln/71005.sql). Diese Objektarten werden in Rheinland-Pfalz geführt und fehlten bisher in der Verarbeitung.

Da auch die Bodenschätzung auf unserem System prozessiert und bereitgestellt werden soll, wurde ein Skript von Frank Jäger entsprechend angepasst. Da diese Berechnungen erst durchgeführt werden können, wenn andere Schritte durchgelaufen sind, findet dessen Abarbeitung nicht gemeinsam mit den anderen Objektarten statt. Derzeit befindet sich das Skript unter postprocessing.d/7_bodenschaetzung.sql und berücksichtigt die rheinland-pfälzischen Spezifika.

Berechnung der Flurstücksabschnitte

Die Informationen zum Flurstück spielen bei der Beauskunftung eine wichtige Rolle. Hierin werden die tatsächliche Nutzung, die Bewertung, die öffentlich-rechtlichen Festlegungen und die Bodenschätzung mit den Flurstücksgeometrien verschnitten. Diese komplexe Aufgabe wurde mittels eines zusätzlichen Verabeitungsschrittes außerhalb des ALKIS-Imports umgesetzt. Die Umsetzung des „FAF“ (Flächen-Abschnitte der Flurstücke) wurde bereits unter GeoInfoDok 6.0.1 vom Landesamt für Vermessung und Geobasisinformation beauftragt und eingesetzt. Die Änderungen für die erfolgreiche Prozessierung der Daten der GeoInfoDok 7.1.2 beschränkten sich lediglich auf eine Steuerdatei (faf.ctl). Hierin mussten nur die Verschneidungsanweisungen aufgrund geänderter Tabellen- und Attributnamen berücksichtigt werden.

	Layer	FID	Attribut	Wert
1	ave:Flurstuecksabschnitte	15	flstkennz *	07119300500062000l__
2	ave:Flurstuecksabschnitte	15	status	A
3	ave:Flurstuecksabschnitte	15	flaeche	3875643
4	ave:Flurstuecksabschnitte	15	tng	3339966 m² See|109365 m² Grünland|50533 m² Grünland|45972 m² Laubholz|45440 m² Ackerland|45268 m² Gehölz|40927 m² Laubholz|28858 m² Campingplatz|27253 m² Laubholz|25197 m² Grünland|22404 m² Grünland|20529 m² Laubholz|16484 m² Gebäude- und Freifläche Land- und Forstwirtschaft|13417 m² Grünland|7804 m² Grünland|5478 m² Laubholz|4617 m² Gebäude- und Freifläche Sport, Freizeit und Erholung|3111 m² Laubholz|2864 m² Gehölz|2719 m² Gehölz|2485 m² Fußweg|2006 m² Laubholz|1836 m² Fahrweg|1815 m² Fußweg|1676 m² Gebäude- und Freifläche Sport, Freizeit und Erholung|1501 m² Fahrweg|1286 m² Graben|1122 m² Fußweg|994 m² Fahrweg|982 m² Campingplatz|699 m² Fahrweg|468 m² Gebäude- und Freifläche Versorgungsanlage, Elektrizität|362 m² Laubholz|204 m² Fahrweg
5	ave:Flurstuecksabschnitte	15	bod	10165 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (3), Grünlandgrundzahl 048, Grünlandzahl 042, Ertragsmesszahl 4269|10594 m² Ackerland (A), Bodenart Sandiger Lehm (sL), Zustandsstufe (4), Alluvium (Al), Bodenzahl 056, Ackerzahl 053, Ertragsmesszahl 5615|10945 m² Ackerland (A), Bodenart Sandiger Lehm (sL), Zustandsstufe (3), Alluvium (Al), Bodenzahl 064, Ackerzahl 064, Ertragsmesszahl 7005|12946 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 056, Grünlandzahl 052, Ertragsmesszahl 6732|1641 m² Grünland (Gr), Bodenart Moor (Mo), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 032, Grünlandzahl 030, Ertragsmesszahl 492|18872 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 048, Grünlandzahl 037, Ertragsmesszahl 6983|1911 m² Grünland (Gr), Bodenart Lehmiger Sand auf Moor (lS/Mo), Bodenstufe Misch- und Schichtböden sowie künstlich veränderte Böden (-), Klimastufe 8° C und darüber (a), Wasserstufe (3), Grünlandgrundzahl 038, Grünlandzahl 036, Ertragsmesszahl 688|21427 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 062, Grünlandzahl 061, Ertragsmesszahl 13070|217 m² Ackerland (A), Bodenart Stark lehmiger Sand (SL), Zustandsstufe (4), Diluvium (D), Bodenzahl 045, Ackerzahl 042, Ertragsmesszahl 91|2282 m² Grünland (Gr), Bodenart Moor (Mo), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (4), Streuwiese (Str), Grünlandgrundzahl , Grünlandzahl 012, Ertragsmesszahl 2742|548 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 062, Grünlandzahl 061, Ertragsmesszahl 1554|2954 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (3), Grünlandgrundzahl 048, Grünlandzahl 046, Ertragsmesszahl 1359|333 m² Grünland (Gr), Bodenart Moor (Mo), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (4), Streuwiese (Str), Grünlandgrundzahl , Grünlandzahl 012, Ertragsmesszahl 40|3414 m² Ackerland (A), Bodenart Stark lehmiger Sand (SL), Zustandsstufe (4), Diluvium (D), Bodenzahl 045, Ackerzahl 042, Ertragsmesszahl 1434|36716 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 062, Grünlandzahl 061, Ertragsmesszahl 2239|5467 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 058, Grünlandzahl 056, Ertragsmesszahl 3062|5716 m² Grünland (Gr), Bodenart Lehmiger Sand auf Moor (lS/Mo), Bodenstufe Misch- und Schichtböden sowie künstlich veränderte Böden (-), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 038, Grünlandzahl 036, Ertragsmesszahl 2058|5884 m² Ackerland (A), Bodenart Sandiger Lehm (sL), Zustandsstufe (3), Alluvium (Al), Bodenzahl 064, Ackerzahl 061, Ertragsmesszahl 3589|609 m² Grünland (Gr), Bodenart Moor (Mo), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (3), Streuwiese (Str), Grünlandgrundzahl , Grünlandzahl 008, Ertragsmesszahl 49|6840 m² Grünland (Gr), Bodenart Moor (Mo), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 032, Grünlandzahl 030, Ertragsmesszahl 2052|76198 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 056, Grünlandzahl 052, Ertragsmesszahl 3962|39036 m² Grünland (Gr), Bodenart Moor mit Lehm (MoL), Bodenstufe Misch- und Schichtböden sowie künstlich veränderte Böden (-), Klimastufe 8° C und darüber (a), Wasserstufe (3), Grünlandgrundzahl 034, Grünlandzahl 026, Ertragsmesszahl 2349|9785 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 056, Grünlandzahl 052, Ertragsmesszahl 5088|9935 m² Grünland (Gr), Bodenart Moor (Mo), Bodenstufe (III), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 032, Grünlandzahl 030, Ertragsmesszahl 2981|9987 m² Grünland (Gr), Bodenart Lehm (L), Bodenstufe (II), Klimastufe 8° C und darüber (a), Wasserstufe (2), Grünlandgrundzahl 050, Grünlandzahl 042, Ertragsmesszahl 4195
6	ave:Flurstuecksabschnitte	15	gmz	137049
7	ave:Flurstuecksabschnitte	15	bew	107203 m² Landwirtschaftliche Nutzung (LN)|49915 m² Landwirtschaftliche Nutzung (LN)|49595 m² Landwirtschaftliche Nutzung (LN)|45972 m² Forstwirtschaftliche Nutzung (H)|45266 m² Forstwirtschaftliche Nutzung (H)|40043 m² Forstwirtschaftliche Nutzung (H)|27422 m² Landwirtschaftliche Nutzung (LN)|27253 m² Forstwirtschaftliche Nutzung (H)|25567 m² Landwirtschaftliche Nutzung (LN)|20529 m² Forstwirtschaftliche Nutzung (H)|16484 m² Nutzungsart Hofstelle (NF)|12556 m² Landwirtschaftliche Nutzung (LN)|8217 m² Forstwirtschaftliche Nutzung (H)|4617 m² Nutzungsart Hofstelle (NF)|4160 m² Landwirtschaftliche Nutzung (LN)|3756 m² Nicht klassifiziert (NA)|3111 m² Forstwirtschaftliche Nutzung (H)|2719 m² Forstwirtschaftliche Nutzung (H)|905 m² Forstwirtschaftliche Nutzung (H)|772 m² Nicht klassifiziert (NA)|362 m² Forstwirtschaftliche Nutzung (H)|75 m² Nutzungsart Hofstelle (NF)
8	ave:Flurstuecksabschnitte	15	osf	3875643 m² Baulast|3875643 m² Benachteiligtes landwirtschaftliches Gebiet; Ministerium für Wirtschaft, Verkehr, Landwirtschaft und Weinbau|3875643 m² Flora-Fauna-Habitat-Gebiet; Struktur- und Genehmigungsdirektion Nord, Bezeichnung FFH-5509-301|380 m² Gemeindestraße|1056 m² Gemeindestraße|2244 m² Gewässer III. Ordnung|3339966 m² Gewässer II. Ordnung|3875643 m² Naturschutzgebiet; Struktur- und Genehmigungsdirektion Nord, Bezeichnung NSG-7131-006|3553874 m² Vogelschutzgebiet; Struktur- und Genehmigungsdirektion Nord, Bezeichnung DE-5509-401
9	ave:Flurstuecksabschnitte	15	osf_vkz	NULL
10	ave:Flurstuecksabschnitte	15	osf_stelle	Struktur- und Genehmigungsdirektion Nord|Struktur- und Genehmigungsdirektion Nord|Struktur- und Genehmigungsdirektion Nord|Ministerium für Wirtschaft|||
11	ave:Flurstuecksabschnitte	15	osf_bez	NSG-7131-006|DE-5509-401|FFH-5509-301||||
12	ave:Flurstuecksabschnitte	15	vorgaenger_flur...	NULL
13	ave:Flurstuecksabschnitte	15	nachfolger_flur...	NULL
14	ave:Flurstuecksabschnitte	15	gid	1601689

Abbildung 1: Ergebnis des FAF-Prozesses, Abschnittsflächen des Laacher Sees

Im Ergebnis liefert der FAF eine Punktgeometrie für jedes Flurstück zurück, an der zusätzliche Attribute anhängen. Diese Attribute weisen die gewünschten Inhalte zu den Abschnittsflächen der tatsächlichen Nutzung, der Bodenschätzung und Bewertung sowie der öffentlich-rechtlichen Festlegungen auf (s. Abbildung 1).

Erstellung von AdV-konformen und weiteren Produkten

Die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) stellt für konforme Produkte und Dienste konkrete Forderungen [7] auf, die zu erfüllen sind. Die Verarbeitung der Daten durch das Postprocessing des ALKIS-Imports oder den FAF-Prozess liefert keine AdV-konformen Datenstrukturen, auf denen ein ALKIS-WFS aufgesetzt werden könnte.

Die Verarbeitung der importierten ALKIS-Daten zur konformen Darstellung gemäß Signaturenkatalog erfolgt bereits im Postprocessing. Jedoch traten in der Anfangszeit hohe Serverlasten der Datenbankserver auf, die auf aufwändige SQL-Statements zurückzuführen waren. Es handelte sich um die Abfrage und Darstellung der Labels aus der Tabelle „po_labels“, die über das Mapfile unter Nutzung diverser Verschneidungen genutzt wurden. Da diese Tabelle in unserer Datenbank bis zum nächsten Im-

port statisch ist, werden die SQL-Anweisungen aus dem Mapfile nun vorher prozessiert und in dedizierten Tabellen abgelegt. Die Abfragen des Mapfiles reduzieren sich somit auf einen simplen SELECT. Die Serverlast konnte somit signifikant verringert werden. Des Weiteren sollen die Layer der Flurstücke und der tatsächlichen Nutzung per GetFeatureInfo abfragbar sein und aggregierte Informationen darstellen. Um dies zu realisieren, wird die Tabelle „po_polygons“ im Nachgang um Attribute erweitert.

Zudem sollen aus der vorhandenen Datenbank weitere Produkte abgeleitet werden:

- Hauskoordinaten und Hausumringe
- WFS für öffentlich-rechtliche Festlegungen
- WFS für die Flurstücksabschnitte
- Diverse WFS zur Nutzung als interne und externe Suchdienste
- WMS und WFS für die Umlegungsgebiete
- WMS und WFS für die Bodenschätzung

Sämtliche Tabellen, die eigene Erweiterungen beinhalten, werden mit dem Präfix „pt_“ (für Präsentations-Tabelle) versehen. Somit ist eine Unterscheidung zu den übrigen Tabellen in der Datenbank gesichert.

WMS mit Informationen zum Liegenschaftskataster

Das Mapfile zur Darstellung der Objekte aus dem Liegenschaftskataster erzeugt in seiner ursprünglichen Form 84 Layer, einschließlich der Gruppenlayer! Dieser WMS ist für die Nutzung durch Bürgerinnen und Bürger zu undurchsichtig. Die Anordnung der Layer folgt auch nicht durchgehend einem eindeutigen Schema: So stehen die Flächen der tatsächlichen Nutzung im ersten Gruppenlayer. Erst im weiteren Verlauf folgen die Gruppenlayer für die jeweiligen Nutzungsarten, in denen wiederum Grenzen, Punkte und Beschriftungen signaturiert werden. Ist lediglich eine Darstellung der tatsächlichen Nutzung gewünscht, so müssten die Layer individuell ausgewählt werden und die Bedeutung der einzelnen Themen auch verstanden sein. Da dies für Nicht-Fachanwender wenig benutzerfreundlich erscheint, wird ein kaskadierter WMS erstellt.

Das ursprüngliche Mapfile wird mittels eines weiteren Mapfiles über den localhost des Serversystems abgefragt. Hierbei werden die Layer des Quell-Mapfiles mittels „wms_name“ thematisch zusammengefasst und entsprechend benannt. Bei der Benennung wird den Namen des AdV-Produktes zum ALKIS-WFS soweit wie möglich gefolgt. Da bei dieser Methode die Legende nicht mehr korrekt wiedergegeben werden kann, muss ein eigenes Legendenbild unter Verwendung von „wms_style_default_legendurl_href“ hinzugefügt werden. Listing 3 zeigt ausschnittsweise die METADATA-Sektion des Layers „Nutzung“, wodurch diese Darstellungen allesamt zusammengefasst werden.

Listing 3: METADATA-Sektion des Layers „Nutzung“ im akkumulierenden Mapfile

```
        METADATA
            "wms_title" "Nutzung"
            […]
            "wms_connectiontimeout" "3"
            "wms_dataurl_format" "text/html"
            "wms_style_default_title" "default"
            "wms_feature_info_mime_type" "text/html"
            "wms_include_items" "all"
            "wms_name" "fh,fb,fe,fw,fs,fi,fv,gh,gb,ge,gs,gw,gi,gv,bh,bb,be,bs,bw,bi,bv"
            "wms_server_version" "1.3.0"
            "tile_map_edge_buffer" "100"
            "tile_metatile_level" "0"
            "wms_style" "default"
            "wms_style_default_legendurl_format" "image/png"
            "wms_style_default_legendurl_height" "510"
            "wms_style_default_legendurl_width" "470"
              "wms_style_default_legendurl_href" "https://geo5.service24.rlp.de/liegenschaf-
ten/liegenschaften_legende_nutzung.png"
        END # METADATA
```

AFIS-Import und AdV-konformer WMS

Der Import der NAS-Daten aus AFIS findet analog zu dem ALKIS-Import statt. Die Nutzung des norBIT ALKIS-Imports bietet sich deshalb an, da im Datenschema von AFIS einige Ausprägungen vorliegen, die durch diesen Import korrekt abgebildet werden. Allerdings wird das Postprocessing nicht ausgeführt, da es keine Objektarten in AFIS gibt, die zu verarbeiten wären. Stattdessen wird unmittelbar nach dem Import der AFIS-Daten ein eigenes Skript zur Aufbereitung der Punktdaten gestartet. Dieses Skript erstellt für die AFIS-Punktarten eigene Tabellen mit der Punktkennung, der Ordnung und der Geometrie. Derzeit findet diese Prozessierung für den AdV-konformen AFIS-WMS und für einen internen WFS-Suchdienst Verwendung.

ATKIS-Import und tagesaktueller WMS BasisDLM

Im Gegensatz zu den beiden anderen Modellen wird beim Import von NAS-Daten aus ATKIS auf die Verwendung des norBIT ALKIS-Imports verzichtet. Stattdessen wird der Import mittels purem OGR realisiert, da für den ATKIS-Import keine Funktionalitäten des norBIT ALKIS-Imports benötigt werden. Hierzu müssen einige Umgebungsvariablen gesetzt sowie die gzip-Dateien jeweils entpackt und verarbeitet werden (s. Listing 4). Durch das Setzen des Parameters „-update“ werden die zu importierenden Datensätze hinzugefügt, ggf. bestehende Datensätze gelöscht oder verändert. Bei einem Erstimport kann der Update-Parameter weggelassen werden. Diese Einstellungen und Verarbeitungsschritte führt der norBIT ALKIS-Import in ähnlicher Weise durch. Die weitere Verarbeitung findet mittels eigener Skripte statt, die eine inhaltliche und geometrische Generalisierung der Geodaten vornehmen. Aktuell werden nahezu alle Objektarten des BasisDLM verarbeitet und können entsprechend präsentiert werden. Aufgrund der geringeren Komplexität des ATKIS-Modells können Differenzdaten schnell eingespielt werden. Somit ermöglicht dieses Verfahren eine tagesaktuelle WMS-Darstellung des BasisDLM.

Listing 4: Ausschnitt aus dem ATKIS-Importskript

```
[...]
export GML_FIELDTYPES=ALWAYS_STRINGS
export OGR_SETFIELD_NUMERIC_WARNING=ON
export OGR_ARC_MINLENGTH=0.1
export OGR_SKIP=GML
export NAS_INDICATOR="NAS-Operationen;AAA-Fachschema;aaa.xsd;aaa-suite"
export NAS_GFS_TEMPLATE=alkis-schema.gfs
export NAS_NO_RELATION_LAYER=NO
export EPSG="-a_srs EPSG:25832"
export DBUSER=import
export DBNAME=atkis

[...]
   ####################
   # insert NAS-ATKIS #
   ####################

   process_xml()
   {
      gzdatei=$1
      echo $gzdatei
      gunzip -c $gzdatei > $gzdatei.xml

      # OGR-Import #
      ogr2ogr -update -append -f "PostgreSQL" -skipfailures -gt 65000 -ds_transaction \
         –config  PG_USE_COPY YES -nlt CONVERT_TO_LINEAR \
         PG:"dbname=$DBNAME user=$DBUSER" $EPSG ${gzdatei}.xml

      rm $gzdatei.xml
   }

   export -f process_xml
   cd $insert_folder
   parallel --jobs $cores process_xml ::: *.xml.gz 2>> $errprot
```

Laufzeiten und Ressourcen

Abschließend stellt sich die Frage nach den Laufzeiten der jeweiligen Prozesse. Hierbei ist von den folgenden Rahmenbedingungen des Serversystems auszugehen: Virtuelle Maschine mit 32 GB RAM, 16 Cores und SAN-Speicher.

Bei dem verwendeten System liegen die Laufzeiten für den **AFIS-Vollimport** und die Nachverarbeitung bei knapp 15 Minuten. Für das **BasisDLM** liegt die Dauer eines Vollimportes bei ca. 90 Minuten, tägliche Differenzdaten werden innerhalb weniger Minuten verarbeitet. Der **ALKIS-Vollimport** benötigt mit sämtlichen Prozessierungsschritten rund 65 Stunden. Für einen Differenzdatensatz über den Zeitraum von 7 Tagen benötigte die gesamte Prozessierungskette rund 15 Stunden, wovon der norBIT ALKIS-Import dank der inkrementellen Abarbeitung zirka 8 Stunden Laufzeit hatte.

Weitere Entwicklungen und Fazit

Aktuell befindet sich ein WFS zur Nutzung der ALKIS-Punkte in Entwicklung, der objektartengetrennt die wesentlichen Eigenschaften der Punkte beinhalten soll. Zudem ist geplant, den Workflow für die Bodenschätzung in das inkrementelle Postprocessing vollständig zu implementieren, sodass auch hier Zeitvorteile genutzt werden können. Aktuell plant das LVermGeo RP keine Entwicklung und Bereitstellung von Eigentümerdaten über einen abgesicherten WFS. Aufgrund der Nachweispflicht gemäß DSGVO wird die bisherige Auskunftsmethode favorisiert. Die Nutzung der Differenzdaten und der inkrementellen Abarbeitung bietet nun die Möglichkeit, dass die ALKIS-Dienste von einem monatlichen auf einen wöchentlichen Zyklus umgestellt werden können.

Im Bereich ATKIS ist eine direkte Ableitung des AdV-Produktes „ATKIS-Basis-DLM-Shape“ [8] anvisiert, welche dann ebenfalls als WFS verfügbar gemacht werden soll. Aktuell wird der entsprechende WFS mit monatsaktuellen Daten betrieben, die mittels eines anderen Workflows in FME abgeleitet werden.

Abschließend bleibt festzuhalten, dass der Import, das Postprocessing und die eigenen Skripte zu einer nahezu vollumfänglichen Verarbeitung und Bereitstellung von 3A-Daten verwendet werden können. Die meisten hierzu verwendeten Skripte und Konfigurationen werden unter der Lizenz „GNU GPL v2“ zur freien Verwendung bereitgestellt [9]. Ergänzungen, Verbesserungsvorschläge und Weiterentwicklung des gesamten Projektes sind immer herzlich willkommen!

Kontakt zum Autor:

Oliver Schmidt
Landesamt für Vermessung und Geobasisinformation Rheinland-Pfalz
Von-Kuhl-Straße 49, 56070 Koblenz
Telefon: 0261/492-346
eMail: oliver.schmidt@vermkv.rlp.de

Literatur und Verweise

[1] https://lvermgeo.rlp.de/geodaten-geoshop/open-data

[2] Schmidt, Oliver: FOSS hoch 5 - Praxisbericht eines Serversystems mit OpenSource-GIS, Tagungsband FOSSGIS 2021, FOSSGIS e.V., S. 78-84.

[3] https://www.adv-online.de/GeoInfoDok/Aktuelle-Anwendungsschemata/AAA-Anwendungsschema-7.1.2-Referenz-7.1/

[4] https://postnas-suite.github.io/postnas-suite/

[5] https://github.com/norBIT/alkisimport/tree/gid7i

[6] https://www.norbit.de/68/

[7] https://www.adv-online.de/AdV-Produkte/Standards-und-Produktblaetter/Standards-des-Liegenschaftskatasters/

[8] https://www.adv-online.de/AdV-Produkte/Standards-und-Produktblaetter/Standards-der-Geotopographie/

[9] https://github.com/postnas-suite

GIS-Schulprojekte in Zusammenarbeit mit kommunalen Gebietskörperschaften

Dietmar Holzner [1], Christian Steiner [1],

[1] Technologische Fachoberschule „Oskar von Miller“-Meran/Südtirol/Italien–TFO Meran

Zusammenfassung: Anhand von zwei Projekten stellen Professoren und Schüler/-innen vor, wie an der TFO- Meran GIS praxisnahe im Unterricht eingesetzt wird.

Schlüsselwörter: Ausbildung, Lehre, Vermessung, GIS, QGIS, PostGIS, postgreSQL

Die Schule: Die TFO-Meran ist eine fünfjährige Fachoberschule mit zwei Fachrichtungen: „Bauwesen, Umwelt und Raumplanung“ und „Elektronik und Elektrotechnik“ und schließt mit der Matura (Abitur) ab. Technik ist an der Schule wichtig – aber nicht alleine. Angestrebt wird eine umfassende Ausbildung, bei der auch Sprachen und allgemeinbildende Fächer neben den technischen Fächern Platz haben.

Das Fach: Der Unterricht im Fach Vermessung der Fachrichtung Bauwesen vermittelt grundlegende und vertiefte Kenntnisse über das Messen, Analysieren und Darstellen von geographischen und geometrischen Daten. Ziel des Unterrichts ist es, den Schülerinnen und Schülern die theoretischen Grundlagen sowie die praktische Anwendung von Vermessungstechniken beizubringen. Inhalte und Schwerpunkte sind die Grundlagen der Vermessungstechnik, Einführung in Messinstrumente und deren korrekte Handhabung, geodätische Berechnungen, Kartografie und Plandarstellung, Anwendung von klassischen und modernen Messmethoden, Datenerfassung und -auswertung. Der Unterricht zeichnet sich durch einen hohen Praxisanteil aus. Die Studierenden führen eigenständig Messprojekte im Freien durch, analysieren die Ergebnisse und präsentieren sie in Form von Plänen oder Berichten.



Abb. 1: Vermessung von Lichtpunkten

In der Abschlussklasse findet eine Vermessungswoche statt, bei der praxisbezogene Projekte in Zusammenarbeit mit öffentlichen Körperschaften, Gemeinden, Freiberuflern oder Firmen des Bausektors ausgeführt werden. In den vergangenen Jahren wurde der Bereich der geografischen Informationssysteme in den Unterricht und die Praxis aufgenommen. Damit lernen die Absolventen der Oberschule digitale Werkzeuge zur Erfassung, Analyse, Visualisierung und Verwaltung räumlicher Daten kennen. Sie verbinden geografische Informationen durch Aufnahme und Erstellung von Lageplänen mit Attributdaten, um je nach Aufgabenstellung Zusammenhänge zwischen Standorten, Landschaften und Objekten darzustellen und zu untersuchen.

Die Projekte: Die Grundlagen der Projektarbeiten werden im Zuge einer Vermessungswoche gesammelt, bei der sich die Studierenden eine ganze Woche im Projektgebiet aufhalten, um sich intensiv mit der praktischen Erhebung, Auswertung und Darstellung befassen zu können. Die Vermessungen erfol-

gen traditionell mit Maßband, Theodolit und GPS-Empfängern, während die Erfassung der Daten anfänglich auf Basis von Textfiles und Tabellenkalkulationen beruht.

Anhand von zwei Projektarbeiten wird die Zusammenarbeit zwischen Kommunen und Schule präsentiert.

Projekt 1: Friedhofs-Informations-System Gemeinde Ulten - Friedhof St. Gertraud. Die Topografie des Friedhofs wurde vermessen und als georeferenzierte CAD-Zeichnung in QGIS importiert. Die Grabstellen wurden mittels GPS erfasst und die Koordinaten in einer Geodatenbank (postgreSQL mit Geomodul PostGIS) abgelegt. Diese wird auf einem von der Schule betriebenen LINUX-Server gehostet und ist damit multiuserfähig. Die grafische Darstellung in QGIS erfolgt durch die Verbindung zur angelegten Datenbank. In diese wurden zusätzlich die anagrafischen Daten der Verstorbenen und der Nutzungsberechtigten, ein Rechnungsarchiv und eine Verknüpfung mit Bildern der Grabstellen mit aufgenommen. Das neu erstellte Informationssystem ersetzt die bisher händisch angelegte Grabbuchführung und ermöglicht nunmehr der Gemeinde ein umfangreiches Datenmanagement.

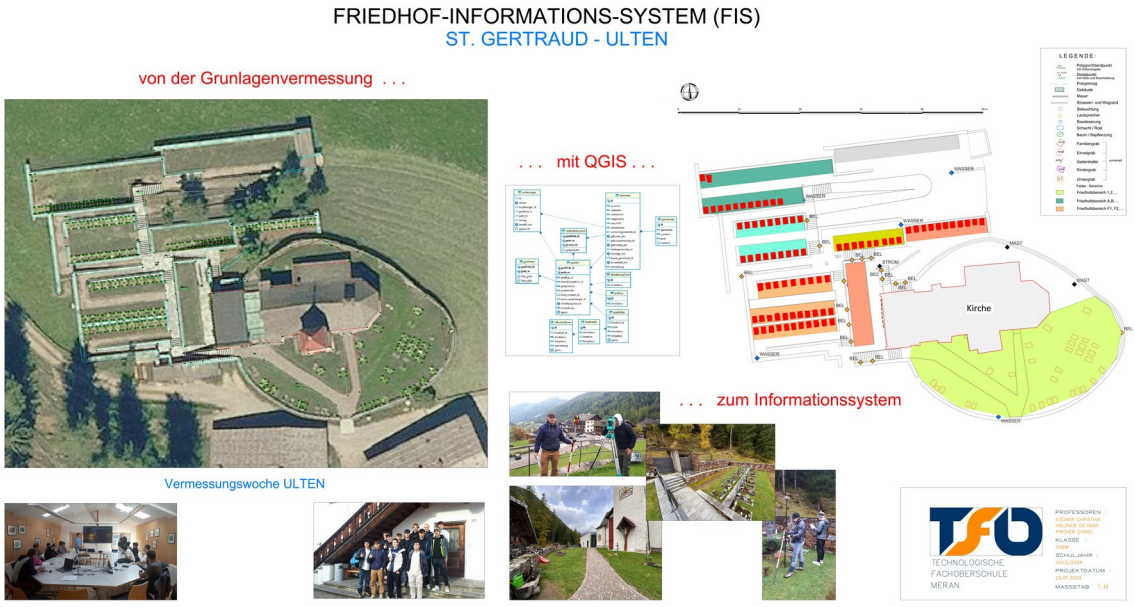


Abb. 2: Plakat Friedhofs-Informations-System (Friedhof St. Gertraud/Ulten)

Projekt 2: Aktualisierung des Lichtplanes der Gemeinde St. Martin in Passeier. Der durch ein Pilotprojekt im Jahre 2014 genehmigte Aktionsplan zur Erstellung des kommunalen Lichtplans muss spätestens alle 10 Jahre aktualisiert werden. In Zusammenarbeit mit der Gemeindeverwaltung konnten sämtliche Lichtpunkte der öffentlichen Außenbeleuchtung und die zugehörigen Leitungsstränge und Kabelschächte georeferenziert erfasst und in einer PostGIS-Datenbank abgelegt werden. Die Kodifizierungen wurden an die Vorgaben für den Datenaustausuch mit der

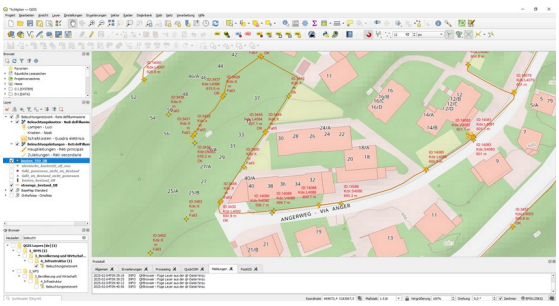


Abb. 3: Screenshot QGIS ver. 3.36.0-maidenhead

Landesverwaltung der Autonomen Provinz Bozen-Südtirol und dem Südtiroler Gemeindenverband angepasst.

Die Bestandsdaten wurden als Shape-File in QGIS übernommen und mit den aktuellen Daten abgeglichen. Die Anwendung geeigneter Filter erlaubt eine Überprüfung und Kontrolle der Bestände. Abweichungen zwischen „Bestand“ und „Neu“ werden im Zuge einer neuerlichen Begehung mit den Technikern der Gemeinde überprüft und in der Datenbank berichtigt.

In beiden gezeigten Projekten erfolgt die grafische Datendarstellung und der Export in übliche GIS-Formate mittels QGIS. Die Direktanbindung an die Datenbanken sorgt für Konsistenz im Datenbestand. QGIS ermöglicht zudem einen effizienten Datenaustausch zwischen den CAD- und GIS-Systemen der Südtiroler Landesverwaltung. Die Autonome Provinz Bozen-Südtirol stellt den „GeoBrowser MapView“ und den „GeoKatalog“ den Bürgerinnen und Bürgern zum freien Zugriff zur Verfügung. Mit dem QGIS-Plugin "Qlr Browser" können diese Geodaten als separate Layer direkt in die personalisierten Projekte eingebunden werden.

Fazit: Die Fallbeispiele zeigen, dass die Verarbeitung von räumlichen Daten mit Quantum GIS (QGIS) einen konsistenten Datenfluss erlauben, der bei entsprechender Einschulung junge Techniker/-innen zu erfolgreichen Ergebnissen führt. Werden praxisnahe, alltagsrelevante oder schulspezifische Themen behandelt, können Schüler/-innen motiviert und das Interesse für diesen Themenbereich geweckt werden. Auf eine vertiefende Einführung in den Entwurf, die Erstellung und das Management relationaler und georeferenzierten Datenbanken muss in Zukunft besonderes Augenmerk gelegt werden.

Kontakt zu den Autoren:

Dietmar Holzner, Christian Steiner
Technologische Fachoberschule Meran
Karl-Wolf-Straße 36
I-39012-Meran (BZ)
+39-473-200489
os-rg-tfo.meran@schule.suedtirol.it

Links (Abruf am 03.02.2025):

[1] Technologische Fachoberschule Meran „Oskar von Miller“: https://www.tfo-meran.it

[2] GeoBrowser MapView der Autonomen Provinz Bozen-Südtirol: https://mapview.civis.bz.it

[3] QGIS Plugin „Qlr Browser“: https://plugins.qgis.org/plugins/QlrBrowser/

eoAPI - eine skalierbare Geodateninfrastruktur

Felix Delattre

In diesem Vortrag geht es um eoAPI, ein Open-Source-Baukasten zur schnellen Bereitstellung einer skalierbaren Geodateninfrastruktur auf Grundlage der STAC-Spezifikation. eoAPI ermöglicht die einfache Auffindbarkeit und breite Nutzung von Erdbeobachtungsdaten. Es werden ihre leistungsstarken Funktionen sowie die Komponenten TiTiler, Tipg, pgSTAC und stac-fastapi erläutert.

In diesem Vortrag wird eoAPI vorgestellt, es handelt sich um einen modularen Baukasten aus Open-Source-Komponenten zur schnellen Bereitstellung einer skalierbaren, cloud-basierten und offenen Geodateninfrastruktur auf Grundlage der STAC-Spezifikation. Erfahren Sie, wie Sie mit eoAPI sowohl kleine als auch große Archive von Erdbeobachtungsdaten schnell und einfach auffindbar und interoperabel machen und einer breiten Anwenderbasis zur Verfügung stellen können.

Das Ökosystem von eoAPI wurde von Development Seed in Zusammenarbeit mit der NASA, Microsoft Planetary Computer und anderen Partnern entwickelt. Entsprechende Backends, die mit eoAPI laufen, liefern derzeit Metadaten für Milliarden von STAC-Elementen und generieren täglich dynamisch Hunderttausende von Bildkacheln.

Es werden die leistungsstarken Funktionen von eoAPI aufgezeigt und die einzelnen Komponenten wie TiTiler, Tipg, pgSTAC und stac-fastapi erläutert.

aviary - ein generisches Python-Framework zur KI-Inferenz für Fernerkundungsdaten

Marius Maryniak

Zusammenfassung: aviary [1] ist ein generisches Python-Framework, das die Inferenz von KI-Modellen für Fernerkundungsdaten vereinfacht. Es bietet verschiedene Pipelines mit austauschbaren, erweiterbaren Komponenten. Neben der Nutzung als Python-Package können vorgefertigte Pipelines über die Kommandozeile verwendet werden. Künftig sollen vortrainierte Modelle für diverse Anwendungsfälle sowie speziell auf Fernerkundungsdaten trainierte Foundation-Modelle bereitgestellt werden.

Schlüsselwörter: Machine Learning, Deep Learning, Computer Vision, Python, Luftbilder, Satellitenbilder, Fernerkundung

Wir, das geospaitial lab [2] der Westfälischen Hochschule, sind ein Forschungsteam für Künstliche Intelligenz und kommunale Geoinformationen und haben in den letzten Jahren verschiedene Auftragsforschungen und Förderprojekte im Bereich automatisierter Auswertung von Mobile-Mapping-Daten und Luftbildern umgesetzt. Darunter war eine Auftragsforschung des Kreises Recklinghausen, in der versiegelte Flächen aus Luftbildern abgeleitet werden sollten. Dazu wurde ein KI-Modell trainiert und eine entsprechende Software (adois [3, 4]) entwickelt, mit der sich die KI-Inferenz ausführen lässt.

Bei adois handelt es sich jedoch um eine prototypische Implementierung, die beispielsweise nur die Einbindung von Luftbildern über einen Web Map Service erlaubt und nicht als generische Lösung konzipiert wurde. Als der Regionalverband Ruhr uns daraufhin beauftragt hat, adois auch mit lokalen Daten anwendbar zu machen, haben wir uns daher entschieden, die Software neu zu konzipieren und zu entwickeln.

Daraus ist aviary entstanden - ein neues Framework, das uns und anderen Forscher:innen und Anwender:innen in der Geoinformatik die Nutzung von KI-Modellen erleichtern soll - schließlich wollen wir uns aufs Forschen, Entwickeln und Trainieren konzentrieren.

Die Inferenz von KI-Modellen hat verschiedenste Anforderungen, folgt jedoch stets einem ähnlichen Ablauf:

- **Definition des Prozessierungsbereichs:** Welcher Bereich soll ausgewertet werden?
- **Definition der Eingangsdaten:** Woher kommen die Eingangsdaten? Liegen sie lokal vor, werden sie über einen Web Service abgerufen oder sind sie in der Cloud?
- **Definition der Vorverarbeitung der Eingangsdaten:** Wie werden die Eingangsdaten vorverarbeitet?
- **Definition des KI-Modells:** Welches KI-Modell soll verwendet werden?
- **Definition der Nachverarbeitung der Ausgangsdaten:** Wie werden die Ausgangsdaten des KI-Modells nachverarbeitet?
- **Definition des Exports:** Wohin werden die Ergebnisse exportiert? Werden sie lokal gespeichert oder in die Cloud hochgeladen?

aviary abstrahiert diese Schritte und bietet Pipelines, die aus verschiedenen Komponenten bestehen. Diese Komponenten sind austausch- und erweiterbar, sodass Anwender:innen eigene Pipelines bauen und bei Bedarf Funktionalitäten hinzufügen können. Dabei ist aviary so konzipiert, dass es möglichst effizient und skalierbar ist. Die Pipelines können beispielsweise auf einem lokalen Rechner oder je nach Bedarf in der Cloud ausgeführt werden.

Neben der Möglichkeit, aviary als Python-Package zu nutzen, können Anwender:innen auch vorgefertigte Pipelines über die Kommandozeile verwenden. Dazu werden deklarative Konfigurationsdateien verwendet, mit denen die Pipelines auf einfache Weise definiert werden können. Dies macht aviary auch ohne Programmierkenntnisse nutzbar.

Das Modell zur Ableitung versiegelter Flächen aus Luftbildern wurde in aviary integriert. Weitere Modelle und Pipelines sollen folgen, darunter auch vortrainierte Modelle für diverse Anwendungsfälle sowie speziell auf Fernerkundungsdaten trainierte Foundation-Modelle. Dies soll im Rahmen unserer Forschungstätigkeiten in URBAN.KI, der Deutschen KI-Initiative für Kommunen, umgesetzt werden.

Kontakt zum Autor:

Marius Maryniak
Westfälische Hochschule, geospaitial lab
Neidenburger Straße 43
45897 Gelsenkirchen
+492099596-4854
marius.maryniak@w-hs.de

Literatur

[1] Maryniak, Marius; geospaitial lab: aviary, https://www.github.com/geospaitial-lab/aviary, 2024.

[2] geospaitial lab: Website, https://github.com/geospaitial-lab

[3] Maryniak, Marius; geospaitial lab: adois, https://github.com/geospaitial-lab/adois, 2022.

[4] Maryniak, Marius et al., VDVmagazin 3/23, S. 198 - 205: Wie eine Maschine Versiegelungskarten erstellt, 2023 (abrufbar unter: https://webgis-re.de/fileadmin/user_upload/artikel_wie_eine_maschine_versiegelungskarten_erstellt_aus_vdvmagazin_3_23.pdf).

GDI per Knopfdruck: Automatisierung mit DevOps und Infrastruktur als Code

Jakob Miksch

Der Vortrag zeigt auf wie eine Geodateninfrastruktur(GDI) automatisiert über Code eingerichtet werden kann und geht dabei auf die von uns genutzten Software-Bausteine und deren Alternativen ein.

Es gibt viele Ansätze, eine Geodateninfrastruktur bereitzustellen. Durch die starke DevOps-Kultur bei siticom ist unsere hauseigene Geodateninfrastruktur komplett über Code definiert und stark automatisiert. Dieser Vortrag beschreibt die Vorzüge und die Herausforderungen dieser Herangehensweise.

Unsere Dienste laufen auf virtuellen Linux-Servern bei einem Cloud-Provider. Diese werden allerdings nicht manuell über eine Benutzeroberfläche „zusammengeklickt“, sondern mit dem Tool Terraform über Code definiert. Das ermöglicht Reproduzierbarkeit und einfache Anpassbarkeit.

Das Einrichten der Server funktioniert über Ansible. Diese auf Python basierende Konfigurationssprache ist verhältnismäßig lesbar und stellt sicher, dass alle benötigten Programme auf den Servern installiert und alle notwendigen Dateien vorhanden sind.

Sowohl die Terraform- als auch die Ansible-Konfiguration sind in Git-Repositories versioniert und können über die GitLab CI ausgeführt werden. Dies ermöglicht eine hohe Transparenz der Infrastruktur bei allen Teammitgliedern. Außerdem stellt es die Reproduzierbarkeit der Geodateninfrastruktur sicher, da alle alten Stände wiederhergestellt werden können.

Die benötigten Passwörter und Zugänge werden in einem eigenen Secret-Speicher vorgehalten und bei jedem Deployment per API abgefragt. Dieser Ansatz erleichtert das Ändern von Zugängen wesentlich, was zu erhöhter Sicherheit beiträgt.

Wir deployen unsere Software ausschließlich über Docker-Images, die wir sowohl aus öffentlichen als auch internen Docker-Registries beziehen. Die Konfiguration und Zusammenstellung wird über Docker-Compose-Dateien definiert, die selbstverständlich per Git versioniert sind.

Um unsere Software aktuell zu halten, nutzen wir den Renovate Bot. Dieser prüft regelmäßig, ob alle Docker-Images bzw. Software-Bibliotheken auf dem neuesten Stand sind. Etwaige Updates können dann nach menschlicher Prüfung per Knopfdruck in die Konfiguration gemergt und direkt deployt werden.

Ein Anwendungsfall dieser Automatisierung ist das Veröffentlichen von QGIS-Plugins im Intranet. Sobald eine neue Version eines QGIS-Plugins bereitsteht, wird dieses per GitLab-Release als Download bereitgestellt und in das QGIS-Plugin-Repository eingepflegt, von wo aus es für QGIS-Nutzer komfortabel genutzt werden kann.

basemap.de als Open Data - Neue Stile und Anwendungsbeispiele

Arnulf B. Bichler (aka Christl)

Die Grundlage der basemap.de sind amtliche Daten der Länder, die seit Mitte 2024 als Open Data bereitgestellt werden. In dem Vortrag werden neue Möglichkeiten gezeigt, die sich daraus ergeben und neue Stile vorgestellt. In Anwendungsbeispielen werden die Potenziale für Behörden und eigene Entwicklungen gezeigt.

Die Kartendienste basemap.de sind von Bund und Ländern gemeinsam entwickelte und durch die Zentrale Stelle Geotopographie (ZSGT) im Auftrag der Länder bereit gestellte Internet-Kartendienste. Auf der Grundlage amtlicher Geobasisdaten vermitteln sie eine attraktive, deutschlandweit einheitliche Kartendarstellung in Zoomstufen vom Einzelgebäude bis zur Deutschlandübersicht. Die leistungsfähige Realisierung auf Basis internationaler Standards und ihrer modernen kartographischen Gestaltung unterstützen vielfältige Anwendungsbereiche.

Gemeinsam ist allen basemap.de-Produkten, dass sie aus einem zentralen Verfahren und Datenbestand abgeleitet werden, wodurch eine schnelle, flexible und wirtschaftliche Erzeugung von kartographischen Produkten und Diensten ermöglicht wird. Neben Webkarten unterschiedlicher Formate (Vektor und Raster) und Ausprägungen werden Reliefinformationen bundesweit berechnet und bereit gestellt.

Automatischer Import und Veröffentlichung von Betriebsmittelgeometrien mittels PyQGIS

Christoph Jung[1], Philipp Opitz[1]

[1] SachsenEnergie AG, Dresden

Zusammenfassung: Ziel von Versorgungsunternehmen ist ein minimaler Zeitraum zwischen Beendingung einer Baumaßnahme und Integration der Geodaten in die Leitungsauskunft. Eine Möglichkeit bietet die automatische Integration der Geometrien als Vorabauskunft. In der Umsetzung bei SachsenEnergie wurde für die Datenveröffentlichung ein ETL-Prozess auf QGIS/PyQGIS-Basis entwickelt, welcher differenziell und performant nächtlich die DXF-Daten in eine Datenbank überführt. Die Publikation der Daten erfolgt mittels WMS.

Schlüsselwörter: DXF, PyQGIS, Schnittstelle, Leitungsauskunft

1. Motivation

Erdarbeiten, egal ob auf öffentlichem oder privatem Grund, erfordern gemäß DIN VDE vorab eine Leitungsauskunft im Sinne einer Sorgfaltspflicht, damit es nicht zur Beschädigung von Versorgungsinfrastruktur infolge der Bauarbeiten kommt (SachsenEnergie, 2024). Hierfür stellen Versorgungsunternehmen IT-Systeme zur Beantragung einer Auskunft über die vorhandenen Betriebsmittel in einem Gebiet zur Verfügung. Grundlage der Auskunft ist der dauerhaft gepflegte Bestand an Betriebsmittelinformationen, welcher i.d.R. in einem GIS verwaltet wird. Kritisch ist dabei der Zeitraum zwischen Beendigung einer Baumaßnahme und der Aktualisierung des Geodatenbestandes, da der Prozess der Einmessung neuer Betriebsmittel, deren Übernahme in das Zielsystem sowie die Qualitätskontrolle eine gewisse Zeit benötigen. Während dieses Prozesses sind neue Leitungen im Erdreich vorhanden, werden aber noch nicht beauskunftet. Um diesen Zeitraum zu Verringern und damit die Wahrscheinlichkeit von Schäden an Versorgungsleitungen, aber auch „[…] zur Abwendung von Gefahren, für die auf der Baustelle beschäftigten […]“ (SachsenEnergie, 2024) bedingt durch fehlende Informationen in der Leitungsauskunft zu reduzieren, wurde bei der SachsenEnergie AG eine Schnittstelle zur automatischen Integration der Geometriedaten der Betriebsmittelinformationen realisiert.

2. Systemvoraussetzungen

Die Geodateninfrastruktur (GDI) bei der SachsenEnergie AG ist durch den Einsatz verschiedener Technologien geprägt.

Zur Verwaltung und Analyse der georeferenzierten Betriebsmittel wird Smallworld GIS inkl. spezifischer Fachschalen eingesetzt. Da man nur schwer mit Werkzeugen Dritter auf Smallworld-Datenbanken zugreifen kann, werden zusätzlich Oracle- und PostgreSQL-Datenbanken inkl. deren räumlichen Erweiterungen eingesetzt. Für die Realisierung der Schnittstelle wurde Oracle als Datenbank ausgewählt, wobei keine Oracle-eigenen Funktionen genutzt werden, die es in ähnlicher Form nicht auch in PostgreSQL und PostGIS gibt. Eine Realisierung der Schnittstelle lässt sich demnach auch mit PostgreSQL umsetzen.

Die Vermessungsdaten liegen bei der SachsenEnergie AG als DXF-Dateien vor. Die Integration von DXF-Objekten in eine Geodatenbank ist eine seit Jahrzehnten bekannte Herausforderung, u.a. da in GIS und Geodatenbanken geografische und projizierte Koordinaten für eine Kartendarstellung gespeichert werden, während in DXF-Dateien Objektkoordinaten hinterlegt sind, welche vielseitig genutzt

und interpretiert werden können (Borşan & Dimén, 2021). Die Integration der DXF-Daten stellt die eigentliche Herausforderung für die Erstellung der Schnittstelle dar.

Als zusätzliche GIS-Plattform wird bei der SachsenEnergie QGIS Desktop für die Arbeit mit Geodaten und QGIS Server für die Publikation von Geodaten über OGC Services. Um auf bereits im Einsatz befindliche Technologien zu setzen, wurde für die Schnittstellenentwicklung die PyQGIS-API genutzt. Diese API weist alle Werkzeuge zur Verarbeitung von DXF-Dateien und dem Import der enthaltenen Objekte in eine Geodatenbank auf und bietet eine hinreichend hohe Abstraktion in der GIS-basierten Datenverarbeitung, um den zu entwickelnden Code möglichst gering zu halten.

3. Technischer Aufbau

3.1. Vorbemerkungen

QGIS bietet die Funktion des Imports von DXF-Daten über den Menüpunkt Project – Import/Export – Import Layers from DWG/DXF… (Abb. 1), welcher jedoch nicht direkt über die API verfügbar ist. Mittels PyQGIS wurde ein ETL-Prozess zur Überführung der Geometrien aus den DXF-Daten nach Oracle erstellt, welcher sich an der DXF-Import-Funktion orientiert. Die entsprechenden Daten im Oracle werden über den QGIS Server als WMS publiziert und an die Leitungsauskunft angebunden.

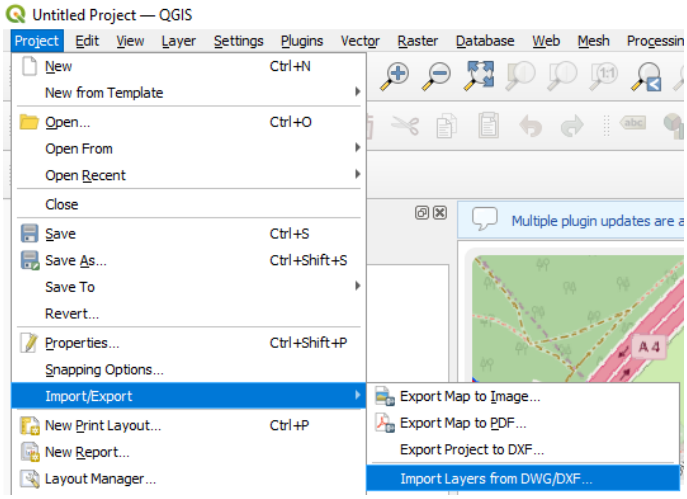


Abbildung 1: Menüpunkt zum Import von DXF-Dateien in QGIS

3.2. Datenmodell

Für die Speicherung der Vermessungsdaten wird ein relationales Datenmodell benötigt, welches einen effizienten Umgang mit mehreren Millionen Datensätzen sicherstellt. Hierfür werden insgesamt drei Tabellen zur Speicherung der Metadaten der DXF-Dateien, deren Bounding Boxes und der eigentlichen Betriebsmittelgeometrien genutzt (siehe Abbildung 2).

Die Metadatentabelle bildet die Grundlage für einen differenziellen Abgleich, welcher nötig ist, um die Laufzeit der Schnittstelle gering zu halten. Für jede DXF-Datei wird ein Datensatz in der Metadatentabelle geführt. Die 1:n- bzw. 1:1-Beziehungen der Metadatentabelle zu den Bounding Boxes und Betriebsmittelgeometrien sind jeweils mit einem Constraint versehen, sodass beim Löschen eines Metadatensatzes alle referenzierten Datensätze der anderen beiden Tabellen gelöscht werden. Dies erhöht die Performance der Schnittstelle, da die Löschaktionen speziell der Betriebsmittelgeometrien direkt

auf Datenbankebene vorgenommen werden und nicht über den Python-Code ausgelöst werden müssen. Bei mehreren Millionen Datensätzen in der Tabelle der Betriebsmittelgeometrien ist darauf zu achten, rechenintensive Transaktionen möglichst nah an der Datenhaltungskomponente zu realisieren. Des Weiteren garantieren diese Constraints eine konsistente Datenhaltung, in der Bounding Boxes und Betriebsmittelgeometrien zwingend einem Metadatensatz zugewiesen sind.

Zur Anbindung an verschiedene Systeme der GDI können Views auf Grundlage des Datenmodells und seiner Beziehungen erstellt werden. Dies ist eine individuelle Anforderung und geht über den Aufbau der eigentlichen Schnittstelle hinaus und wird daher an dieser Stelle nicht weiter beleuchtet.



Abbildung 2: Datenmodell

3.3. ETL-Prozess

DXF-Dateien können wie andere unterstützte Vektordatenformate als QgsVectorLayer repräsentiert werden, wobei bei DXF-Dateien Features mit unterschiedlichen Geometrietypen enthalten sein können. Iteriert man über die Features von DXF-Dateien, ist der Typ der zugehörigen Geometrie abzufragen, um die Geometrie dem korrekten Feld in der Zieldatenbank zuzuordnen. Nachfolgend der Python-Code zur Interation über die Features einer DXF-Datei und der Abfrage des Geometrietyps:

```
vl = QgsVectorLayer('[PFAD_ZU_DXF_DATEI]', 'dxf_layer')
for f in vl.getFeatures():
    print(f.geometry().type())
```

Da die Abfrage des Geometrietyps bei jedem Feature teuer ist, wurde für die Umsetzung der Schnittstelle auf Sublayer zurückgegriffen. QGIS interpretiert DXF-Dateien auch als Sammlung von Sublayern reiner Geometrietypen. Sublayer für die Geometrietypen Punkt, Linie und Polygon werden wie folgt im PyQGIS initialisiert:

```
vl_point = QgsVectorLayer('[PFAD_ZU_DXF_DATEI]|layername=entities|geometrytype=POINT', 'layer_name', 'ogr')

vl_linestring = QgsVectorLayer('[PFAD_ZU_DXF_DATEI]|layername=entities|geometrytype=LINESTRING, 'layer_name', 'ogr')

vl_polygon = QgsVectorLayer('[PFAD_ZU_DXF_DATEI]|layername=entities|geometrytype=POLYGON, 'layer_name', 'ogr')
```

Durch die Verwendung dieser Sublayers ist die korrekte Zuordnung der Features zu den Feldern in der Zieltabelle per Definition gegeben und muss nicht für jedes Feature einzeln ermittelt werden. Zusätzlich kann ein Sublayer vom Geometrietyp GeometryCollection enthalten sein. Eine die einzelnen Geometrietypen einer GeometryCollection können im PyQGIS auch per Sublayers extrahiert werden.

Der gesamte ETL-Prozess hat im Minimum folgende Aufgaben zu erfüllen:

- Auslesen aller DXF-Dateien aus dem Quellverzeichnis
- Differenzieller Abgleich zwischen Metadaten und DXF-Dateien
    - Metadatensätze löschen, für die es keine DXF-Dateien mehr gibt

  - Metadatensätze für DXF-Dateien anlegen, die noch nicht in die Datenbank importiert wurden
- Sublayer der neuen DXF-Dateien auslesen und importieren
- Bounding Box aus den importieren Objektgeometrien berechnen und in die entsprechende Tabelle importieren

Durch den differenziellen Abgleich müssen einmal importierte DXF-Dateien kein zweites Mal importiert werden, wodurch die benötigte Zeit für einen Schnittstellenlauf nach dem initialen Schnittstellenlauf deutlich reduzieren lässt. Auch das automatische Löschen der Betriebsmittelgeometrien nach Entfernen des zugehörigen Metadatensatzes verringert die Laufzeit der Schnittstelle.

QGIS ermittelt automatisch gewisse Sachdaten aus den DXF-Dateien, u.a. die Namen der DXF-Ebenen sowie Werte für punktförmige Objekte wie z.B. Höhenpunkte. Diese Werte werden in die Zieldatenbank importiert, andere Werte wie der Paperspace werden im ETL-Prozess ignoriert.

Während der Entwicklung hat sich herausgestellt, dass regelmäßige Datenbank-Commits von vergleichsweise wenigen Datenänderungen die Schnittstelle stabil laufen lassen im Vergleich zu einem einzigen Commit am Ende des Schnittstellenlaufs. Um in diesem Fall die Datenbanktransaktionen möglichst performant zu halten, v.a. das Einfügen, ist es notwendig, zu Beginn des Schnittstellenlaufes bestehende räumliche Indizes zu löschen und am Ende wieder zu erstellen.

Das folgende Aktivitätendiagramm stellt den kompletten ETL-Prozess und Schnittstellenlauf bei der SachsenEnergie dar:

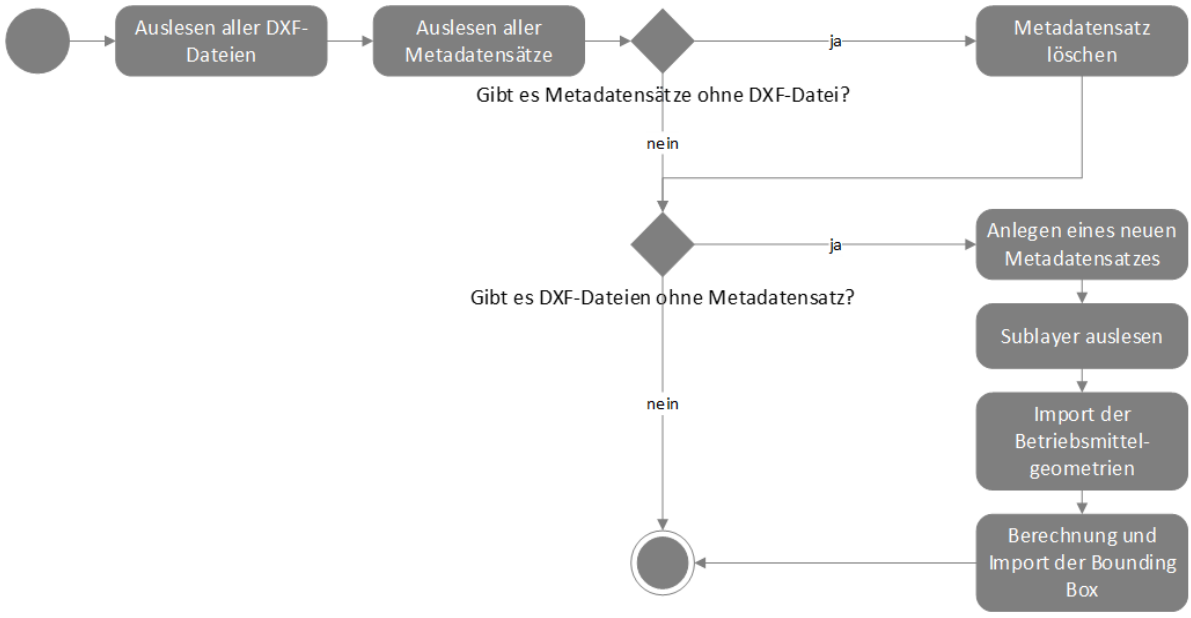


Abbildung 3: Aktivitätendiagramm der Schnittstelle

3.4. WMS-Dienst

Zur Einbindung der Daten in die Leitungsauskunft der SachsenEnergie werden die Betriebsmittelgeometrien in einem QGIS-Projekt mit einer kartografischen Visualisierung und Sichtbarkeitsregeln versehen und anschließend über den QGIS Server per WMS publiziert. Der QGIS Server kann Bilder für GetMap-Requests auch mit hoher Auflösung rendern, sodass die Daten des Dienstes in den Plots der Leitungsauskunft sehr gut lesbar sind.

Zur Anbindung der Daten in GIS-Applikationen und Drittsysteme werden auch die Bounding Boxes als WMS publiziert. Diese sind in allen Maßstabsbereichen sichtbar, während die Betriebsmittelgeomtrien auf große Maßstäbe beschränkt sind. Dies erhöht die Geschwindigkeit des serverseitigen Renderns und verringert Last auf den QGIS Server.

4. Ergebnisse

Da die Schnittstelle jede Nacht läuft, sind nach manueller Ablage der DXF-Dateien im Quellverzeichnis die eingemessenen Geometrien neuer Betriebsmittel am nächsten Tag in der Leitungsauskunft in Form einer Vorabinformation enthalten. Gleichzeitig können die importierten Daten für den Prozess der Betriebsmitteldokumentation weiterverwendet werden. Ein bisher bestehender manueller bzw. teilautomatischer Import der DXF-Daten für die Beauskunftung konnte mit dieser Entwicklung abgelöst werden. Dabei konnte die Entwicklung komplett intern ohne Einbezug eines Dienstleisters erfolgen. Auch werden keine Lizenz- oder Wartungskosten im laufenden Betrieb fällig.

5. Diskussion

Um den Betrieb und die Kontrolle der Schnittstelle zu gewährleisten, ist es nötig, im Skript das Erstellen und Befüllen einer Log-Datei zu implementieren. Welche Informationen in der Log-Datei geschrieben werden sollen, ist eine individuelle Entscheidung.

Bei der Schnittstelle der SachsenEnergie wird Oracle als Datenbank genutzt, wobei PostgreSQL ohne Abstriche ebenfalls genutzt werden könnten. Da PostgreSQL und PostGIS in QGIS stärker integriert sind als Oracle stellt sich die Frage, ob man für die Schritte des Datenimports in die Datenbank auf Standardwerkzeuge des Processing-Frameworks zurückgreifen könnte.

Kontakt zum Autor

Christoph Jung
SachsenEnergie AG
Friedrich-List-Platz 2, 01069 Dresden
0351 5630 25655
christoph.jung@sachsenenergie.de

Philipp Opitz
SachsenEnergie AG
Friedrich-List-Platz 2, 01069 Dresden
0351 563054722
philipp.opitz@sachsenenergie.de

Literatur und Quellen

SachsenEnergie AG (2024). Leitungsauskunft über Internet, abgerufen am 02.12.2024 von https://www.sachsen-netze.de/wps/portal/netze/cms/menu_main/service/onlineservice/leitungsauskunft

Borşan, T., & Dimén, L. (2021). CAD-GIS INTEROPERABILITY. Pangeea, 62-70.

Superset - Business Intelligence meets Cartography

Jan Suleiman

Superset ist eine der meistgenutzten Open-Source Tools im Bereich der Business Intelligence (BI). Wir haben Superset um die Erstellung thematischer Karten erweitert, wodurch auch raum-zeitliche Aspekte besser analysiert werden können. In diesem Vortrag zeigen wir, wie sowohl Kartodiagramme, Choroplethenkarten, sowie Proportional Symbol Maps in Superset erstellt und für raum-zeitliche Analysen verwendet werden können.

In diesem Vortrag wird das Open-Source-Tool Apache Superset vorgestellt, das als beliebte BI-Plattform vielseitige Datenvisualisierungen und interaktive Dashboards ermöglicht. Unsere Erweiterung fokussiert sich auf die Integration von Geodatenvisualisierungen und die Unterstützung raum-zeitlicher Analysen, was besonders für geografisch geprägte Anwendungen wie Stadtplanung und Umweltschutz wertvoll ist.

Durch unsere Implementierung können u.a. folgende thematischen Kartentypen in Superset genutzt werden: Kartodiagramme, Choroplethenkarten und Proportional Symbol Maps.

Mit diesen Kartentypen wird es möglich, zeitliche Trends und geografische Muster gleichzeitig zu analysieren. Der Vortrag zeigt praktische Anwendungsbeispiele, die das Potenzial der erweiterten Superset-Funktionalität für die Datenanalyse in der Business Intelligence.

FOSS-basierte Schnittstelle zum Management von Heritage BIM Modellen

Güren Tan Dinga, Laura Fernandez Resta

Die Integration von Heritage BIM in die Erhaltung historischer Gebäude ist insbesondere bei der zukünftigen Bewirtschaftung dieser von hoher Relevanz. Am Beispiel des UNESCO-Welterbes Speicherstadt Hamburg wurde eine benutzerorientierte Schnittstelle entwickelt, welche es Nutzer:innen ohne BIM-Kenntnisse ermöglicht, Zugang zu strukturierten Gebäudeinformationen zu gewähren. Dafür wurden offene Standards wie IFC und FOSS genutzt, wodurch die Interoperabilität und Kollaboration gefördert wurden.

Im Lightning Talk wird ein webbasiertes Dashboard (entwickelt mit Streamlit (Apache-2.0) und IfcOpenShell (LGPL-3.0)) vorgestellt, welches am Beispiel des Blocks H des UNESCO-Welterbes Speicherstadt Hamburg entwickelt wurde. Das Dashboard ermöglicht es, Nutzer:innen auf das HBIM-Modell zuzugreifen, ohne (proprietäre) Autorensoftware zu verwenden. Über das Dashboard können Gebäudeinformationen, abgelegt in sogenannten PropertySets, gelesen, exportiert und modifiziert werden. Der niedrigschwellige Zugriff auf strukturierte Gebäudedaten verbessert die kollaborative Zusammenarbeit und ermöglicht insbesondere das Verwalten von Aufgaben und Zuständigkeiten. So können Veränderungen an Gebäudeeigenschaften direkt durch die Verantwortlichen eingetragen werden, ohne die BIM-Modellierenden zu beauftragen. Dies ermöglicht das Einsparen eines Arbeitsschrittes, die Interpretation von Daten durch eine dritte Person und das Vermeiden von Fehlern.

Amtliche Orthofotos zentral verfügbar: Eine Open-Source-Lösung für Deutschland

Moritz Lucas

In Deutschland werden amtliche Orthophotos mittlerweile in allen Bundesländern frei zugänglich bereitgestellt, geregelt durch die INSPIRE-Richtlinie. Doch die föderale Struktur führt zu 16 unterschiedlichen Diensten, die den Zugriff auf diese hochwertigen Bilddaten erschweren. Die vorgestellte Software bietet eine zentrale Lösung, die den deutschlandweiten Download vereinfacht. Ziel ist es, mehr Aufmerksamkeit für Orthophotos zu schaffen und deren Nutzung sowie Weiterentwicklung zu fördern.

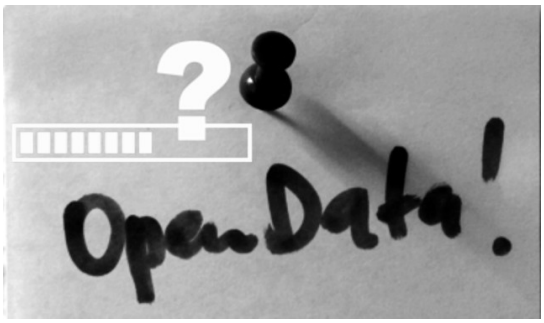
Orthophotos bieten durch hohe Bodenauflösung, spektrale Information und Lagegenauigkeit eine wertvolle Datenbasis für verschiedenste Anwendungsbereiche in der Fernerkundung, von Umweltforschung bis hin zur Stadtplanung. In Deutschland werden diese Aufnahmen im Zweijahresrhythmus erstellt und seit kurzem frei zugänglich gemacht, um der INSPIRE-Richtlinie zu entsprechen. Da Geobasisdaten in Deutschland Ländersache sind, wird die Aufnahme, Speicherung und Bereitstellung dieser Daten innerhalb der Bundesländer organisiert. Dies führt zu 16 unterschiedlichen Plattformen und Schnittstellen, die von modernen STAC-Katalogen bis hin zu WMS-Diensten reichen.

Dieses Projekt stellt eine Softwarelösung vor, die den deutschlandweiten Zugriff auf Orthophotos vereinfacht und Nutzer die Möglichkeit bietet, Bilder unabhängig vom jeweiligen Dienst herunterzuladen. Die Software unterstützt das Herunterladen von RGB- oder RGBI-Daten und ermöglicht flexible geografische Abfragen.

Diese Lösung richtet sich sowohl an Forschende als auch an Entwickler und zielt darauf ab, die Nutzung und Weiterentwicklung zu fördern. Die zentrale Bereitstellung und der vereinfachte Zugriff können dabei helfen, die föderal organisierten Daten für größere, regionsübergreifende Projekte leichter nutzbar zu machen und neue Möglichkeiten zur Datennutzung in Wissenschaft und Praxis zu erschließen.

Open Data in D: Perfekte Idee, halbherzige Umsetzung? Ein Erfahrungsbericht.

Mike Elstermann



Open Data ist gut, aber nur, wenn ...

Lange haben wir sie gefordert und darum gekämpft und zum Glück gibt es sie jetzt, diese deutschen frei verfügbaren Geobasisdaten, in allen Bundesländern, aber mit unterschiedlicher Ausprägung. Ein persönlicher Erfahrungsbericht soll den aktuellen Status sowie das Verbesserungspotenzial zeigen und zur Diskussion zwischen Nutzern und Anbietern anregen.

Ich möchte meine Wahrnehmungen zum aktuellen Status der deutschen offenen Geobasisdaten beschreiben. Diese basieren auf den persönlichen Erfahrungen bei der Erstellung und Befüllung des QGIS-Plugins "GeoBasis_Loader" [1], [2] mit den Open Data Quellen vor allem des Bundes und der Landesvermessungsämter. Vorgestellt werden der Status Quo, also, was ist schon richtig gut und wo sollte (dringend) nachgebessert werden. Der Vortrag soll als konstruktive Kritik verstanden werden mit dem Ziel, Mängel transparent und nachvollziehbar zu beschreiben und den Anbietern Input zu geben sowie für die nähere Zukunft zu klären, was wir als Nutzer fordern können und sollten? Wie also reagieren wir als FOSSGIS Community auf die aktuelle und zukünftige Situation?

Kriterien sind dabei beispielweise die Verfügbarkeit, die Auffindbarkeit, Layerstrukturen, Visualisierung, Projektionen, Maßstabsgrenzen, Stabilität, Performance, Formate sowie die noch fehlende Vereinheitlichung und vielleicht auch wünschenswerte Namenskonventionen.

Kontakt zum Autor:

Mike Elstermann
IT-Consult Halle GmbH (und #geoObserver)
Bornknechtstraße 5, 06108 Halle (Saale)
+49 581 7128
mike.elstermann@itc-halle.de

Literatur:

[1] https://geoobserver.de/qgis-plugin-geobasis-loader/

[2] https://plugins.qgis.org/plugins/GeoBasis_Loader/

QGIS-Werkzeuge und Python

Isabelle Korsch

Mit Python kann der Funktionsumfang von QGIS erweitert werden, aber es muss nicht immer ein QGIS-Plugin sein: Häufig lohnt es sich auch mit Python neue Werkzeuge in QGIS zu erstellen. Wie funktioniert das?

In diesem Einführungsvortrag wird zunächst die Python-Schnittstelle der QGIS-Werkzeugkiste vorgestellt und es werden Möglichkeiten und Grenzen der Werkzeugentwicklung diskutiert (zum Beispiel im Vergleich zum Model Builder und zur Pluginentwicklung).

Anschließend wird auf der Basis eines Templates in einem einfachen Codebeispiel ein neues QGIS-Werkzeug erstellt und gezeigt, an welchen Stellen Python-Code angepasst und was dabei beachtet werden muss.

Visualisierung von historischen Schiffsrouten mit unscharfer Datengrundlage

Stefan Fuest[1], Maximilian Herbers[2], Andreas Gollenstede[2], Jennifer Tadge[3], Rieke Marie Kaiser[4]

[1] Leibniz Universität Hannover, Institut für Kartographie und Geoinformatik (ikg)

[2] Jade Hochschule, Institut für Angewandte Photogrammetrie und Geoinformatik (IAPG)

[3] Landesmuseum Natur und Mensch Oldenburg (LMNM)

[4] Carl von Ossietzky Universität Oldenburg, FK IV, Institut für Geschichte (IfG)

Zusammenfassung

Im Forschungsverbund DiViAS [1] werden Quellen wie Logbücher und Journale aus dem 18./19. Jh. ausgewertet und analysiert. Im Fokus dieses Vortrags stehen dabei verschiedene Möglichkeiten der Visualisierung von Schiffsrouten mit unscharfer Datengrundlage, welche mit QGIS und Mapbox GL JS erstellt werden. Grundlage für die kartographische Darstellung ist eine KI-gestützte Extraktion von unscharfen Orts- und Zeitangaben aus diesen Quellen und deren Modellierung in einer PostgreSQL-Datenbank.

Schlüsselwörter: Kartographie, raumzeitliche Unschärfe, Datenmodellierung, transdisziplinäre Analyse, Logbücher, NLP, koloniale Kontexte

Im Forschungsverbund DiViAS werden bislang kaum verknüpfte wissenschaftliche Methoden und Praktiken in der Digitalisierung, Forschung und Repräsentation von Sammlungsgut aus kolonialen Kontexten und deren Bewegungen in einem zukunftsweisenden transdisziplinären Forschungsprofil systematisch zusammengeführt. Den Ausgangspunkt bilden die großen Sammlungen und archivalischen Überlieferungen des Landesmuseums Natur und Mensch Oldenburg und des Akademie-Projekts Prize Papers, deren Entstehungskontexte eng mit europäischer Expansion und Kolonialismus verbunden sind. Bei den zu visualisierenden Schiffsrouten handelt es sich zum einen um deutsche Strafexpeditionen in der Bismarcksee im heutigen Papua-Neuguinea aus dem 19. Jahrhundert und zum anderen um Handels- und Kaperfahrten europäischer Seemächte aus dem 18. Jahrhundert (siehe Abb. 1). Die aufgezeichneten Logbücher, Berichte und Journale werden transkribiert und mithilfe von Sprachmodellen und Verfahren des Natural Language Processing, wie der Named Entity Recognition, nach Orts- und Zeitangaben durchsucht. Zusätzlich werden räumliche und zeitliche Präpositionen wie z.B. „nach“ oder „bei“ extrahiert.

Eine große Herausforderung stellt die inhärente Unschärfe dar, die sich sowohl auf Geometrien (z. B. Schiffsrouten, Orte) als auch auf Zeitangaben bezieht. Ursächlich für die Unschärfe können beispielsweise Messfehler bei der Positionsbestimmung, ungenaue Angaben oder Datenlücken sein. Die Unschärfe in Standortbeschreibungen kann, abhängig von der textlichen Beschreibung, vielfältige Ausprägungen annehmen. Daher werden zunächst sogenannte Möglichkeitsräume definiert, die die gesamte räumliche Ausdehnung der entsprechend einer Ortsbeschreibung möglichen Standorte umfassen. Zur visuellen Kommunikation von Unschärfe werden anschließend mittels QGIS Ansätze der kartographischen Symbolisierung unter Nutzung visueller Variablen (z. B. Blurring, Transparenz) herangezogen [2-4] (siehe Abb. 2), aber auch Möglichkeiten dynamischer Effekte wie Animationen exploriert [5].

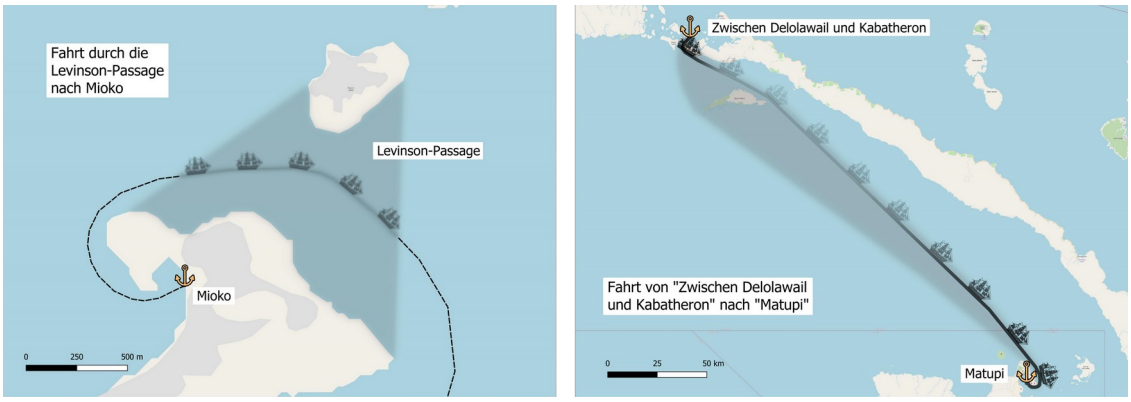


Abb. 2: Visualisierungsbeispiele für die Darstellung räumlicher Unschärfe mithilfe von Möglichkeitsräumen und der visuellen Variablen Blurring und Transparenz. Quelle: Eigene Erstellung mit QGIS.

Um den unterschiedlichen Ansprüchen verschiedener Nutzergruppen bei der Erkundung der visualisierten Informationen gerecht zu werden, werden unterschiedliche kartographische Aufbereitungen der Inhalte auf OSM-Basis aber auch unter Verwendung einer frei zugänglichen historischen Kartenbasis genutzt. So ergeben sich z. B. durch die Anwendung narrativer Techniken beim kartenbasierten Storymapping mit Mapbox GL JS unter Nutzung der Schnittstellen zum raumzeitlichen Datenmanagement mit PostgreSQL/PostGIS neue Möglichkeiten der Vermittlung von historischen Zusammenhängen. Im Rahmen des Vortrags werden verschiedene Arten der kartographischen Repräsentation von raumzeitlicher Unschärfe gezeigt und deren Anwendbarkeit in verschiedenen Nutzungskontexten diskutiert.

Kontakt zum Autor:

Stefan Fuest
stefan.fuest@ikg.uni-hannover.de
Leibniz Universität Hannover
Institut für Kartographie und Geoinformatik (ikg)

Literatur

[1] DiViAS. 2025. „Digitalisierung, Visualisierung und Analyse von Sammlungsgut (DiViAS).“ In divias.de, https://divias.de (zugegriffen: 30. Januar 2025).

[2] Hildebrand, H. et al. 1979. Die deutschen Kriegsschiffe: Biographien: ein Spiegel der Marinegeschichte von 1815 bis zur Gegenwart. Band 1. Herford: Koehler.

[3] Kinkeldey, C., MacEachren, A. M., & Schiewe, J. 2014. „How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies.“ In The Cartographic Journal, 51(4), 372-386.

[4] MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., & Hetzler, E. 2005. „Visualizing geospatial information uncertainty: What we know and what we need to know.“ In Cartography and Geographic Information Science, 32(3), 139-160.

[5] Lobben, A. 2003. „Classification and application of cartographic animation“ In The Professional Geographer, 55(3), 318-328.

[6] Baudissin, Friedrich. 1886. „Bericht über die Thätigkeit im Bismarck-Archipel in Erledigung von Requisitionen ausgeführter Landungen, Kämpfe und Gefechte“ In Bundesarchiv, RM1/2625.

Wie MapLibre und Vektorkarten die Welt übernehmen

Bart Louwers, Just van den Broecke

Vektorbasierte Karten sind die Zukunft! Oder vielleicht sogar schon die Gegenwart? In diesem Vortrag werden beide Perspektiven beleuchtet! Bart spricht aus der Perspektive der Entwicklung von MapLibre, und gibt einen Einblick in den neuesten Stand. Just erzählt von seinen Erfahrungen als Benutzer der MapLibre-Stack um vektorbasierte Karten zu gestalten.

Bart Louwers, der seit zwei Jahren Maintainer von MapLibre Native ist, gibt ein Update über spannende neue Entwicklungen bei MapLibre. In der C++-Version von MapLibre stehen jetzt nicht nur eine, sondern zwei neue Rendering-Backends zur Verfügung (Vulkan und Metal). Seit Version 5.0.0 unterstützt MapLibre Web zudem die Darstellung der Erde als Globus.

Just van den Broecke, ein erfahrener niederländischer Kartograph, spricht über die neuen Herausforderungen, die die Erstellung von vektorbasierten Karten mit dem MapLibre-Stack mit sich bringt – im Gegensatz zur Arbeit mit Rasterkarten in Tools wie Mapnik.

Open Data des BKG (II)

Joachim Eisenberg

Das Bundesamt für Kartographie und Geodäsie (BKG) hat ein breites Open Data-Angebot, das auf der letzten FOSSGIS-Konferenz in Hamburg vorgestellt wurde. Was hat sich mit der Umsetzung der HVD-Verordnung an diesem Angebot getan?

Im Vortrag wird kurz die Durchführungsverordnung 2023/138 zur Festlegung bestimmter hochwertiger Datensätze und der Modalitäten ihrer Veröffentlichung und Weiterverwendung (HVD-Verordnung) vorgestellt. Ergänzend wird geschaut, ob durch sie neue „High-Value Datasets“ als Open Data zur Verfügung gestellt wurden und werden.

Schnupperkurs: Das Potential von QGIS mit der Python-Konsole freischalten

Gordin Schlolaut

Programmierung mit Python ermöglicht es, sich eigene Werkzeuge in QGIS zu erstellen und so den Funktionsumfang von QGIS individuell zu erweitern und den eigenen Bedürfnissen anzupassen. Aber wie funktioniert das alles eigentlich und womit fängt man an? Dieser Vortrag gibt einen kurzen und verständlichen Überblick auch für alle ohne Programmier-Vorkenntnisse. Wir werden live ein kleines Werkzeug schreiben, an dem wir die grundlegenden Konzepte vorstellen und zeigen: es ist gar nicht so schwierig.

Programmierung mit Python ermöglicht es, sich eigene Werkzeuge in QGIS zu erstellen und so den Funktionsumfang von QGIS individuell zu erweitern und den eigenen Bedürfnissen anzupassen. Aber wie funktioniert das alles eigentlich und womit fängt man an? Dieser Vortrag gibt einen kurzen und verständlichen Überblick auch für alle ohne Programmier-Vorkenntnisse. Wir werden live ein kleines Werkzeug schreiben, an dem wir die grundlegenden Konzepte vorstellen und zeigen: es ist gar nicht so schwierig.

Automatisierte Verarbeitung von Daten der Meeresbodenkartografie mit QGIS

Dr.-Ing. Helge Städtler, PlanBlue GmbH, Bremen

Für die Entwicklung skalierbarer Karteninhalte, die z.B. validierte, lokal vergleichbare Informationen zu Umweltverschmutzung enthalten, setzt PlanBlue auf eine Datenverarbeitung mit fortschrittlichen Algorithmen der Data Science und maschinengestütztes Lernen (ML), um insbesondere hyperspektrale Daten in aussagekräftige Indikatorkarten zu konvertieren.

PlanBlue ist ein junges Unternehmen aus Bremen im Bereich innovativer, maritimer Umwelttechnologie, dessen Fokus darauf liegt, hochdetaillierte, georeferenzierte Karten des Meeresbodens zu erstellen.

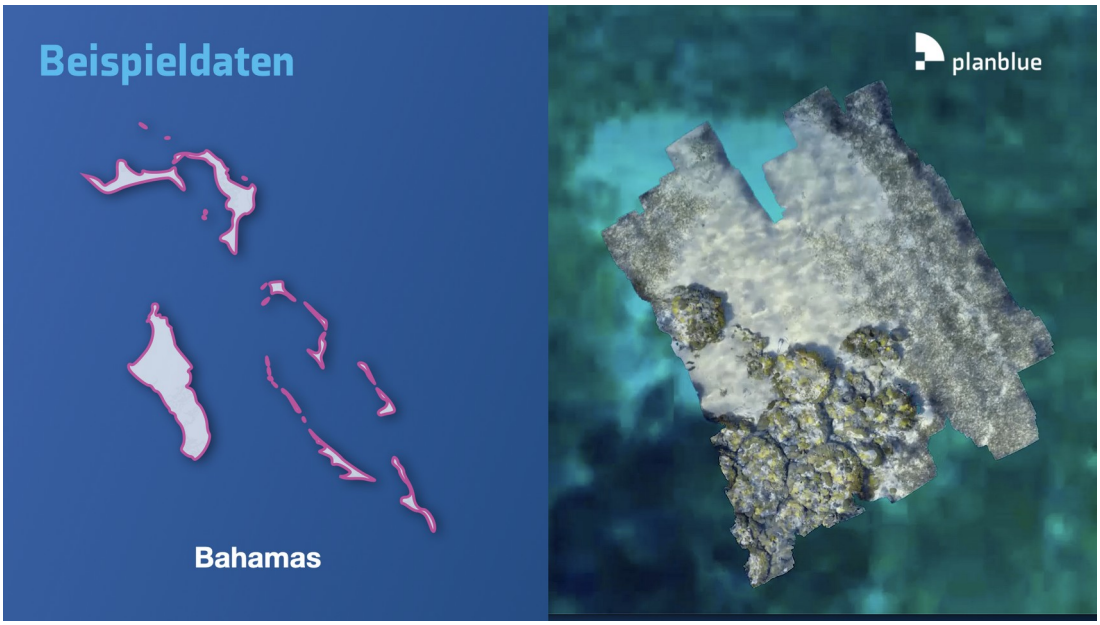


Abb. 1: Meeresboden vor den Bahamas

Um zu veranschaulichen, um welche Daten es im Folgenden geht, zeigt Abbildung 1, links, eine Umrissübersicht der Inselgruppe der Bahamas und rechts eine Basiskarte aus Satellitenfotos mit einem darübergelegten Ausschnitt des in der Region kartografierten Meeresbodens [1].

Es wird das Ziel verfolgt, standardisiert, vollautomatisch, schnell und in großem Maßstab die Vorteile der großflächigen Kartierung durch robotische Erkundungstechnologien, z.B. Autonome Unterwasser Vehikel (AUV), mit der hohen Auflösung und der Genauigkeit der In-situ-Überwachung zu kombinieren. Die daraus resultierenden Geodaten sollen eingesetzt werden, um u.a. Naturschutzbemühungen in den Ozeanen zu unterstützen.

Rasterdaten

Der überwiegende Datentyp bei der Meeresbodenkartografie sind Rasterdaten in Form von RGB-Bildern. Die Bilder werden als Einzelbilder oder Film aufgenommen und anschließend in mehreren Zwischenschritten zu einer Fläche zusammengesetzt. Diese Zwischenschritte enthalten u.a. Belichtungskorrekturen und einen nicht-trivialen Prozess der Projektionskorrektur, um für jeden Rasterpunkt einen

echten 3D-Pixel zu erhalten. Als Zwischenprodukt entstehen **GEOTIFF-Dateien**, die zusammenhängende Flächen des Meeresbodens abbilden.

Automatisierung

Wiederkehrender Schwerpunkt in der Verarbeitung von Rasterdaten des Meeresbodens ist eine möglichst umfangreiche Automatisierung bei der Weiterverarbeitung dieser Zwischenprodukte. Es werden viele GEOTIFF-Dateien in kurzer Zeit generiert, wenn man großflächig den Ozeanboden mit verschiedenen Kamerasystemen (RGB und Hyperspektral) erfasst.

Um die Geodaten für den Gebrauch zugänglich zu machen, haben sich in der Geodaten Community einige Lösungen als sehr praktikabel und nützlich erwiesen. Für einen Fachexperten, der es gewohnt ist, mit Geoinformationssystemen (GIS) zu arbeiten, ist die Hürde gering, Rasterdaten in sein Lieblings-GIS zu laden, um anschließend passgenaue Analysen durchzuführen oder spezifische Aspekte der Daten herauszustellen und in eine ansprechende Abbildung zu überführen. Um die Daten allerdings auch für nicht-fachspezifische Experten zugänglich zu machen, ist die Transformation in einfacher lesbare und interpretierbare Datenformate notwendig.

Die folgenden Lösungen bieten sich hier an:

- **QGIS Druck-Layout Vorlagen** [2] anfertigen und Rasterdaten automatisiert in eine vollständig georeferenzierte und beschriftete Kartenabbildung überführen
- Aus Rasterdaten automatisiert ein Tile Matrix Set [3] zu erstellen, das die Basis bildet um einen **Web Map Tile Service (WMTS)** zu betreiben oder eine statische, webbrowserbasierte, interaktive Karte bereitzustellen

Beide Lösungen wurden in der Datenverarbeitung bei PlanBlue umgesetzt und sind in der Praxis im wiederkehrenden Einsatz.

Georeferenzierte Kartenabbildung

Präsentiert man kartografierten Meeresboden ausschließlich als GEOTIFF-Datei, fehlen jedoch viele wichtige Angaben.

Dazu zählen u.a.:

- Projektion
- Nordausrichtungsindikator

- Skala bzw. Maßstab
- Koordinatensystem
- Datum der Datenerhebung
- Betitelung
- geographische Ortsangabe
- Angaben zum Copyright

QGIS bietet für die Unterbringung dieser Angaben die Möglichkeit, ein sogenanntes **Druck-Layout** zu erstellen. Hierfür öffnet man die Layout-Verwaltung und erstellt zunächst ein neues Layout [4].

Stück für Stück kann man nun mit den zur Verfügung stehenden Layout-Elementen (z.B. Map, Picture,

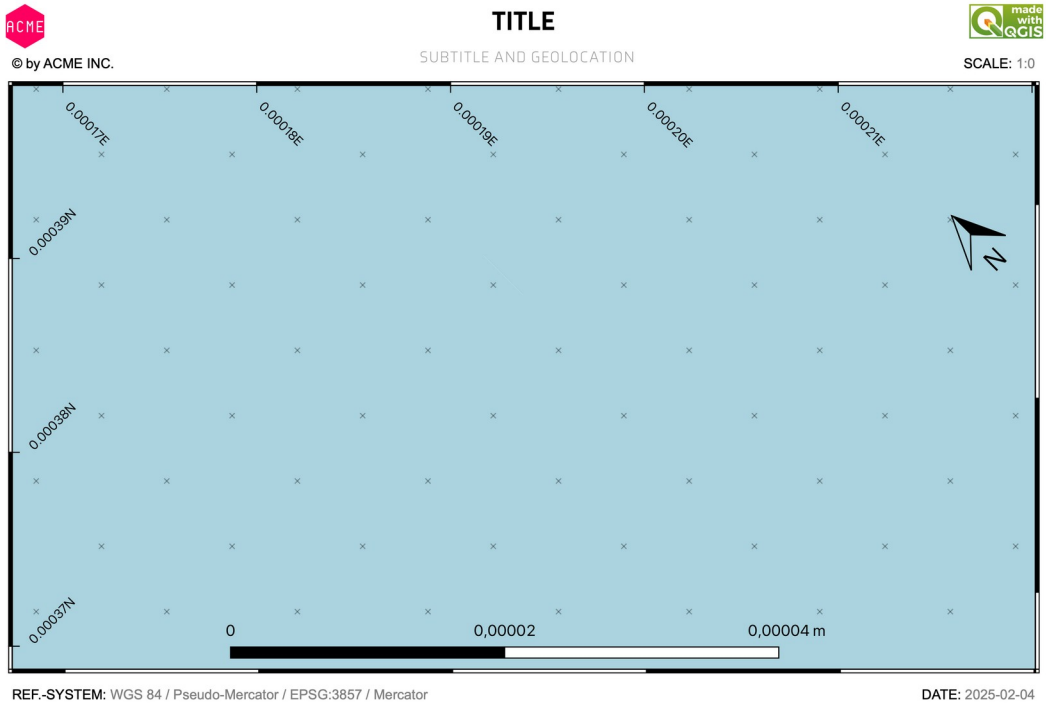


Abb. 2: Druck-Layout als Beispiel mit den Elementen Map, Picture, Scalebar und Label

Scalebar und Label) eine Vorlage gestalten, die alle oben genannten Angaben enthält (Abbildung 2). Als wichtigstes Element bindet man die zu zeigende GEOTIFF-Datei als Rasterlayer aus QGIS in einem Map-Element ein.

Ob das Layout optisch geeignet ist, das gewünschte Präsentationsergebnis zu erzeugen, kann man jederzeit prüfen, indem man einen PDF-Export mit dem bearbeiteten Layout durchführt. Ist man mit dem Darstellungsergebnis des Druck-Layout zufrieden, speichert man dieses. Das gespeicherte Druck-Layout findet sich anschließend im Arbeitsverzeichnis als „QGIS Composer Template“ bzw. als Datei mit der Endung „.qpt“ und enthält die Beschreibung des Layouts, formuliert in der Extensible Markup Language (XML).

Modifizierung des Druck-Layout

Um mit dem erstellten Druck-Layout eine automatisierte Erstellung von z.B. PDF-Dokumenten zu erreichen, ist der Rückgriff auf die **PyQGIS-Schnittstelle** notwendig. Wie man hier einen guten Einstieg

finden kann, ist in dem Tagungsbeitrag *„QGIS fernsteuern: Kickstart mit PyQGIS zur Automatisierung“* nachlesbar [5].

Für eine gelungene Automatisierung kann man nachfolgend beschriebene Schritte (Beschreibung als Pseudocode) nacheinander ausführen und diese Sequenz immer wieder von der Kommandozeile aus mit anderen Eingabedaten wiederholen. Als Eingabeparameter wird eine stets wechselnde GEOTIFF-Datei angenommen, die zentraler Bestandteil der Druck-Layout-Ausgabe werden soll.

Folgende Schritte sollten programmatisch per PyQGIS-Skript erfolgen (Pseudocode):

- Anlegen eines Projektes
- Erstellen einer RasterLayer mit der darzustellenden GEOTIFF-Datei
- Hinzufügen der RasterLayer als MapLayer
- Setzen des CoordinateReferenceSystem (CRS) für das Projekt, basierend auf dem CRS, das in der GEOTIFF-Datei hinterlegt ist
- Erstellen und Konfigurieren eines MapCanvas
- Setzen des „extent“ des MapCanvas auf das „extent()“ der RasterLayer
- Erstellen eines PrintLayout für das bereits erstellte Projekt
- Lesen/Laden des Inhaltes des gespeicherten Druck-Layout (.qpt-Datei)
- Iterieren über alle LayoutItems des geladenen PrintLayout und Einsetzen der Werte für z.B. Label die in der Ausgabe verwendet werden sollen
- Das geladene und modifizierte PrintLayout mit „addLayout()“ zum „layoutManager“ des Projektes hinzufügen
- Dem Map-Element des PrintLayout die erzeugte RasterLayer zuordnen
- Dem Map-Element des PrintLayout das CRS des Projektes zuordnen
- Dem Map-Element mit „zoomToExtent()“ den „extent“ des MapCanvas zuordnen
- Einmal „refresh()“ aufrufen für das Map-Element
- Dem PrintLayout als „referenceMap“ das Map-Element zuordnen
- Einmal „refresh()“ aufrufen für das PrintLayout
- Erstellen eines LayoutExporter und diesem das PrintLayout zuordnen
- Den LayoutExporter mit Hilfe von PdfExportSettings konfigurieren (dies erlaubt z.B. die Angabe der Auflösung in DPI)
- Das fertig konfigurierte PrintLayout exportieren mit Aufruf von „exportToPdf()“
- Das gesamte Projekt speichern

Zusammenfassung

Durch die Kombination des QGIS Druck-Layouts mit der dynamischen Modifikation dieses Layouts über ein Skript das die PyQGIS-Schnittstelle nutzt, kann man eine GEOTIFF-Datei automatisiert in ein vorzeigbares PDF-Dokument umwandeln.

Kontakt zum Autor:

Dr.-Ing. Helge Städtler
Holbeinstr. 10, 28209 Bremen
hs@appdoctors.de

Literatur

[1] Beispiel eines georeferenzierten Datenprodukts über die Bahamas, PlanBlue GmbH

[2] QGIS 3.34 Documentation, Laying out the maps, https://docs.qgis.org/3.34/en/docs/user_manual/print_composer/index.html#laying-out-the-maps, Abruf im Feb 2025.

[3] OpenGIS® Web Map Tile Service Implementation Standard: https://portal.ogc.org/files/?artifact_id=35326, Abruf im Feb 2025.

[4] QGIS 3.34 Dokumentation, Layoutverwaltung, https://docs.qgis.org/3.34/de/docs/user_manual/print_composer/overview_composer.html#the-layout-manager, Abruf im Feb 2025.

[5] „QGIS fernsteuern: Kickstart mit PyQGIS zur Automatisierung“, Tagungsband FOSSGIS 2025

Vektor Tiles für Karten mit Echtzeitdaten

Pirmin Kalberer

Vektorkacheln sind eine effiziente Art, um Karten mit grossen Mengen an Echtzeitdaten bereitzustellen.

IoT-Anwendungen oder Karten mit Wetter-, Verkehrs- oder anderen Live-Daten benötigen ein kompaktes Format zur Übertragung aktualisierter Daten. Vektorkacheln mit ihren leistungsfähigen Styling-Möglichkeiten werden von diversen JavaScript-Kartenviewern unterstützt.

Dieser Vortrag zeigt, wie man Echtzeitdaten mit MapLibre und OpenLayers darstellen kann. Auf der Serverseite wird BBOX verwendet, um Kacheln aus einer PostGIS-Datenbank bereitzustellen.

Small seeds - FOSS Communities stärken!

Paul Robben

Mehr finanzielle Mittel für Freie und Open-Source-Software – die unendliche Geschichte. In einer von Start-ups, VC Funding und datensammelnden Apps geprägten Tech-Welt ist der Kampf um nachhaltige Finanzierungen für ethische Technologieprojekte besonders hart. Nach einigen großen Erfolgen für die Förderung von FOSS in den letzten Jahren geht es in diesem Vortrag nun darum, die kleinen, weniger sichtbaren zivilgesellschaftlichen Projekte nicht zu vergessen.

Wie kann man Technologie nachhaltig finanzieren? Diese Antwort hören wir immer wieder: Infrastruktur und Maintenance fördern, und bitte unbedingt auch dieses eine Projekt unterstützen, das jemand in Nebraska ohne Dank seit 2003 am Laufen hält. Das ist fraglos alles sehr wichtig. In diesem Vortrag wollen wir jedoch der Frage nachgehen, warum eine vielfältige Förderlandschaft, die auch die Unterstützung neuer Zielgruppen mit frischen Ideen ermöglicht, extrem wertvoll für FOSS-Communities sein kann.

Wie können wir bestehende Förderstrukturen nutzen, sie so zurechtbiegen und anpassen, dass sie den tatsächlichen Bedürfnissen der Communities gerecht werden? Wie können wir diese Strukturen sinnvoller für förderferne Projekte und Personen gestalten? Wie können wir Förderinfrastrukturen so aufbauen, dass sie mehr Vielfalt, neue Ideen und Inklusion in unsere Communities bringen? Und wie fördern wir dabei möglichst nachhaltig?

Hintergrund dieses Vortrags sind unsere Erfahrungen bei der Umsetzung des deutschen FOSS-Förderprogramms Prototype Fund, in dessen Rahmen wir gemeinsam mit dem Bundesministerium für Bildung und Forschung in den letzten acht Jahren fast 400 Projekte finanzieren konnten. Von Frontend-Apps wie StreetComplete bis zu Hintergrundinfrastruktur wie OSM2PGSQL durften wir dabei auch zahlreiche Projekte aus der OSM Community unterstützen. Nach einer umfassenden internen und externen Evaluierung haben wir angesichts des Auslaufens der aktuellen Förderperiode im vergangenen Jahr eine neue, nachhaltigere Förderstruktur für sogenannte „early stage“-Innovationsprojekte konzipiert und aufgebaut.

Dieser Vortrag ist ein Aufruf an staatliche Institutionen, Geldgeber und andere Förderorganisationen, ihre Kräfte zu bündeln, über die Grenzen ihrer traditionellen Fördermodelle zu treten und ein breites und dynamisches Netzwerk von kleinen, vielfältigen und leichtgewichtigen Fördertöpfen zu schaffen, die den Bedürfnissen verschiedener Gruppen und Communities gerecht werden. Es ist eine Einladung an die Communities, gemeinsam ihre Bedürfnisse zu formulieren, um Strukturen aufzubauen, die ihre Arbeit auch wirklich unterstützen können. FOSS-Projekte, egal ob gut etabliert oder gerade erst im Entstehen, ob groß oder klein, bergen Hoffnung auf eine bessere Welt. Wir wollen ihnen die Unterstützung geben, die sie brauchen.

Von Proprietär zu QGIS

David Arndt

Regionalverband Ruhr

Open-Source-Software im GIS-Bereich ist beim Regionalverband Ruhr seit Jahren etabliert. Ein Ausbau hin zu mehr freier Software und Abbau von proprietärer Software wird im Sinne der digitalen Souveränität vorangetrieben.

Die voranschreitende Digitalisierung in der Öffentlichen Verwaltung führt dazu, dass immer mehr Verwaltungsmitarbeitende die Anforderung eines leistungsfähigen GIS an ihrem Arbeitsplatz haben. Der dafür nötige Ausbau der Lizenzen proprietärer Software stand dabei aufgrund von Wirtschaftlichkeitsbetrachtungen nicht im Fokus.

Die Übertragbarkeit von Lösungen auf unsere Mitgliedskommunen, ohne dass hier weitere Lizenzkosten entstehen, sind ein weiterer zentraler Mehrwert.

QGIS hat sich so in den letzten Jahren beim Regionalverband Ruhr und einigen Mitgliedskommunen zum Standard-GIS entwickelt.

Eine Migration zu einer neuen Software bedeutet aber auch immer eine intensive Planung und Kommunikation. Die Erfahrungen, die beim Regionalverband Ruhr gemacht wurden, sollen hier geteilt werden.

Folgende Punkte sollen unter anderem betrachtet werden:

- Wirtschaftlichkeit
- Nutzungsszenarien
- Plugins
- Verknüpfung mit mobilen Anwendungen

Aber auch Grenzen und wo QGIS noch Nachholbedarf hat, sollen dargestellt werden.

Kontakt zum Autor:

Dipl. Geogr. David Arndt
Regionalverband Ruhr
Kronprinzenstraße 6
45128 Essen
0201/2069-412
arndt@rvr.ruhr

Literatur

[1] QGIS-Projektseite: https://qgis.org

[2] QGIS-DE Anwendergruppe Deutschland e.V.: https://qgis.de

Open Source and Commercial Large Language Models comparison:

The performance capability and limitation of Open Source Large Language Models for geoparsing tasks

Authors: Juiwen Chang[1], Cheng-Chia Wu[2], Jörg Rainer Noennig[3]

[1] Cooresponding author, Digital City Science (DCS), Hafencity Universität Hamburg (HCU)

[2] Department of Geography, National Taiwan University

[3] Chair of Digital City Science (DCS), HCU. https://digital-city-science.de/team/

To cite this paper: Chang, J.; Wu, C; Noennig, J. (2025) The performance capability and limitation of open source large language models for geoparsing tasks. FOSSGIS 2025 Tagungsband.

https://doi.org/10.5281/zenodo.14616472

Summary:

Geoparsing is a process that combines *Name Entity Recognition* (NER) and *Geocoding* to process free text and identify geolocalized terms in them, which are then converted into spatial coordinates using Geocoder API. On such premise, here presented research project is part of a cooperation with the Real Estate Management Corporation of the city of Hamburg (LIG) and pursues the overarching goal of developing a geoparsing workflow for processing the open source database of the Hamburg Parliament. Specifically, the aim is to link and visualize neighbourhood land parcels that are part of public discussion - for which a large number of PDF files are to be processed and visualized on a WebGIS system. For this aim, we are now investigating the functional capabilities and limitations of Large Language Models (LLM). In order to accurately estimate how many place terms are recognized, we have developed a validation workflow (including a process to spatially join multiple geocoded point coordinates) that automatically compares and calculates the percentage of place terms recognized by each model. With the rapid development of LLM, it has become easy to extract corresponding location lists - but the quality of results has remained an open question. This research paper therefore focuses on the performance of open-source LLMs, of which we have tested Meta *Llama3.1-70b-instruct* and *Mistral-large-2* in order to compare them with *Claude-sonnet-3.5-new* and OpenAI *o1-preview*. For the geocoding process the geocoders *Nominatim* and *Mapbox Geocoder API* are used. For addresses with house numbers, we use the official address sources published by the State Corporation For Geoinformation and Surveying (LGV). The results show that a hybrid approach of combined LLM can extract up to 74,98% of unique locations from the text, including the full address with house numbers while Claude, on the other hand, only achieves 53%. We also found that open source LLMs are limited in extracting the names of institutions (e.g. high schools) without giving concrete address information from the texts. Such edge cases will be further illustrated in the discussion session.

Keywords:

Geoparsing, Name Entity Recognition, Geocoding, Llama, Mistral, Nominatim

1. Introduction and previous research

Geoparsing is a process that combines Name Entity Recognition (NER) and Geocoding to process free-text and identify geolocation terminology, and further uses geocoder API to convert from location

texts into coordinates. revious research by Ventrice & Di Caro (2023) [6] explored text extraction from Wikipedia data, focusing on the public knowledge graph's spatial accuracy,emphasizing advancements in NER to enhance location-based information accuracy. There are many ways to detect location / addresses / place names from long texts. When applying to German texts, the most popular approaches are the *flair-german-large* and *BERT* model (Tao et.al, 2022) [5]. Here represent research cooperation project, we are jointly working on scientific context with Hamburg´s Real Estate Management Corporation (Landesbetrieb Immobilienmanagement und Grundvermögen, LIG) focusing on developing a **geoparsing workflow** for processing data from Hamburg´s open source Parliament Database (Dally et.al, 2024) [3]. To discover public discussions related to real-estate land properties, we aimed at processing large amounts of PDF files and visualizing them on a webGIS platform developed by HCU chair of Digital City Science in cooperation with LIG – the AGORA tool. In an earlier project phase, the *flair/ner-german-large* [7] model was used to extract locations from texts (Mellow & Chang, 2023) [4]. In this research, the focus of the investigation was on an exemplary local newspaper, "Elbewochenblatt"; the geoparsed results (including topic modeling) were visualized as point clusters in the AGORA toolbox [8].

2. Methodologies

a. Geoparsing Workflow

In order to support the routine tasks of the LIG workers, a powerful geoparsing workflow is required, with the focus on 1) location recognition including house numbers, and 2) accurate geocoding of the spatial map coordinates. With the rapid growth of LLMs, we can easily prompt the systems in order to extract location lists.

Below is a prompt example :

Extract all place names or addresses in Hamburg from the input text and enter them in the following format: 'Street name, house number, district'. Please note:

- Each address should be in separate lines.
- If there are several house numbers on the same street, create separate lines with the complete address for each house number.
- Terms that represent both the location and the name of an organisation should be included in the result list.

- Only extract locations and addresses in Hamburg.
- Skip the terms of 'Hamburg' or 'Free and Hanseatic City of Hamburg'.
- The output text should not contain any irrelevant words or text phrases.
- Only use the information available in the input text.

Conversely, we found that the flair model is not able to identify complete addresses including house numbers. Therefore, we investigated whether modern LLM can carry out the same task and checked - in order to better evaluate the results - how many location terms are eventually detected. Figure 1 shows the implementation diagram of our geoparsing workflow:

Open Source and Commercial Large Language Models comparison:

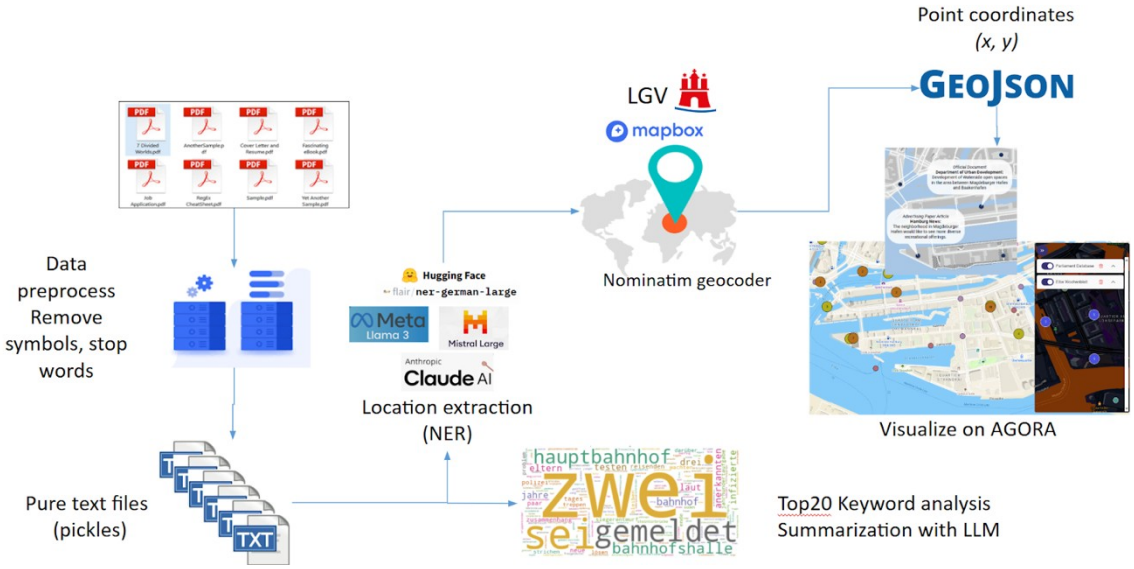


Figure 1 Geoparsing workflow

For the geocoding tasks, we have integrated three approaches. For those addresses in Hamburg having house numbers, we adapt the official published address from the city of Hamburg (Landesbetrieb Geoinformation und Vermessung, LGV). For the remaining addresses, Mapbox geocoder API and the open source geocoder Nominatim were used.

b. NER Validation

To achieve significant savings in terms of human labour costs and to address the challenge of the fast iterated versions of LLMs, we have developed a validation workflow that includes a process for geocoding and spatial joining of the same location term from different model results. This workflow enables the automatic comparison and calculation of the percentage of location terms that have been detected from each LLM model. We have deployed Ollama [9] on High Performance Computing (HPC) clusters [10] to run the open source LLM Mistral-large-2 and LLama 3.1-70b-Instruct.

We have further selected 33 smaller parliamentary requests (Schriftliche Kleine Anfragen, SKA) which include 1,612 locations as samples. Many of them include long appendices that contain most locations and addresses. We have further invited German native speakers to label these texts (as human labels), and deployed different LLM to process the text, document for each location which model successfully detected it, and further join these locations from different models into the same row. We have aligned our validation table as shown in table 1.

Table 1 Validation table

human_label	claude	llama	mistral	o1_preview	Check
Alsterdorfer Straße 39			Alsterdorfer Straße 39	Alsterdorfer Straße 39	TP
Alter Postweg 30	Alter Postweg 30			Alter Postweg 30	TP
Anna Susanna-Stieg 3					FN
Auf dem	Auf dem	Auf dem			TP

Open Source and Commercial Large Language Models comparison:

Sülzbrack 2	Sülzbrack 2	Sülzbrack 2			
Außenmühlenweg 23a	Außenmühlenweg 23a		Außenmühlenweg 23a	Außenmühlenweg 23a	TP
Baakenallee 33					FN
Bad Billstedt Archenholzstraße 50a			Archenholzstraße 50a		TP
Dratelnstraße 21	Dratelnstraße 21	Dratelnstraße 21		Dratelnstraße 21	TP
Georg-Wilhelm-Straße 112	Georg-Wilhelm-Straße 112	Georg-Wilhelm-Straße 112	Georg-Wilhelm-Straße 112	Georg-Wilhelm-Straße 112	TP
		Sozialtherapeutische Anstalt			FP

Column “Check” was first given by automatic script and further checked by a second human validator. TP stands for “True-Positive” and means that both human and LLM have determined this term as location. FN stands for “False-Negative”, meaning that the LLM did not detect this actually existing location; and FP stands for “False-Positive”, meaning that the LLM extracted a location while humans determined this being no location. We do not have TN “True-Negative” information, since we did not preserve texts that are not locations from prompt results.

To align this validation table, it would take a lot of copy and paste time from several location lists of different LLM models. Therefore, we developed a spatial joining approach, that is able to 1) automatically geocode location or address into coordinates, and 2) *spatially outer join* every two of them, with a predefined joining buffer of 5 meters radius. In this approach, when human labels e.g. “Alster-Schwimmhalle Ifflandstraße 21” but LLM only extracts “Ifflandstraße 21”, this location will be labeled as TP. There are also scenarios where: 1) a location can not be geocoded or geocoded outside of Hamburg, and 2) no specific location is given, e.g. “Elbepark” or “Senatsparkplatz” - such edge cases we have eliminated from our validation table.

To better follow European GDPR regulation, before processing commercial LLM we removed all person names from the text. Additionally, the GPU cluster center processing open source LLM is within the EU and is certified with data protection regulation.

3. Results

For ChatGPT o1-preview results, we were unable to process all 33 SKAs due to HCU currently developing new AI regulations and guidelines, that prevented a subscription to OpenAI. Hence, we were only able to process 1 SKA. With Claude from Anthropic AI, however, we were able to process all 33 sample datasets. Mistral and Llama we processed on Karolina HPC cluster with Ollama. For each model, we found that smaller models (such as Mistral-8b or Llama 3.1-8b) are much more likely to produce hallucinations in the prompt result. They also have disadvantages as they easily ignore middle contexts and only extract location from the beginning and the ending of a paragraph. Therefore, we only choose large models to process SKAs.

To tackle hallucination, it is necessary to program a post-processing script to remove location that do not exist in SKAs, so that these false addresses will not be geocoded.

Table 2 illustrates how many percent of the location was extracted for 33 sample SKAs.

Table 3 represent both flair-german-large and o1-preview model, only 1 SKA were evaluated.

Open Source and Commercial Large Language Models comparison:

Table 2 Model results I (33 SKAs, 1612 rows)

Models / Fundamental Metrics tests	**Llama3.1-70b -instruct**	**Mistral Large-2**	**Claude 3.5-Sonnet**	**Hybrid (without deepseek)**	**Deepseek R1:70b**
Sensitivity TP / (TP + FN)	34,16 %	55,10 %	48,90 %	68,41 %	48.15%
Precision TP / (TP + FP)	95,36 %	91,81 %	93,90%	87,52%	87.38%
F1 Score	0.5030	0.6880	0.6431	0.768	0.621

Table 3 Model results II (1 SKA, 341 rows)

Models / Tests	**Claude 3.5-Sonnet**	**Llama3.1-70b-instruct**	**Mistral Large-2**	**Flair-german-large**	**ChatGPT o1-preview**	**Deepseek R1:70b**
TP	213	46	42	58	175	69
FP	11	18	4	615	29	63
FN	128	295	299	283	166	194
Sensitivity TP / (TP + FN)	62,5 %	13,49 %	12,31 %	17,01 %	51,32 %	26,24%
Precision TP / (TP + FP)	95,1 %	71,9 %	91,3 %	8,61 %	85,8 %	52,3%
F1 Score	0.7539	0.2272	0.2171	0.1144	0.6422	0.349

Notes:

F1 Score = 2 * (Precision * Sensitivity) / (Precision + Sensitivity)

Interpretation of F1: Moderate: 0.4-0.6; Good: 0.6-0.8; Excellent: 0.8-1 ; All formula reference Carterette, B. (2009) [2]

Table 3 represents Claude as the best model, according to F1 score. Despite Flair-german-large extract the largest number of TP compared with open source LLMs, its large number of FP (false location that can be geocoded) lead to a very low F1 score.

When combining methods and calculating the sensitivity ratio, a coverage of 74,98% correct location assignments is reached. Even though fake locations (FP) are a concern, however, most false locations

are either being eliminated by prompt response post processing script and thus not geocoded, or being geocoded outside of Hamburg so not included in the WebGIS system.

4. Discussion

Definition of "Edge Cases":

While implementing NER validation, we found there are some special location terms, e.g. Tiefgarage (underground parking) or elementary school Eimsbüttel; which however refer to **non-specific** or **multiple** locations. Such cases may only be geocoded into **a single location** or **not in Hamburg** at all. Some institution names are hard to detect by open source LLM e.g. terms like "**Molotow"**, the Hamburg-Mitte District Office, the Ministry of Urban Development and Housing, the Ministry of Culture and Media, the Ministry of Schools and Vocational Training, or the Jewish Community in Hamburg.. Other difficult cases are planning sites for future city development, e.g. a replacement building for the Finkenwerder fire brigade and rescue station (here only Finkenwerder was geoparsed).

There are also scenarios that a specific area is mentioned, e.g. "**a conjunction of** road entries near town hall and extension to 100 meters **in the direction of** Adolphs Bridge". Our workflow is currently not able to tackle these challenges, however, it is a good use case to further work with spatial AI in the near future.

In terms of open source LLM, we found Mistral is a more "talkative" model than Llama. Sometimes it does not follow our prompt instruction and produces additional information in response. However, through our post processing script, some institution names were captured and geoparsed, which may lead to higher TP for some SKAs.

As a conventional model, the *flair-german-large* brings more **"stable"** results than commercial or open source LLMs. This is to say, no matter how many times the *flair* model was run, it always reproduces the same results, which is not achieved by LLMs - they may generate different results for each run. For open source models this "instability" is more likely to happen than for commercial LLMs.

Our geoparsing and validation approach can be easily adapted to **other languages** and for **a variety of data types**, like journal articles, wikipedia or social media posts. In such scenario, **language-specific** LLM should be considered (for example: to process Chinese traditional datasets one should consider "MediaTek LLMs [11]). It is also important to search for a region specific geocoder (ideally provided by the local government) to improve the sensitivity ratio.

Following the NER validation, the project will focus on fine-tuning an open source language model in the upcoming months. We will test open source models with a parameter size of 7b or 70b- depending on the availability of GPU resources and with the help of model / data parallelism. The medium-term goal is to develop a chatbot that is able to answer questions on specific urban development topics such as "Densification", "Pre-Emption Rights" and "Main Avenue Planning".

5. Funding Statement

The research was supported by funding from Landesbetrieb Immobilienmanagement und Grundvermögen Hamburg (LIG).

6. Acknowledgements

Open Source and Commercial Large Language Models comparison:

This study was part of one work package within the third phase of HCU´s collaborative research project with LIG Hamburg, which aims to discover modern AI using Geoparsing in an urban governance perspective. We also thank Ms. Cheng-Chia Wu, our DAAD summer intern who developed a prototype of the spatial join validation workflow.

Competing interests

The authors have no competing interests to declare. The *Landesbetrieb Immobilienmanagement und Grundvermögen Hamburg (LIG)* had no influence on the content of this article.

Contact to Autor:

(M.Sc.) Juiwen Chang
Digital City Science (DCS), Hafencity Universität Hamburg (HCU)
Henning-Voscherau-Platz 1, 20457 Hamburg
+49 (0) 40 30088 04568
juiwen.chang@hcu-hamburg.de
https://orcid.org/0000-0002-7896-5609

Literature and Source

[1] Berragan, C. et. al (2023) Transformer based named entity recognition for place name extraction from unstructured text. International Journal of Geographical Information Science, vol. 37, no. 4, 747–766. https://doi.org/10.1080/13658816.2022.2133125

[2] Carterette, B. (2009). Precision and Recall. In: LIU, L., ÖZSU, M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_5050

[3] Dally, B.; Hernandez Leal, H.; Noennig, J. (2024): Explorative Research and Development for Urban Land Development: Digital Tools to Tackle Land Scarcity. In: Translating Knowledge into Innovation Dynamics - Proceedings of the 19th International Forum on Knowledge into Innovation Dynamics (IFKAD 2024): 2081 – 2090, ISBN: 978-88-96687-17-8, ISSN: 2280-787X.

[4] Mello Rose, F., ; Chang, J. (2023). Urban data: harnessing subjective sociocultural data from local newspapers. Buildings and Cities, 4(1), pp. 369–385. DOI: https://doi.org/10.5334/bc.300

[5] Tao, L.; Xie, Z.; Xu, D.; Ma, K.; Qiu, Q.; Pan, S.; Huang, B. (2022) Geographic Named Entity Recognition by Employing Natural Language Processing and an Improved BERT Model. ISPRS Int. J. Geo-Inf., 11, 598. https://doi.org/10.3390/ijgi11120598

[6] Ventrice, L.; Di Caro, L. (2023) Enriching Wikipedia Texts through Geographic Information Extraction -. DOI: 10.1145/3625007.3630110

[7] flair german large model: https://huggingface.co/flair/ner-german-large

[8] AGORA toolbox: https://github.com/digitalcityscience/AGORA

[9] Ollama: https://ollama.com/

[10] Karolina HPC: https://www.it4i.cz/en/infrastructure/karolina

[11] MediaTek LLMs: https://huggingface.co/MediaTek-Research

Open ALKIS? – Oder was passiert, wenn der deutsche Föderalismus auf EU-Recht trifft

Stefan Zaunseder

GISCAD Technologies

GISCAD Technologies, früher GISCAD-Institut, beschäftigt sich seit über 30 Jahren mit Geoinformatik und GIS, in der Vergangenheit insbesondere im kommunalen Bereich. Unser aktueller Entwicklungsschwerpunkt sind Big GeoData und Open Data mit Open Source GIS, z.B. landes- und bald bundesweites ALKIS. Nicht zuletzt dadurch bieten wir umfangreiches Know-How, insbesondere mit sehr großen Geodatenmengen in PostGIS und GeoServer und wie diese performant gehostet werden können, sowie mit benutzerfreundlichen WebGIS-Apps auf OpenLayers-Basis.

Ausgangssituation

Am 9. Juni 2024 trat die Durchführungsverordnung (EU) 2023/138 der Kommission vom 21. Dezember 2022 zur Festlegung bestimmter hochwertiger Datensätze und der Modalitäten ihrer Veröffentlichung und Weiterverwendung [1, 2] in Kraft, die sich auf die Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors [3] stützt. Diese Durchführungsverordnung steht auch in Zusammenhang mit Richtlinie 2007/2/EG des Europäischen Parlaments und des Rates vom 14. März 2007 zur Schaffung einer Geodateninfrastruktur in der Europäischen Gemeinschaft (INSPIRE) [4, 5] und legt für die Datenkategorien Georaum, Erdbeobachtung und Umwelt, Meteorologie, Statistik, Unternehmen und Eigentümerschaft von Unternehmen sowie Mobilität Datensätze fest, die in der Datenkategorie Georaum unter den Bedingungen der Lizenz Creative Commons BY 4.0 oder einer gleichwertigen oder weniger einschränkenden offenen Lizenz in einem öffentlich dokumentierten, unionsweit oder international anerkannten offenen, maschinenlesbaren Format über Anwendungsprogrammierschnittstellen (APIs) und Massen-Download in ihrer jeweils aktuellsten Version zur Weiterverwendung zur Verfügung gestellt werden [1, Anhang 1.2].

Die in der Datenkategorie Georaum von der Durchführungsverordnung erfassten Datensätze umfassen insbesondere auch Gebäude, Flurstücke/Grundstücke (Katasterparzellen), Referenzparzellen und landwirtschaftliche Parzellen [1, Anhang 1.1], die in Deutschland als wesentliche Inhalte des Amtlichen Liegenschaftskatasterinformationssystems (ALKIS) [6] von den Vermessungsverwaltungen der Länder erfasst werden. Die Geltung der o.g. Durchführungsverordnung zum 9. Juni 2024 hat dazu geführt, dass nun (Stand Januar 2025) 15 von 16 Bundesländern entweder die kompletten ALKIS-Daten gemäß GeoInfoDok [7] (aber aus Datenschutzgründen ohne Eigentümerangaben) oder in der Schemavariante „Vereinfachtes Datenaustauschschema“ [8] als Open Data zur Verfügung stellen. Lediglich Bayern schert hier aus.

Aktueller Stand der Verfügbarkeit von ALKIS Open Data

Da in Deutschland die Länder mit ihren Vermessungsverwaltungen für die Führung des Liegenschaftskatasters zuständig sind, kommt es trotz der Zusammenarbeit in der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) [9] zu 16 unterschiedlichen Umsetzungen, die sich insbesondere bzgl. Lizenzen, Datenformaten, Koordinatenreferenzsystemen, Zugang zum Download sowie verfügbaren Diensten und deren Umsetzung teils erheblich unterscheiden.

Verfügbarkeit nach Bundesländern

Hier der aktuelle Stand der Verfügbarkeit von ALKIS Open Data nach Bundesländern in alphabetischer Reihenfolge (Stand Januar 2025), wobei hier der Schwerpunkt auf dem Download liegt, Aussagen zu den Diensten folgen weiter unten:

Baden-Württemberg

Baden-Württemberg gehört zu den Bundesländern, die mit Inkrafttreten der Durchführungsverordnung ihr Open Geodata-Angebot um ALKIS [10] erweitert haben.

- Lizenz (Download und Dienste): Datenlizenz Deutschland – Namensnennung – Version 2.0 (DL-DE→BY-2.0) [11]
- Download-Formate: NAS gemäß GeoInfoDok 7.1.2, Shape
- Koordinatenreferenzsystem Download: ETRS89 / UTM Zone 32N (EPSG:25832) [12]
- Download: Maximal 5 Gemarkungen nach interaktiver Auswahl in Webkarte [13] ohne Katalogdaten (Daten zu Fluren bis Land), Katalogdaten landesweit getrennt in undokumentiertem CSV-Format
- Aktualisierungszyklus: viermal pro Jahr
- Verfügbare Dienste: WMS, WMTS und WFS

Bayern

Der den gemäß Durchführungsverordnung in der Datenkategorie Georaum umfassten Daten am nächsten kommende, als Open Data verfügbaren Datensatz ist in Bayern die sogenannte ALKIS Parzellarkarte [14], die jedoch ausschließlich als Rasterdaten ohne Flurstücksnummer zur Verfügung steht.

Die ALKIS Parzellarkarte erfüllt nicht die Vorgaben der Umsetzungsverordnung, insbesondere genügt das Format nicht den Anforderungen (schon weil keine Attributierung möglich ist), die Flurstücksnummer als Code der Parzelle oder Buchungseinheit ist nicht enthalten, es ist kein Massendownload möglich und auch die Lizenz entspricht nicht den Vorgaben der Umsetzungsvorordnung.

Bayern schert damit als einziges von sechzehn Bundesländern aus und setzt die Umsetzungsverordnung diesbzgl. nicht um, gemäß telefonischer Anfrage beim Kundenservice des Landesamtes für Digitalisierung, Breitband und Vermessung (LDBV) [14] ist dies auch nicht geplant, für weiteres wurde auf den Schriftweg verwiesen. Eine entsprechende schriftliche Anfrage seitens GISCAD Technologies folgt, über das Ergebnis kann hoffentlich im Rahmen des Vortrages auf der FOSSGIS 2025 berichtet werden.

- Lizenz (Download und Dienste): Creative Commons Namensnennung-Keine Bearbeitungen 4.0 International (CC BY-ND 4.0) [15]
- Download-Formate: GeoTIFF
- Koordinatenreferenzsystem Download: EPSG:25832
- Download: max. 10 km² nach interaktiver Polygondigitalisierung
- Aktualisierungszyklus: täglich
- Verfügbare Dienste: WMS und WMTS

Berlin

Berlin hat bereits vor Inkrafttreten der Durchführungsverordnung ALKIS als Open Data zur Verfügung gestellt [16].

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: DXF, NAS gemäß GeoInfoDok 7.1.2, Shape
- Koordinatenreferenzsystem Download: ETRS89 / UTM Zone 33N (EPSG:25833) [17]
- Download: landesweit
- Aktualisierungszyklus: unbekannt
- Verfügbare Dienste: WMS und WFS

Brandenburg

Auch Brandenburg hat bereits vor Inkrafttreten der Durchführungsverordnung ALKIS als Open Data zur Verfügung gestellt [18].

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: NAS gemäß GeoInfoDok 7.1.2, Shape
- Koordinatenreferenzsystem Download: EPSG:25833
- Download: landkreisweit
- Aktualisierungszyklus: vierteljährlich
- Verfügbare Dienste: WMS und WFS

Bremen

Die Hansestadt Bremen stellte ALKIS bereits vor Inkrafttreten der Durchführungsverordnung als Open Data zur Verfügung [19]. Es muss abgewartet werden, ob sich bzgl. der folgenden Angaben mit Einführung der GeoInfoDok 7.1.2 Änderungen ergeben.

- Lizenz (Download und Dienste): Creative Commons Namensnennung 4.0 International (CC BY 4.0) [20]
- Download-Formate: DXF, NAS gemäß GeoInfoDok 6.0.1, Shape
- Koordinatenreferenzsystem: EPSG:25832
- Download: getrennt für die Städte Bremen und Bremerhaven
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Hamburg

Auch die Hansestadt Hamburg stellte ALKIS bereits vor Inkrafttreten der Durchführungsverordnung als Open Data zur Verfügung [21], die Umstellung auf GeoInfoDok 7.1.2 erfolgte im Januar 2025.

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: JPEG, NAS gemäß GeoInfoDok 7.1.2
- Koordinatenreferenzsystem: EPSG:25832
- Download: landesweit

- Aktualisierungszyklus: vierteljährlich
- Verfügbare Dienste: WMS, WFS

Hessen

Das Land Hessen stellt ALKIS seit Februar 2022 als Open Data zur Verfügung [22].

- Lizenz (Download und Dienste): jede Nutzung ist ohne Einschränkung oder Bedingung erlaubt [22]
- Download-Formate: NAS gemäß GeoInfoDok 7.1.2
- Koordinatenreferenzsystem: EPSG:25832
- Download: landkreisweit
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Mecklenburg-Vorpommern

Das Land Mecklenburg-Vorpommern führte ALKIS als Open Data mit Inkrafttreten der Durchführungsverordnung ein [23].

- Lizenz (Download und Dienste): CC BY 4.0
- Download-Formate: NAS gemäß GeoInfoDok 7.1.2, Shape
- Koordinatenreferenzsystem: EPSG:25833
- Download: gemeindeweise nach interaktiver Auswahl in Webkarte
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Niedersachsen

Auch das Land Niedersachsen führte ALKIS Open Data mit Inkrafttreten der Durchführungsverordnung ein [24]. Bzgl. Download besteht hier die Besonderheit, dass als Massendownload ausschließlich GeoPackage in einem eigenen Schema (landkreisweise verlinkt über ein GeoJSON-Dokument) zur Verfügung steht, Daten im NAS-Format oder im vereinfachten Schema sind über einen WFS mit. max. 10.000 Features je Request verfügbar.

- Lizenz (Download und Dienste): CC BY 4.0
- Download-Formate: GeoPackage
- Koordinatenreferenzsystem: EPSG:25832
- Download: landkreisweise über Links in GeoJSON-Dokument
- Aktualisierungszyklus: wöchentlich
- Verfügbare Dienste: WMS, WFS

Nordrhein-Westfalen

Das Land Nordrhein-Westfalen stellte ALKIS bereits vor Inkrafttreten der Durchführungsverordnung als Open Data zur Verfügung [25]. Die Umstellung auf GeoInfoDok 7.1.2 ist noch nicht abgeschlossen, erst ein Teil der Landkreise ist bereits im neuen Format verfügbar.

- Lizenz (Download und Dienste): Datenlizenz Deutschland – Zero – Version 2.0 (DL-DE→Zero-2.0) [26]
- Download-Formate: NAS gemäß GeoInfoDok 6.0.1
- Koordinatenreferenzsystem: EPSG:25832
- Download: landkreisweise
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WMTS, WFS

Rheinland-Pfalz

Das Land Rheinland-Pfalz gehört zu den Bundesländern, die ALKIS Open Data mit Inkrafttreten der Durchführungsverordnung bereitgestellt haben [27].

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: NAS gemäß GeoInfoDok 7.1.2
- Koordinatenreferenzsystem: EPSG:25832
- Download: gemarkungsweise
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Saarland

Das Saarland stellt ALKIS Open Data seit Inkrafttreten der Durchführungsverordnung zur Verfügung [28], allerdings nicht gemäß GeoInfoDok, sondern im vereinfachten Schema. Auch steht keine direkte Download-Möglichkeit zur Verfügung, sondern ein WFS, der max. 10.000 Features pro Request liefert.

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: GML (vereinfachtes Schema)
- Koordinatenreferenzsystem: EPSG:25832
- Download: kein direkter Download
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Sachsen

Der Freistaat Sachsen stellte ALKIS Open Data bereits vor Inkrafttreten der Durchführungsverordnung zur Verfügung [29]. Allerdings ist der Abschluss der Umstellung auf GeoInfoDok 7.1.2 erst bis Ende 2026 geplant [30].

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: NAS gemäß GeoInfoDok 6.0.1
- Koordinatenreferenzsystem: EPSG:25833
- Download: ohne Punktinformationen landesweit, mit Punktinformationen gemeindeweise
- Aktualisierungszyklus: unklar

- Verfügbare Dienste: WMS, WFS

Sachsen-Anhalt

Das Land Sachsen-Anhalt stellt ALKIS Open Data in einem reduzierten vereinfachten Schema in genau dem Minimalumfang zur Verfügung, der gemäß Durchführungsverordnung erforderlich ist (nur Flurstücke, Gebäude und tatsächliche Nutzung) [31].

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: GML (reduziertes vereinfachtes Schema)
- Koordinatenreferenzsystem: EPSG:25832
- Download: landesweit oder landkreisweise, jeweils getrennt nach Flurstücken, Gebäuden und tatsächlicher Nutzung
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Schleswig-Holstein

Schleswig-Holstein hat ALKIS Open Data bereits vor Inkrafttreten der Durchführungsverordnung zur Verfügung gestellt [32]. Hier besteht allerdings die Besonderheit, dass die vollständigen Daten im NAS-Format nur einzeln flurweise nach interaktiver Auswahl auf einer Webkarte herunterladbar sind (Schleswig-Holstein hat ca. 18.000 Fluren!), während die Daten im vereinfachten Schema gemeindeweise heruntergeladen werden können.

- Lizenz (Download und Dienste): CC BY 4.0
- Download-Formate: GML (vereinfachtes Schema), NAS gemäß GeoInfoDok 7.1.2
- Koordinatenreferenzsystem: EPSG:25832
- Download: vereinfachtes Schema gemeindeweise, im NAS-Format flurweise nach interaktiver Auswahl in Webkarte, Katalogdaten getrennt
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Thüringen

Thüringen hat ALKIS als Open Data bereits vor Inkrafttreten der Durchführungsverordnung zur Verfügung gestellt [33].

- Lizenz (Download und Dienste): DL-DE→BY-2.0
- Download-Formate: NAS gemäß GeoInfoDok 7.1.2
- Koordinatenreferenzsystem: EPSG:25832
- Download: flurweise oder katasterbereichsweise, max. 20 (flurweise) bzw. 8 Dateien (katasterbereichsweise) nach interaktiver Auswahl in Webkarte
- Aktualisierungszyklus: unklar
- Verfügbare Dienste: WMS, WFS

Fazit

Aufgrund der föderalen Struktur und der Zuständigkeit der Länder ist es bei der Umsetzung der Durchführungsverordnung in Deutschland bzgl. ALKIS zu einem Sammelsurium von Umsetzungen gekommen. Die Durchführung jedes Bundeslandes unterscheidet sich.

Insbesondere sind folgende Punkte zu nennen:

- Bayern verweigert die Umsetzung der Durchführungsverordnung bzgl. ALKIS.
- Fünf verschiedene Lizenzen:
  - CC BY 4.0
  - CC BY-ND 4.0 (nur Bayern, entspricht nicht der Durchführungsverordnung)
  - DL-DE→BY-2.0
  - DL-DE→Zero-2.0
  - „jede Nutzung ist ohne Einschränkung oder Bedingung erlaubt“ (nur Hessen)
- Acht verschiedene Download-Formate, wobei manche Bundesländer nur ein Format anbieten, andere bis zu drei (NAS unterscheidet sich zwischen GeoInfoDok 6.0.1 und 7.1.2 erheblich, daher werden diese als getrennte Formate aufgeführt):
  - DXF
  - GeoPackage (nur Niedersachsen)
  - GeoTIFF (nur Bayern, entspricht nicht der Durchführungsverordnung)
  - GML (vereinfachtes Schema, Sachsen-Anhalt verwendet ein reduziertes vereinfachtes Schema, bei den Daten aus dem Saarland kommt es zu Abweichungen vom Schema)
  - JPEG (nur Hamburg)
  - NAS gemäß GeoInfoDok 6.0.1
  - NAS gemäß GeoInfoDok 7.1.2 (bei den Daten aus Thüringen treten Abweichungen vom Schema bzw. zusätzliche Feature-Typen auf)
  - Shape
- Zwei Koordinatenreferenzsysteme:
  - EPSG:25832
  - EPSG:25833
- Der Download ist je Bundesland unterschiedlich erreichbar bzw. durchzuführen:
  - Einige Bundesländer bieten nur interaktive Auswahl auf verschiedenen Ebenen (Fluren, Gemarkungen, Gemeinden, Katasterbereiche) auf einer Webkarte an.
  - Andere Bundesländer bieten nur einen WFS mit verschiedenen Schemata (NAS oder vereinfachtes Schema, immer mit begrenzter Featurezahl pro Request, meist 10.000) als „Download-Dienst“ an.
  - Niedersachsen steckt Download-Links auf GeoPackage in GeoJSON.
  - Bei manchen Bundesländern sind die Katalogdaten getrennt herunterzuladen, im Fall Baden-Württemberg in einem undokumentierten CSV-Format.
- Ihren Aktualisierungszyklus geben die meisten Bundesländer gar nicht erst an, bei denen, die ihn angeben, unterscheidet er sich erheblich (täglich bis vierteljährlich).

- Die verfügbaren Dienste sind oft nicht performant und unterscheiden sich zwischen den Bundesländern.
  - WMS:
    - Es werden unterschiedliche Layergliederungen angeboten.
    - Oft ist der Root-Layer unbenannt und kann daher nicht abgerufen werden.
    - Die Darstellungen unterscheiden sich erheblich (teilweise relativ nahe an der GeoInfoDok, aber diesbzgl. nicht vollständig, teilweise kaum GeoInfoDok-konform).
    - Eine aus Performanzgründen zu bevorzugende Nutzung in Tiles (auch mit dem WMS-Parameter TILED=true) ist aufgrund abgeschnittener und/oder wiederholter Signaturen oder Beschriftungen oft nicht möglich.
    - Die Antwort auf GeoFeatureInfo-Requests ist uneinheitlich und meist unvollständig.
  - WMTS: Obwohl WMTS aus Performanzgründen, insbesondere aufgrund der gut cachebaren Tiles, WMS deutlich vorzuziehen wäre, bieten nur drei Bundesländer einen WMTS an.
  - WFS:
    - Es werden unterschiedliche Schemata angeboten, z.B. NAS-konform, vereinfachtes Schema oder nicht näher definiertes „simple feature“.
    - Alle untersuchten WFS haben ein angesichts der ALKIS-Datenmengen recht niedriges Limit auf der Anzahl der mit einem Request anforderbaren Features (meist 10.000).

ALKIS stellt aufgrund der hohen Genauigkeit im cm-Bereich und des umfangreichen Inhalts (auch wenn auch dieser zwischen den Bundesländern etwas schwankt) in Deutschland die wichtigste Open Data-Datengrundlage im Bereich amtlicher Open Data dar. Insgesamt ist festzustellen, dass diese, dem Föderalismus und der mangelnden Koordination zwischen den Bundesländern zu schuldenden, umfangreichen Unterschiede in Sachen ALKIS Open Data die Nutzung dieser wertvollen Open Data-Resource massiv erschweren und damit auch verteuern.

Dass es auch anders geht, zeigt z.B. unser Nachbar Frankreich, wo Katasterdaten für ganz Frankreich aus einer Quelle in den Formaten DXF, EDIGEO [34], GeoJSON, MBTiles, Shape und TIFF heruntergeladen werden können [35].

GISCAD Maps OpenData++

Um Open Data, insbesondere amtliche offene Geodaten in Deutschland besser zugänglich zu machen, entwickelt GISCAD Technologies daher einen auf Open Source- Softwarekomponenten basierenden Dienst: GISCAD Maps OpenData++. In einem ersten Schritt werden dazu die ALKIS-Daten aller Bundesländer (nach aktuellem Stand außer Bayern, siehe oben) einheitlich bereitgestellt.

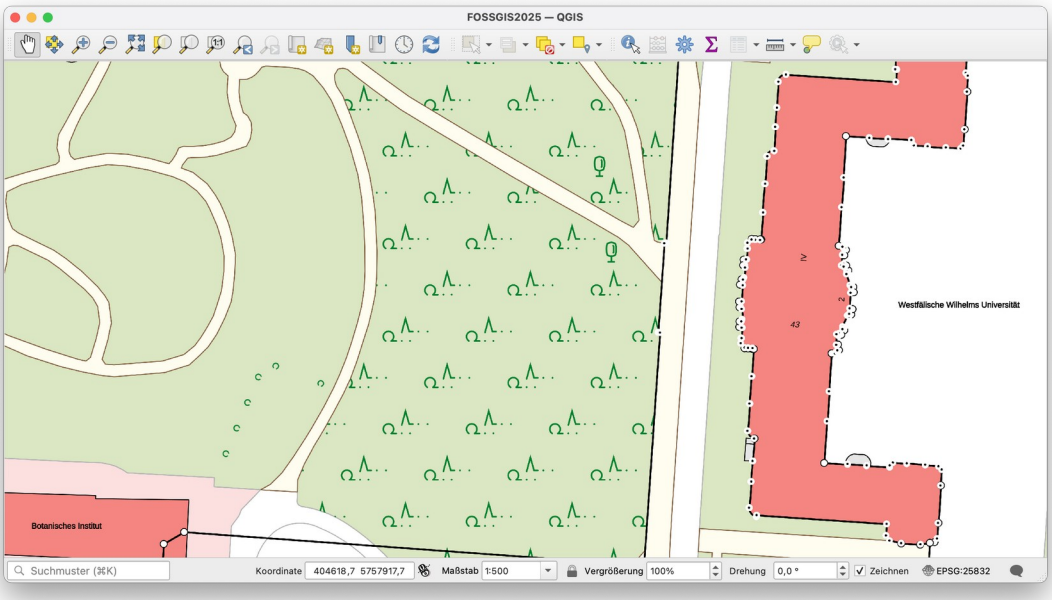
GISCAD Maps OpenData++ ALKIS bietet wesentliche Vorteile, z.B.:

- Der performante Zugriff ist aus allen Anwendungen und Bibliotheken möglich, die WMS, WMTS oder WFS unterstützen, z.B. QGIS oder OpenLayers.
- Parallel steht eine eigene, komfortable Web-App zur Verfügung (für Desktop und mobil).
- Einheitlicher, performanter Zugriff unabhängig vom Bundesland.
- Einheitliche und vollständige, farbige und graue Darstellung exakt gemäß GeoInfoDok mit Varianten (aktuell mit/ohne tatsächlicher Nutzung, mit/ohne Bodenschätzung, weitere folgen),

z.Zt. optimiert für den Maßstab 1 : 1.000, zusätzliche Maßstäbe und eine bildschirmoptimierte Darstellung sind geplant.

- Unterstützung von HiDPI-Bildschirmen bis zum Auflösungsfaktor 4x (1x, 1.5x, 2x, 3x und 4x), dadurch qualitativ optimale Darstellung auch auf sog. Retina-Bildschirmen. Gleichzeitig sind diese hochauflösenden Tiles optimal geeignet auch für eine Druckausgabe.
- GeoFeatureInfo-Requests liefern die komplette ALKIS-Attributierung.
- WMTS für alle Bundesländer
- Service Level Agreement
- Datenschutzfreundliche Umsetzung

Start des Pilotbetriebs zur FOSSGIS 2025



GISCAD Maps OpenData++ ALKIS:
Der Veranstaltungsort der FOSSGIS 2025 in QGIS (WMTS)

GISCAD Technologies wird den Pilotbetrieb von GISCAD Maps OpenData++ zur FOSSGIS 2025 starten. Nähere Informationen dazu sind ab der FOSSGIS auf unserer Webseite https://giscad.de zu finden.

Verwendete Open Source-Komponenten

Serverseitig werden im Rahmen von GISCAD Maps PostgreSQL mit PostGIS sowie GeoServer als Open Source GIS-Komponenten eingesetzt. Die clientseitige Web-App verwendet OpenLayers.

Für die Zukunft würden wir seitens GISCAD Technologies gerne zu verwendeten Open Source-Projekte beitragen. So haben wir für die korrekte und vollständige GeoInfoDok-konforme Darstellung in GeoServer z.B. GeoTools-Patches entwickelt, die auch für andere GeoTools- bzw. GeoServer-Anwender interessant sein dürften. Ähnliches gilt für ein von uns entwickeltes Kommandozeilen-Tool, mit dem sich GeoServer komplett über seine REST-API verwalten lässt. Da eine solche Beteiligung an Open Source-Projekten für uns neu ist, suchen wir Kontakt zu Entwicklern dieser Projekte, insbesondere GeoServer und GeoTools, um unsere Beiträge geeignet vorzubereiten und einzubringen.

Kontakt zum Autor:

Stefan Zaunseder
GISCAD Technologies GmbH & Co. KG
Moosburger Str. 8 – 84169 Altfraunhofen
+49 170 4822783
sz@giscad.de

Literatur

[1] EUR-Lex: https://eur-lex.europa.eu/eli/reg_impl/2023/138/oj?locale=de

[2] Wikipedia: https://de.wikipedia.org/wiki/Durchführungsverordnung_(EU)_2023/138

[3] EUR-Lex: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=CELEX:32019L1024

[4] EUR-Lex: https://eur-lex.europa.eu/legal-content/DE/TXT/?uri=celex%3A32007L0002

[5] Wikipedia: https://de.wikipedia.org/wiki/Infrastructure_for_Spatial_Information_in_the_European_Community

[6] Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV): https://www.adv-online.de/AdV-Produkte/Liegenschaftskataster/

[7] AdV: https://www.adv-online.de/GeoInfoDok/

[8] AdV-Produktspezifikation ALKIS-WFS und Ausgabeformate (Shape, CSV): https://www.adv-online.de/icc/extdeu/med/f49/f49502a0-36fa-6b61-c2d2-1bf43b36c4c2,11111111-1111-1111-1111-111111111111

[9] Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV): https://www.adv-online.de/

[10] Landesamt für Geoinformation und Landentwicklung: https://www.lgl-bw.de/Produkte/Liegenschaftskataster/ALKIS/index.html

[11] Datenlizenz Deutschland – Namensnennung – Version 2.0 (DL-DE→BY-2.0): https://www.govdata.de/dl-de/by-2-0

[12] epsg.io: https://epsg.io/25832

[13] Open GeoData Portal des Landesamt für Geoinformation und Landentwicklung (LDBV): https://opengeodata.lgl-bw.de/

[14] LDBV ALKIS-Parzellarkarte: https://geodaten.bayern.de/opengeodata/OpenDataDetail.html?pn=parzellarkarte

[15] Creative Commons Namensnennung-Keine Bearbeitungen 4.0 International (CC BY-ND 4.0): https://creativecommons.org/licenses/by-nd/4.0/deed.de

[16] ALKIS Berlin: https://fbinter.stadt-berlin.de/fb/loginkey=showMap&mapId=wmsk_alkis@senstadt

[17] epsg.io: https://epsg.io/25833

[18] Landesvermessung und Geobasisinformation Brandenburg (LGB): https://geobroker.geobasis-bb.de/gbss.php?MODE=GetProductInformation&PRODUCTID=6de36219-3e68-489e-8ebc-632e5ff-b6dc9

[19] Metadatenverbund: https://www.metaver.de/trefferanzeige?docuuid=CE29A1C3-C82C-4600-A8B9-04CBDE26F6D8

[20] Creative Commons Namensnennung 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/deed.de

[21] Metadatenverbund: https://metaver.de/trefferanzeige?docuuid=DC71F8A1-7A8C-488C-AC99-23776FA7775E

[22] Einführung von Open Data zum 01.02.2022: https://hvbg.hessen.de/geoinformation/open-data

[23] Landesamt für innere Verwaltung, Amt für Geoinformation, Vermessungs- und Katasterwesen, Open-Data-Angebot: https://www.laiv-mv.de/Geoinformation/Open_Data_Angebot/

[24] Landesamt für Geoinformation und Landesvermessung Niedersachsen, OpenGeoData.NI, ALKIS: https://ni-lgln-opengeodata.hub.arcgis.com/apps/lgln-opengeodata::alkis/about

[25] Information und Technik Nordrhein-Westfalen, OpenGeodata.NRW, Grundrissdaten: https://www.opengeodata.nrw.de/produkte/geobasis/lk/akt/gru_xml/

[26] Datenlizenz Deutschland – Zero – Version 2.0 (DL-DE→Zero-2.0): https://www.govdata.de/dl-de/zero-2-0

[27] Landesamt für Vermessung und Geobasisinformation, Open Data: https://lvermgeo.rlp.de/geodaten-geoshop/open-data

[28] GeoPortal Saarland, Geobasisdaten (opendata): https://geoportal.saarland.de/app-article/geobasisdatenuebersicht/

[29] geodaten.sachsen.de: https://www.geodaten.sachsen.de/liegenschaftskataster-3990.html

[30] landesvermessung.sachsen.de: https://www.landesvermessung.sachsen.de/geoinfodok-neu-referenz-7-1-referenzversion-ab-31-12-2023-8706.html

[31] lvermgeo.sachsen-anhalt.de: https://www.lvermgeo.sachsen-anhalt.de/de/gdp-open-data.html

[32] Landesamt für Vermessung und Geoinformation Schleswig-Holstein: https://www.schleswig-holstein.de/DE/landesregierung/ministerien-behoerden/LVERMGEOSH/Themen/themaOpenGbd

[33] Geoportal Thüringen: https://geoportal.thueringen.de/gdi-th/download-offene-geodaten

[34] Wikipedia: https://fr.wikipedia.org/wiki/EDIGEO

[35] cadastre.data.gouv.fr: https://cadastre.data.gouv.fr/datasets

Governance von Open-Source-Software im öffentlichen Sektor: Make, Buy or Contribute?

Christian Weidner[1]

[1] Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS

DOI: 10.5281/zenodo.14616445

Zusammenfassung: Die Beschaffenheit von OSS-Projekten im öffentlichen Sektor variiert stark: Während einige Projekte den häufig proklamierten offenen und kollaborativen Bazaar leben, bleiben andere hinter einem Schaufenster verschlossen. Dieser Beitrag untersucht, warum sich Behörden für oder gegen die Öffnung ihrer Softwareprojekte entscheiden und gibt Einblick in ausgewählte OSS-Projekte der öffentlichen Verwaltung in Deutschland. Die Unsicherheit über ausbleibende Projektbeteiligung wird als entscheidender Faktor identifiziert, weshalb Open Source häufig nur als Transparenzinitiative und nicht als Methode der kollaborativen Innovationsstrategie eingesetzt wird.

Schlüsselwörter: Open-Source-Software, Transaktionskostenansatz, Governance, öffentliche Verwaltung, Bazaar, Softwareentwicklung, Kollaboration

1. Problemaufriss und Forschungsgegenstand

Die öffentliche Hand steht vor der Herausforderung, ihre digitale Infrastruktur nachhaltig und effizient zu gestalten. Im Rahmen des Onlinezugangsgesetzes (OZG) wurde hierfür das kontrovers diskutierte Prinzip *Einer für Alle* etabliert, welches Redundanzen und Ungleichgewichte bei der Entwicklung von OZG-Lösungen durch die Zuweisung von *Themenfeldern* an einzelne Bundesländer verhindern soll [1]. Während die Entwicklungs- und Betriebsverantwortung jeweils eines federführenden Bundeslandes zunächst ein inhärentes System kostengünstiger Nachnutzung skizziert, stößt laut Kritiker*innen aus der Verwaltungspraxis diese Vorgehensweise schnell an seine Grenzen [2–4]. Stattdessen wird die konsequente Etablierung von Mechanismen der Entwicklungslogik von Open-Source-Software (OSS) gefordert. Die kollaborative Weiterentwicklung von Software soll das Innovationspotenzial von Software stärken und die Entwicklung langfristig sicherstellen [5]. In jüngeren Untersuchungen wurde auch ein volkswirtschaftlich positiver Effekt durch OSS nachgewiesen, der den lokalen IT-Sektor maßgeblich stärkt [6].

Die öffentliche Verwaltung steht OSS oftmals janusköpfig entgegen: Während es Anfang der 2000er Jahre einen regelrechten Hype um die Migration auf OSS-Produkte wie Linux oder OpenOffice.org gab, schliefen viele dieser Projekte wieder ein [7]. Zuletzt wandelte sich die Auseinandersetzung im öffentlichen Sektor Jahren immer weiter in Richtung einer umfangreichen strategischen Ausrichtung *Open Source* als Methode zu begreifen, sich in Communitys zu beteiligen oder selbst solche aufzubauen. Mit zahlreichen aktuellen strategischen Initiativen wie *Open Code* [8], *ZenDiS* oder *Sovereign Tech Agency* [9] sowie einer Vielzahl von OSS-Projekten treiben Akteur*innen im öffentlichen Sektor die Verbreitung von OSS voran. Open-Source-Governance [10, 11] sowie der charakteristische *Bazaar* [12, 13] als Steuerungsmodus scheint öffentlichen Einrichtungen jedoch schwer zu fallen, obwohl bereits zuvor Entscheidungen für die offene Lizenzierung und Veröffentlichung positiv ausgefallen sind [14]. Vor diesem Phänomen beantwortet dieser Beitrag die Frage, warum sich öffentliche Einrichtungen für den Bazaar als Steuerungsform in OSS-Projekten (nicht) entscheiden.

2. Theoretischer Hintergrund

Transaktionskostenansatz – Eine kollaborative Governance ermöglicht es Projektverantwortlichen, ein Produkt selbst herzustellen und gleichzeitig bei geringer Anreiz- und Kontrollintensität durch externe Beteiligte innerhalb der Projektgrenzen Innovationen zu akquirieren. Die Koordination solcher Projekte kostet allerdings Zeit und Ressourcen. Während in der Organisationsforschung der Transaktionskostenansatz (TKA) bislang weitreichend erklärt, wann sich Organisationen für *Make* (Eigenfertigung) oder *Buy* (Auslagerung) entscheiden [15, 16], wurde das Erklärungsangebot auf offene Entwicklungsgemeinschaften bisher eher theoretisch und empirisch nur unzureichend angewandt [13]. Ob es sich lohnt, das eigene Projekt zu öffnen, hängt im Sinne des TKA von Transaktionskosten ab, wie etwa Informations- und Suchkosten, Personalkosten oder Kosten für das Ausräumen rechtlicher Unsicherheiten. In der Softwareentwicklung können solche Kosten z. B. aufgrund hochgradig (organisations-)spezifischer, einzigartiger Software ausgelöst werden [17]. Die Bazaar-Governance reiht sich zwischen die hierarchische Eigenfertigung und das marktorientierte Outsourcing ein [18], ohne die Verbindlichkeiten von Netzwerken zu erfahren [19] (Abb. 1). Demil und Lecocq merken jedoch an, dass die Bazaar-Governance in der Praxis eher Mischformen mit den angrenzenden Governance-Modi einnehmen wird [13, 19].

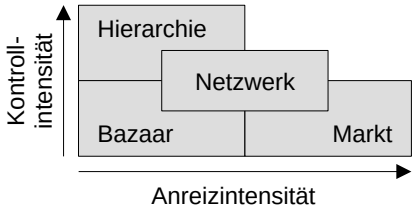


Abb. 1: Governance-Modi.

Innovationsabhängigkeit – Im Vergleich zum Markt und zur hierarchischen Eigenfertigung weist der Bazaar durch seine freien Lizenzen eine Offenheit gegenüber potenziellen Beitragenden auf. Offenheit wird in der Organisationsforschung in verschiedenen Dimensionen begriffen, darunter fallen Transparenz, Partizipation und Entscheidungsbefugnisse, die sich in Konzepten wie Open Access, Open Data, Open Innovation oder Open Strategy konstituieren [20]. Durch ihre standardisierten Projektstrukturen besitzen Open-Source-Projekte die Besonderheit sowohl Transparenz herzustellen, Beteiligung zuzulassen, als auch einen Weg in die strategische Governance ebnen zu können [21]. Insbesondere durch die Öffnung der Beteiligungs- und Entscheidungsstrukturen können so Trends und Innovationen akquiriert werden, die bei einer rein internen Unternehmenssicht ungesehen bleiben könnten [22]. Die Ressourcenabhängigkeitstheorie berücksichtigt die Abhängigkeit einer Organisation von den Ressourcen ihrer Umwelt [23] und kann die Projektbindung externer *Lead User* an das eigene Projekt erklären [24].

Isomorphismus und Pfadabhängigkeit – Die ökonomische Sicht der (begrenzt) rationalen Entscheidung [25] kann die Fehlannahme reproduzieren, dass Akteure überhaupt einen Entscheidungsspielraum haben. Der institutionelle Isomorphismus skizziert einen neoinstitutionalistischen Vorschlag, weshalb Organisationen sich an ihre Umwelt anpassen (müssen), der sich laut DiMaggio und Powell in drei Spielarten unterteilt [26]: *koerzive* Prozesse, also solche die z. B. aufgrund politischer, rechtlicher oder finanzieller Vorgaben entstehen, *normative* Prozesse, die z. B. auf ähnliche Lebensläufe der Organisationsmitglieder zurückzuführen sind sowie *mimetische* Prozesse, die bei Unsicherheit über das richtige Vorgehen entstehen, häufig durch das unhinterfragte adaptieren sogenannter *Best Practices*. Ähnlichkei-

ten zum Isomorphismus weist auch die Theorie der Pfadabhängigkeit auf [27], die insbesondere im Zusammenhang mit den Linux-Strategien deutscher Kommunalverwaltungen Anfang der 2000er Jahre diskutiert wurde [28].

3. Feldzugang und Datenerhebung

Tabelle 1: Fallstudienauswahl und Ergebnisübersicht

	Masterportal	**KoliBri**	**ZMS**
Website	www.masterportal.org	public-ui.github.io	nicht vorhanden
Zweck	Geoportalanwendung	Programmbibliothek	Terminverwaltungssystem
Plattformen	BitBucket, Trello	GitHub	GitLab
Lizenz	MIT	EUPL 1.2	EUPL 1.2 + MIT
*Initiator*in*	Landesbetrieb	Bundesoberbehörde	landeseigene AöR
Teilnehmende Behörden	> 49	< 10	2
Feldzugang	Dokumentenanalyse, zwei Interviews, Onlinebefragung (N=38).	Dokumentenanalyse, zwei Interviews, eine teiln. Beobachtung.	Dokumentenanalyse, zwei Interviews.
Strategische Governance	Netzwerk	Hierarchie	Hierarchie
Technische Entwicklungszusammenarbeit	Formalisierter Bazaar	Bazaar	Informelles Netzwerk
Investitionsspezifität (nicht getätigte Invest.)	eher hoch (gering)	eher gering (moderat)	eher hoch (hoch)
Unsicherheit	niedrig	niedrig	hoch
Ressourcenabhängigkeit	hoch	eher hoch	niedrig
Isomorphie	positiver Effekt	ambivalent	teils/ teils

Für das vergleichende Fallstudiendesign wurden drei verwaltungseigene OSS-Projekte ausgewählt [29], die unterschiedliche Steuerungsstrukturen aufweisen (siehe Tab. 1 oben). Zunächst wurden die Governance-Modi mittels qualitativer Inhaltsanalyse von Projektdokumenten, Websites und Repositorien analysiert und anschließend in fünf semi-strukturierten Interviews mit Projektbeteiligten vertiefend erhoben. Die Interviews fanden zwischen dem 20.06. und 26.07.2024 online statt und dauerten zwischen 45 und 75 Minuten. Sie dienten der qualitativen Erhebung der Einflussfaktoren auf die Wahl der Governance und ermöglichten einen tiefer gehenden Einblick in Steuerungsstrukturen. Die Interview-Leitfäden stützen sich auf eine vorherige Operationalisierung der zu prüfenden unabhängigen Variablen. Die Gespräche wurden gemäß gängiger inhaltlich-semantischer Transkriptionsregeln verschriftlicht [30], um anschließend theoriegeleitet kodiert werden zu können [31] (siehe Anhang, Tab. 3).

Im Rahmen der Untersuchung des Masterportals wurde zudem eine Onlinebefragung zwischen dem 24.07. und 02.09.2024 unter den Projektbeteiligten durchgeführt (siehe Anhang, Tab. 4). Die Befragungsteilnehmenden wurden quantitativ mithilfe von Likert-Skalen (7 Punkte) u. a. zur Wahrnehmung der Projektspezifität, zu getätigten Investitionen, organisationalen Hürden und der Motivation im Bezug auf das Masterportal-Projekt befragt [vgl. Surveydesigns 32–34]. Die quantitativen Daten ermöglichten auch die Prüfung moderierender Variablen wie der Lokalität oder der Verwaltungsebene der einzelnen Beteiligten. Hierfür wurden (a) die Verwaltungsebene in Anlehnung an die *Nomenclature des unités territoriales statistiques*, (b) die Digitalisierung des Standortes, unter Zuhilfenahme des *Smart City Index* (SCI) 2024 [35, 36]) und (c) die Nähe der Beteiligten zur initiierenden sowie Lead-User-Behörde, um mögliche lokale Clustereffekte prüfen zu können [37], berücksichtigt. Letzte Variable bezieht sich auf die zeitliche Entfernung zwischen den (Haupt-)Bahnhöfen der jeweiligen Standorte mit Verweis auf

die Fahrzeiten der Deutschen Bahn [38]. Die Ergebnisse (siehe Tab. 1 unten) werden nachfolgend anhand der untersuchten Projekte vorgestellt und anschließend zusammenfassend diskutiert

4. Ergebnisdarstellung

Verbindliches Netzwerk und formalisierter Bazaar – Das *Masterportal* ist eine Webanwendung für den zentralen, visualisierten Zugang zu Geodaten, die seit 2014 vom Landesbetrieb Geoinformation und Vermessung (LGV) der Freien und Hansestadt Hamburg entwickelt wird. Die Software befindet sich dort seit 2015 im produktiven Einsatz und wurde frühzeitig unter einer OSS-Lizenz veröffentlicht. Es verfügt über ausgeprägte, transparente Steuerungsstrukturen, die seit 2018 über die *Implementierungspartnerschaft* (IPM) verbindlich institutionalisiert sind. Die IPM wurde von fünf Behörden initiiert und soll die langfristige gemeinschaftliche Finanzierung und Weiterentwicklung sicherstellen. Sie gilt im öffentliche Sektor als Erfolgsmodell [39], das durch weitere Projekte – wie *DIPAS* und *Civitas Connect* – adaptiert wurde. Die Zahl der Mitglieder ist seit dem auf 49 öffentliche Einrichtungen gestiegen [40]. Zentrales Organ der Projektsteuerung ist das *Strategische Komitee*, welches viermal im Jahr online tagt und an dem jedes IPM-Mitglied stimmberechtigt teilnehmen kann. Die Beteiligung an den Sitzungen ist in der Regel hoch. Entscheidungen über technische Weiterentwicklungen werden unter Beschlussvorlage des *Technischen Komitees*, welches das zweite relevante Organ der IPM bildet, per einfacher Mehrheit demokratisch entschieden. Der LGV behält sich als Projektgründer ein Vetorecht für diese Entscheidungen vor.

Obwohl die Projektstruktur durch die IPM hochgradig formalisiert ist, haben sich auch informelle Netzwerke herausgebildet, die wiederum selbstverstärkende Effekte aufweisen. Regressionsanalysen zum Antwortverhalten zeigen, dass eine Nähe zum Lead User einen leichten Einfluss auf gesunkene Informations- und Suchkosten im Vorfeld des Beitritts zur IPM hat (Abb. 2). Im Zuge der anonymisierten Onlinebefragung wurde den Befragten zusätzlich auch Anmerkungen als Freitextantwort ermöglicht. Hier wurden solche Netzwerke vereinzelt kritisiert, da sich informelle Hierarchien in den letzten Jahren herausgebildet hätten, die die Entwicklungs- und Entscheidungsprozesse maßgeblich beeinflussten.

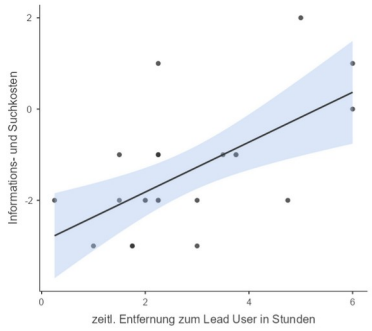


Abb. 2: Streudiagramm Cluster-Effekte.

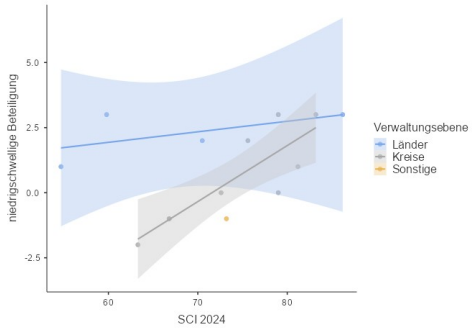


Abb. 3: Smart Citys fördern digitale Kompetenzen.

Investitionen für eine intakte Projektsteuerung wurden in hohem Umfang getätigt. Insbesondere die Öffentlichkeitsarbeit auf Konferenzen, die Präsenz in Fachzeitschriften sowie das Produktmanagement als zentrale Anlaufstelle zeichnen das Projekt aus. Ein Gesprächspartner deutete die Verringerung von Unsicherheiten für potenzielle Beteiligte dank dieser Maßnahmen an. Dem Projekt wurde durch eine breite Beteiligung angesehener Organisationen eine höhere Legitimität zugesprochen. Die Verbindlichkeit der IPM war auch für die Wahl der Software des späteren Lead Users ausschlaggebend, welcher deshalb das Masterportal gegenüber Marktalternativen bevorzugte. Während die Projektkommunikation durch Beauftragungen hochgradig professionalisiert wurde, blieben Investitionen z. B. in marktübliche Plattformen anstelle von Bitbucket auf der Strecke, was in Einzelfällen die Beteiligung verhindern kann (Abb. 4). Der eingesetzte Software-Stack, darunter *OpenLayers* und *Vue.js,* ist laut Befragungsteilnehmenden hingegen durchweg branchenüblich (Abb. 5).

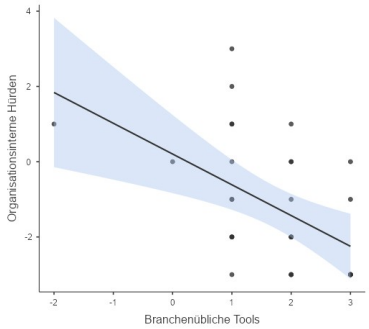


Abb. 4: Branchenübliche Tools erleichtern die Beteiligung.

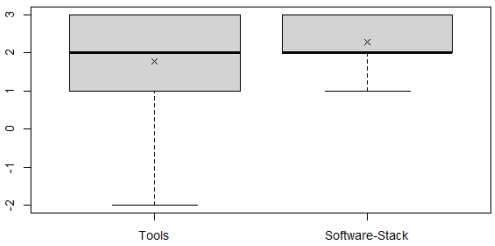


Abb. 5: Wahrnehmung der Branchenüblichkeit der eingesetzten Tools und des Software-Stacks.

Das Projekt zeichnet sich durch hohe Gegenseitigkeit der Community aus. Die Befragung zeigt, dass eigene Entwicklungen an das Projekt zurückgespielt werden (Abb. 6), zugleich ermöglicht die IPM aber auch eine nachhaltige Finanzierung beigetragener Komponenten, sofern diese durch das StraKom bestätigt wurden. Normative Prozesse begünstigten die kollaborative Entwicklung, da vielen Beteiligten erfolgreiche quelloffene GIS-Anwendungen bereits seit ihrem Studium vertraut sind. Unter den Mitgliedern wurden entsprechend wenig Bedenken aufgrund der offenen Lizenzierung des Masterportals vor ihrem Projektbeitritt geäußert. Die Mehrheit der IPM-Mitglieder (ca. 77 %) findet die bazaartypische Entwicklungsgemeinschaft wichtig, wenngleich die Relevanz der Vertragsstrukturen der IPM erheblich wichtiger eingeschätzt wird (ca. 92 %). Bemerkenswert ist allerdings, dass beteiligungsstarke Mitglieder die Bazaar-Governance durchweg wichtig finden (Abb. 7).

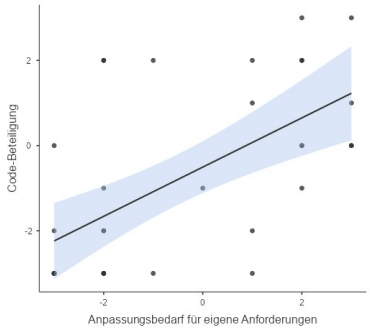


Abb. 6: Reziprozität der IPM.

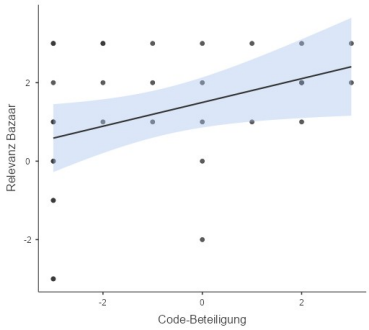


Abb. 7: Relevanz des Bazaars.

Vertrauen in die Universalität der Software – *KoliBri* ist eine Programmbibliothek für barrierefreie *Web Components*, die seit 2020 vom ITZBund entwickelt wird. Die Bibliothek wird in der Entwicklung von Websites eingesetzt, um eine barrierefreie Gestaltung der Benutzungsschnittstelle mit Hilfe standardisierter Webkomponenten zu ermöglichen. Dies stellt eine verhältnismäßig neue Entwicklung im Bereich der Webentwicklung dar, die ab 2014 durch das W3C empfohlen wurde [41, 42]. An der Entwicklung von KoliBri ist eine mittlere einstellige Zahl weiterer öffentlicher Einrichtungen beteiligt. Ein prominentes nachnutzendes Projekt stellt der UX-Standard KERN dar, welcher u. a. von den Ländern Hamburg und Schleswig-Holstein quelloffen entwickelt wird und seit Dezember 2024 als Basis für die *Digitale Dachmarke* der Bundesrepublik eingesetzt wird [43].

Das Projekt wurde durch einzelne Mitarbeitende als Bottom-up-Projekt innerhalb des ITZBund initiiert und zunächst behördenintern als Bibliothek für die standardisierte Einbindung barrierefreier Komponenten in Webanwendungen eingesetzt, ehe es unter einer OSS-Lizenz veröffentlicht wurde. Bereits der Veröffentlichungsprozess weist partizipative Züge externer Beteiligter auf. Die strategische Entscheidung das Projekt zu veröffentlichen wurde nicht zuletzt auch aufgrund einiger externer Anfragen legitimiert [44].

Der Innovationsprozess des Projektes ist seitdem weitestgehend geöffnet. Durch das Projektmanagement wird zweiwöchentlich eine öffentliche Community of Practice veranstaltet, in der Fragen zum Umgang mit KoliBri gestellt sowie selten auch Wünsche hinsichtlich der Weiterentwicklung geäußert werden können. Diese werden durch den Maintainer geleitet, protokolliert und im Repository veröffentlicht. Eine umfangreiche Dokumentation über Einrichtung und Betrieb der Bibliothek ist auf einer eigenen Website zu finden. Strategische Entscheidungen werden jedoch geschlossen durch das Produktmanagement im ITZBund getroffen, das sich laut Gesprächspartner als „cross-funktionales Team“ aus drei Personen zusammensetzt.

Die Entscheidung, KoliBri offen zu entwickeln, liegt maßgeblich an der Erwartung Innovationen extern akquirieren zu können, die durch die universelle Nachnutzbarkeit möglich ist. Investitionen bezüglich des Governance-Aufbaus insbesondere in Kommunikationsstrukturen blieben weitestgehend aus und fallen laut Gesprächspartner auch nicht in das Aufgabenspektrum einer öffentlichen Einrichtung. Aktuelle Sparmaßnahmen werden als Chance wahrgenommen redundante Softwareentwicklungen in Behörden zu minimieren, welche auch das Projekt durch einen Zustrom potenzieller neuer Beitragender positiv beeinflussen können.

Kirchturmpolitik am Fernsehturm – Das *ZMS* ist das Termin- und Zeitmanagementsystem des Landes Berlin, das seit 2014 für die Publikumssteuerung in den Berliner Behörden mit Bürgerkontakt entwickelt und seit 2016 sukzessive ausgerollt wird. Es stellt eine Neuentwicklung einer gleichnamigen proprietären Software dar, die auf Grund technischer Abhängigkeiten parallel weiterhin eingesetzt wird. Die OSS wird alleine von der Landeshauptstadt München (LHM) nachgenutzt, die die Software separat als Fork pflegt und nur punktuell zum originären Berliner Branch beiträgt. Das System hat seit 2016 in Berlin die gesetzlich verankerte Stellung eines IKT-Basisdienstes, der besondere Anforderungen an Bereitstellung und Pflege stellt und die Nutzung in Berliner Behörden verpflichtet.

Die Zusammenarbeit innerhalb des ZMS-Projektes ist aufgrund der Abspaltung durch die LHM in hohem Maße gehemmt und findet nur noch unregelmäßig statt. Aufgrund der gesetzlichen Stellung als IKT-Basisdienst sind strategische Entscheidungen der Berliner Projektkoordination stark durch die ministerielle Verantwortung reglementiert, der darüber hinaus externe Steuerungsbeteiligungen deutlich einschränkt. Die technische Produktpflege wird durch ein beauftragtes landeseigenes Unternehmen gewährleistet, das auch die Repositorien auf GitLab betreibt. Der Austausch zwischen der LHM und dem Berliner Produktmanagement ist auf fachlich-technischer Ebene informell, aber intensiv. Fachliche Anforderungen wurden zum Teil auch durch regelmäßige Anwendertreffen innerhalb der Berliner Verwaltungen identifiziert. Hierbei wurde auch eine hohe Anforderungsvielfalt und -volatilität festgestellt.

Auf strategischer Ebene findet laut Aussage eines Gesprächspartners kein Austausch statt. Erschwerend werden von beiden Behörden fachlich und technisch hochgradig konkurrierende Anforderungen an das System gestellt, die eine Zusammenführung der Forks kurz- bis mittelfristig eher unwahrscheinlich machen. Ein Gesprächspartner bezeichnet die engstirnigen Entscheidungen beider Behörden bezugnehmend auf die Wahrzeichen der beteiligten Städte scherzhaft als „Fernsehturm-“ beziehungsweise „(Frauen-)Kirchturmpolitik“.

Das Projekt ist durch die Ziehkräfte der beteiligten Städte geprägt: Während in Berlin ministerielle Anforderungen und die Pfadabhängigkeit zum Altsystem eine Öffnung blockieren, werden Innovationen und offenere Zusammenarbeit durch die langjährigen OSS-Erfahrungen der LHM forciert. Unsicherheiten über geringe Projektbeteiligung aufgrund hoher Projektkomplexität und -spezifität hemmt die Berliner Bestrebungen eine größere Community anzusprechen. Ein Gesprächspartner sieht hierin allerdings auch eine systematische Fehlannahme öffentlicher Einrichtungen, sehr unflexibel mit Anforderungen umzugehen.

5. Diskussion

Tabelle 2: OSS-Reifegrad

Offenheit [20]	Vergleichbare Konzepte	Fallbeispiel
3: externe Beteiligte an Strategie teilhaben lassen	Open Strategy [45]	Masterportal
2: Einforderung von Communitybeteiligung	Open Innovation [46]	KoliBri
1: Transparenz durch OSS-Lizenzierung	Open Access, Open Data	ZMS
0: interne hierarchische Entwicklung (*Projektstart*)		

Open-Source-Projekte im öffentlichen Sektor starten zumeist intern und geschlossen. Erst im Verlauf bilden sich Transparenz, Beteiligung oder Strategieöffnung aus (siehe Tab. 2). Einen archetypischen und vollumfänglichen Bazaar findet man indes in keinem der untersuchten Projekte. Unterschieden werden sollte hierbei jedoch der Kompetenzbereich. Während Innovationsprozesse in Entwicklungsgemeinschaften eher bazaartypische Formen annehmen und fachliche Zusammenarbeit über Organisationsgrenzen hinweg ermöglicht wird, finden sich auf strategischer Ebene bislang nur punktuell Vorhaben, Entscheidungsfindungsprozesse für Externe zu öffnen. Für Außenstehende kann jedoch gerade die strategische Mitbestimmung im Projekt relevant sein, um eigene Anforderungen im Projekt platzieren zu können.

Transaktionsspezifische Investitionen, die für die Umsetzung der Projekte notwendig waren, sind allesamt moderat zu bewerten. Die unterschiedlich hohen, tatsächlich getätigten Investitionen – insbesondere in Kommunikationsmaßnahmen, ermöglichen oder bremsen eine lebendigere OSS-Community. Die Untersuchung zeigt, dass ein hoher Zusammenhang zwischen der Unsicherheit über Projektbeteiligung und dem Willen, einen Bazaar zu etablieren, besteht. Vermeintliche Anforderungsspezifität und wahrgenommene Überkomplexität der Software hemmen ebenfalls die Entscheidung in Richtung des Communityaufbaus. Zugleich fördern hoher Innovationsdruck sowie normative Erwartungen eine stärkere Öffnung von Projekten. So wäre ohne die Bemühungen der LHM das ZMS-Projekt weitaus weniger bekannt und offen.

Das KoliBri-Projekt hat trotz hoher Universalität, hoher Innovationsabhängigkeit und geringer Komplexität eine recht überschaubare Community. Entgegen der Erwartung nehmen weniger Organisationen am Projekt teil und entwickeln die Komponenten eher selbst weiter. Das Not-invented-here-Syndrom [47, 48] könnte hier im Verdacht stehen, eine Beteiligung unattraktiv erscheinen zu lassen. Beim Masterportal nahm der Lead User die Rolle eines Multiplikatoren ein, um den sich informelle Netzwerke bildeten. Für die Identifikation des Lead Users lohnt auch ein Blick in die weiteren OSS-Aktivitäten und -Erwartungen der Beteiligten. Organisationen mit höheren OSS-Aktivitäten in anderen Projekten trugen in der Regel auch dem untersuchten OSS-Projekt mehr Code bei. Diesen Organisationen war auch die offene Bazaar-Governance wichtiger als die formalisierte Steuerung eines vertraglich gestalteten Netzwerkes. Schließlich kann auch das digitalpolitische Umfeld einer Behörde Digitalkompetenzen beeinflussen und so einen positiven Effekt auf kollaborative Softwareprojekte haben (Abb. 3).

Limitationen – Trotz sorgfältiger Überlegungen zur Vorgehensweise der Datenerhebung und ihrer Interpretation, weist die Untersuchung einige Limitationen auf. Die Studie basiert auf einer kleinen Stichprobe, bestehend aus drei Fallbeispielen mit fünf Interviews sowie 38 Befragungsteilnehmenden. Diese begrenzte Anzahl könnte die Generalisierbarkeit der Ergebnisse einschränken. Zudem wurden ausschließlich wahrgenommene Faktoren berücksichtigt, während objektive Daten zu Transaktions- und Produktionskosten nicht erhoben wurden. Dies könnte zu einer verzerrten Darstellung der Transaktionscharakteristika führen. In der Fallauswahl wurde der Governance-Modus „Markt“ nicht berücksichtigt, ebenso wenig wie eine Kontrollgruppe, etwa ein gescheitertes oder geschlossenes Softwareprojekt. Dieses Fehlen könnte die Vergleichbarkeit der Ergebnisse und die umfassende Analyse von Governance-Typen in Softwareprojekten einschränken. Bei den Interviews wurden ausschließlich männliche Partner befragt, was zu einer Geschlechterverzerrung führen könnte. Einige der Gespräche fan-

den nach Bekanntgabe der Kürzungen im Bundeshaushalt für das Jahr 2025 statt, was das Antwortverhalten der Interviewpartner beeinflusst haben könnte. Bei der Online-Befragung ist das Geschlechterverhältnis der Teilnehmenden unbekannt, was eine umfassende demografische Analyse erschwert. Wie bei den Interviews wurden auch hier nur wahrgenommene Faktoren berücksichtigt. Es besteht zuletzt die Möglichkeit, dass die Teilnehmenden sich selbst über- oder unterschätzen und die Likert-Skala unterschiedlich interpretieren.

6. Implikationen für Praxis und Wissenschaft

Öffentliche Einrichtungen scheinen sich dann für *Contribute* zu entscheiden, wenn sie sich einem hohen Innovationsdruck ausgesetzt sehen oder einzelne Organisationsmitglieder viel Sympathie für OSS-Communitys entgegenbringen. Besonders erfolgreich im Sinne hoher Nachnutzung sind jene Projekte, die klar definierte, inkludierende Steuerungsstrukturen aufweisen. Die Identifikation von trendführend nachnutzenden Organisationen kann dabei helfen, Innovationen in das Projekt zu holen. Empfehlenswert ist auch die Adaption und Erweiterung der Make-or-Buy-Entscheidung um klare Vorgaben zur Nachnutzung von OSS und die Abhängigkeit von externen Ressourcen im Anforderungs- und Beschaffungsprozess (siehe hierzu auch Münchens MBUC-Methode [49]). Die Methode scheint geeignet, um die Besonderheiten der Softwareentwicklung umfangreich berücksichtigen zu können.

Der ausgeprägte – wenngleich stark formalisierte – Bazaar des Masterportals ermöglichte eine hohe Nachnutzbarkeit im Sinne „alle für einen“. Gerade die Steuerungsgremien machen das Masterportal für Nachnutzende interessant. Inwieweit die Offenheit in Entscheidungsprozessen die Teilnahmemotivation steigern kann, ist zukünftig genauer zu untersuchen. Gleiches gilt auch für den Zusammenhang zwischen Entscheidungsbeteiligung und Nachnutzung, insbesondere vor dem Hintergrund des EfA-Prinzips. Die Ergebnisse deuten zudem darauf hin, dass normative Prozesse, wie branchenspezifische Erwartungen, den Erfolg des Bazaar-Modells beeinflussen können. So zeigt sich etwa, dass sich im Geosektor offene Organisationsformen leichter etablieren können, als unter Büroanwendungen. Eine vertiefende Untersuchung unter Zuhilfenahme bereits definierter Steuerungskategorien [50] könnte aufschlussreich sein, um organisationale Gepflogenheiten in unterschiedlichen Fachressorts besser zu verstehen. Schließlich wäre auch das *Ansteckungspotenzial* [51] der Offenheit in die vielen nichttechnischen Bereiche der Verwaltung spannend, Demil und Lecocq deuteten bereits das Potenzial des Bazaars in der Low-Tech-Industrie an [13].

Kontakt zum Autor:

Christian Weidner
Fraunhofer-Institut für Offene Kommunikationssysteme FOKUS
Kaiserin-Augusta-Allee 31
10589 Berlin
christian.weidner@fokus.fraunhofer.de

Literatur und Quellen

[1] Berger, Ariane: Die Entfesselung von D’Artagnan. In: Seckelmann; Margrit und Brunzel; Marco (Hrsg.) *Handbuch Onlinezugangsgesetz: Potenziale - Synergien - Herausforderungen*. S. 75–93. Springer, Berlin, Heidelberg **2021**, https://doi.org/10.1007/978-3-662-62395-4_4.

[2] KGSt: Anforderungen an das OZG aus kommunal-fachlicher Sicht zur Diskussion im Kommunalgremium des IT-Planungsrats. Kommunale Gemeinschaftsstelle für Verwaltungsmanagement, Köln, **2020**.

[3] Halsbenning, Sebastian: Digitalisierung öffentlicher Dienstleistungen: Herausforderungen und Erfolgsfaktoren der OZG-Umsetzung in der Kommunalverwaltung. *HMD* 58 (5), 1038–1053, **2021**, https://doi.org/10.1365/s40702-021-00765-5.

[4] Frenzel, Torsten; Köster, Thomas: Das „Einer für Alle“-Prinzip ist gescheitert. **2022**, https://background.tagesspiegel.de/smart-city/das-einer-fuer-alle-prinzip-ist-gescheitert.

[5] Bonaccorsi, Andrea; Rossi, Cristina: Why Open Source software can succeed. *Research Policy* 32 (7), 1243–1258, **2003**, https://doi.org/10.1016/S0048-7333(03)00051-9.

[6] Wright, Nataliya Langburd; Nagle, Frank; Greenstein, Shane: Open source software and global entrepreneurship. *Research Policy* 52 (9), 104846, **2023**, https://doi.org/10.1016/j.respol.2023.104846.

[7] Silic, Mario; Back, Andrea: Open Source Software Adoption: Lessons from Linux in Munich. *IT Professional* 19 (1), 42–47, **2017**, https://doi.org/10.1109/MITP.2017.7.

[8] Burgfried, Matthias; Reckert-Lodde, Andreas: Die Deutsche Verwaltungscloud-Strategie. *Datenschutz Datensich* 46 (10), 611–615, **2022**, https://doi.org/10.1007/s11623-022-1668-6.

[9] Osborne, Cailean; Sharratt, Paul; Foster, Dawn; Boehm, Mirko: A Toolkit for Measuring the Impacts of Public Funding on Open Source Software Development, **2024**, https://doi.org/10.48550/arXiv.2411.06027.

[10] Markus, M. Lynne: The governance of free/open source software projects: monolithic, multidimensional, or configurational? *J Manage Governance* 11 (2), 151–163, **2007**, https://doi.org/10.1007/s10997-007-9021-x.

[11] O’Mahony, Siobhán: The governance of open source initiatives: what does it mean to be community managed? *J Manage Governance* 11 (2), 139–150, **2007**, https://doi.org/10.1007/s10997-007-9024-7.

[12] Raymond, Eric: The cathedral and the bazaar. *Know Techn Pol* 12 (3), 23–49, **1999**, https://doi.org/10.1007/s12130-999-1026-0.

[13] Demil, Benoît; Lecocq, Xavier: Neither Market nor Hierarchy nor Network: The Emergence of Bazaar Governance. *Organization Studies* 27 (10), 1447–1466, **2006**, https://doi.org/10.1177/0170840606067250.

[14] Linåker, Johan; Robles, Gregorio; Bryant, Deborah; Muto, Sachiko: Open Source Software in the Public Sector: 25 Years and Still in Its Infancy. *IEEE Software* 40 (4), 39–44, **2023**, https://doi.org/10.1109/MS.2023.3266105.

[15] Williamson, Oliver E.: The Economics of Organization: The Transaction Cost Approach. *American Journal of Sociology* 87 (3), 548–577, **1981**.

[16] David, Robert J.; Han, Shin-Kap: A systematic assessment of the empirical support for transaction cost economics. *Strategic Management Journal* 25 (1), 39–58, **2004**, https://doi.org/10.1002/smj.359.

[17] Greiner, Martina; Goodhue, Dale: Make-or-Buy in the Age of Open Source: A Transaction Cost Analysis. In: *ICIS 2005 Proceedings*. S. 689–700. , Las Vegas, NV, USA **2005**.

[18] Watson, Richard T.; Boudreau, Marie-Claude; Greiner, Martina; Wynn, Donald; York, Paul; Gul, Rusen: Governance and global communities. *Journal of International Management* 11 (2), 125–142, **2005**, https://doi.org/10.1016/j.intman.2005.03.006.

[19] Powell, Walter W.: Neither Market Nor Hierarchy: Network Forms of Organization. *Research in Organizational Behaviour* 12, 295–336, **1990**.

[20] Splitter, Violetta; Dobusch, Leonhard; von Krogh, Georg; Whittington, Richard; Walgenbach, Peter: Openness as Organizing Principle: Introduction to the Special Issue. *Organ Stud* 44 (1), 7–27, **2023**, https://doi.org/10.1177/01708406221145595.

[21] Oezbek, Christopher; Prechelt, Lutz; Thiel, Florian: The onion has cancer: some social network analysis visualizations of open source project communication. In: *Proceedings of the 3rd International Workshop on Emerging Trends in Free/Libre/Open Source Software Research and Develop-*

ment. S. 5–10. Association for Computing Machinery, New York, NY, USA **2010**, https://doi.org/10.1145/1833272.1833274.

[22] Chesbrough, Henry W.; Appleyard, Melissa M.: Open Innovation and Strategy. *California Management Review* 50 (1), 57–76, **2007**, https://doi.org/10.2307/41166416.

[23] Pfeffer, Jeffrey; Salancik, Gerald R.: The external control of organizations: a resource dependence perspective. Stanford Business Books, Stanford, CA, **2003**.

[24] Hippel, Eric von: Lead Users: A Source of Novel Product Concepts. *Management Science* 32 (7), 791–805, **1986**, https://doi.org/10.1287/mnsc.32.7.791.

[25] Williamson, Oliver E.: The Economics of Governance: Framework and Implications. *Zeitschrift für die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics* 140 (1), 195–223, **1984**.

[26] DiMaggio, Paul J.; Powell, Walter W.: The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48 (2), 147–160, **1983**, https://doi.org/10.2307/2095101.

[27] Sydow, Jörg; Schreyögg, Georg; Koch, Jochen: Organizational path dependence: Opening the black box. *The Academy of Management Review* 34 (4), 689–709, **2009**, https://doi.org/10.5465/AMR.2009.44885978.

[28] Dobusch, Leonhard: Windows versus Linux. VS Verlag für Sozialwissenschaften Wiesbaden, Wiesbaden, **2008**.

[29] Yin, Robert K.: Case study research and applications: design and methods. SAGE, Los Angeles, CA, **2018**.

[30] Dresing, Thorsten; Pehl, Thorsten: Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende. Eigenverlag, Marburg, **2018**.

[31] Mayring, Philipp; Fenzl, Thomas: Qualitative Inhaltsanalyse. In: Baur; Nina und Blasius; Jörg (Hrsg.) *Handbuch Methoden der empirischen Sozialforschung*. S. 633–648. Springer Fachmedien, Wiesbaden **2019**, https://doi.org/10.1007/978-3-658-21308-4_42.

[32] Anderson, Erin; Coughlan, Anne T.: International Market Entry and Expansion Via Independent or Integrated Channels of Distribution. *Journal of Marketing* 51 (1), 71–82, **1987**, https://doi.org/10.2307/1251145.

[33] Klein, Saul; Frazier, Gary L.; Roth, Victor J.: A Transaction Cost Analysis Model of Channel Integration in International Markets. *Journal of Marketing Research* 27 (2), 196–208, **1990**, https://doi.org/10.1177/002224379002700207.

[34] Zaheer, Akbar; Venkatraman, N.: Determinants of Electronic Integration in the Insurance Industry: An Empirical Test. *Management Science* 40 (5), 549–566, **1994**.

[35] Albino, Vito; Berardi, Umberto; Dangelico, Rosa Maria: Smart Cities: Definitions, Dimensions, Performance, and Initiatives. *Journal of Urban Technology* 22 (1), 3–21, **2015**, https://doi.org/10.1080/10630732.2014.942092.

[36] Bitkom: Smart City Index 2024. Ausführliche Ergebnisse. Bitkom e. V., Berlin, **2024**.

[37] Porter, Michael E.: Clusters and the New Economics of Competition: Harvard Business Review. *Harvard Business Review* 76 (6), 77–90, **1998**.

[38] Chronotrains - Zugreisekarte für Europa: https://www.chronotrains.com, zugegriffen 19.09.2024.

[39] Schliffke, Nicholas: Das Beispiel Masterportal – ein OS Erfolgsmodell für die öffentliche Verwaltung? In: *Tagungsband FOSSGIS 2024*. 120–121, **2024**.

[40] Masterportal Hamburg: Dienstleister: https://www.masterportal.org/referenzen/dienstleister, zugegriffen 18.12.2024.

[41] Page, Wilson: The state of Web Components. **2015**, https://hacks.mozilla.org/2015/06/the-state-of-web-components, zugegriffen 27.08.2024.

[42] Rubano, Vincenzo; Vitali, Fabio: Making accessibility accessible: strategy and tools. In: *2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC)*. S. 1–6 **2021**, https://doi.org/10.1109/CCNC49032.2021.9369632.

[43] BMI: Digitale Dachmarke – Start der Pilotphase: https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2024/12/digitale-dachmarke.html?nn=9390260, zugegriffen 13.01.2025.

[44] Pietta, Isabell; Pfaff, Robin; Graf, Christian: Befreit eure Schätze! Open Source in der öffentlichen Verwaltung. **2024**, https://www.oeffentliche-it.de/-/befreit-eure-schaetze-open-source-in-der-oeffentlichen-verwaltung, zugegriffen 04.09.2024.

[45] Whittington, Richard; Cailluet, Ludovic; Yakis-Douglas, Basak: Opening Strategy: Evolution of a Precarious Profession. *British Journal of Management* 22 (3), 531–544, **2011**, https://doi.org/10.1111/j.1467-8551.2011.00762.x.

[46] Chesbrough, Henry W.: The Era of Open Innovation. *MIT SMR* 44 (3), 35–41, **2003**.

[47] Lichtenthaler, Ulrich; Ernst, Holger: Attitudes to externally organising knowledge management tasks: a review, reconsideration and extension of the NIH syndrome. *R&D Management* 36 (4), 367–386, **2006**, https://doi.org/10.1111/j.1467-9310.2006.00443.x.

[48] Stefi, Anisa; Hess, Thomas: To Develop or to Reuse? Two Perspectives on External Reuse in Software Projects. In: Fernandes; João M., Machado; Ricardo J., und Wnuk; Krzysztof (Hrsg.) *Software Business*. S. 192–206. Springer International Publishing, Cham **2015**, https://doi.org/10.1007/978-3-319-19593-3_18.

[49] Strauß, Friedrich; Siemon, Helmut: „Make, Buy, Use oder Compose?“ als zentrales Element des Anforderungsmanagmentprozesses in einer großen Kommune. In: Rausch, Andreas und Sihling, Marc (Hrsg.) *Software & Systems Engineering Essentials 2013*. S. 156–170, **2013**.

[50] Schrape, Jan-Felix: Open-Source-Projekte: vom Nischenphänomen zum integralen Bestandteil der Softwareindustrie. *WSI* 69 (8), 603–612, **2016**, https://doi.org/10.5771/0342-300X-2016-8-603.

[51] Borgatti, Stephen P.; Foster, Pacey C.: The Network Paradigm in Organizational Research: A Review and Typology. *Journal of Management* 29 (6), 991–1013, **2003**, https://doi.org/10.1016/S0149-2063_03_00087-4.

Anhang

Tabelle 3: Operationalisierung der Leitfaden- und Kodierkategorien.

Variable	Dimensionen (pro Variable sortiert nach Kodieranteil)
Investitionsspezifität	1. Kommunikationsmaßnahmen (z. B. Pressemitteilungen, Team-Events etc.), 2. Beauftragungen an externe Dienstleister, 3. allgemeine Projektsteuerung, 4. Tools für Projektzusammenarbeit (z. B. GitLab, Trello, Mailinglisten etc.), 5. Rechtsberatung und Lizenzwahl, 6. Softwarespezifische Abhängigkeiten (z. B. Programmierumgebung, -sprache), 7. bürokratische Hürden und Überzeugungsarbeit in eigener Organisation, 8. Personal und Schulungen, 9. Anpassungskosten für Nachnutzung.

Anhang

Variable	Dimensionen (pro Variable sortiert nach Kodieranteil)
Unsicherheit	1. Einzigartigkeit der Software, 2. Zusammenarbeitspotenzial, 3. Universalität der Software, 4. Komplexität der Software, 5. Vorgängersystem, 6. Rechtsunsicherheiten, 7. Relevanz der Software für eigenen Organisationszweck.
Frequenz	1. Häufig wechselnde/ hinzukommende fachliche Anforderungen, 2. Sicherheitsrisiko, das häufige Änderungen am Programm unerlässlich macht.
Ressourcen-abhängigkeit	1. Wahrgenommene Abhängigkeit externer Innovationskraft 2. Identifikation relevanter Lead User.
Isomorphie	1. Ansichten einzelner Organisationsmitglieder (insb. Aufgabenkritik), 2. akad. oder berufl. Sozialisierung der Mitglieder und informelle Erwartungen, 3. rechtliche Vorgaben und politische Entscheidungen, 4. technische Abhängigkeiten (z. B. zu Altsystemen), 5. finanzielle Abhängigkeiten, 6. mimetisches Verhalten, 7. Nutzungsakzeptanz.
Sonstiges	1. Netzwerkeffekte, 2. ex-ante und ex-post Kosten, 3. Reziprozität, 4. Allmende-Dilemma.

Tabelle 4: Stichprobenbeschreibung. Verwaltungsebene und ostdeutsche Beteiligung mit *Gegenprobe der Gesamtpopulation* kursiv gedruckt.

Variable	N	Fehlend	Mittel	Md.	Std.-Abw.	Min	Max
Organisationale Zuordenbarkeit	38	0	0.658	1	0.474	0	1
Führungsverantwortung	31	7	0.52	1	0.508	0	1
aktuelle Nutzung der Software	38	0	0.947	1	0.227	0	1
Mitglied der Implementierungspartnerschaft	38	0	0.974	1	0.160	0	1
Communitybeteiligung	38	0	0.763	1	0.425	0	1
Smart City Index 2024	14	24	72.30	72.9	9.210	54.7	86.2
Entfernung LGV Hamburg in *h*	24	14	3.820	3.50	1.430	1.00	6.75
Entfernung Lead User (Frankfurt am Main) in *h*	24	14	2.810	2.25	1.480	0.25	6.00
Verwaltungsebene (1: Bund, 2: Land, 3: Kreis, 4: Gemeinde, 5: Sonst.)	38	0	2.870	3	0.906	1	5
IPM gesamt	*49*	*0*	*2.730*	*3*	*0.878*	*1*	*5*
ostdeutsche Beteiligung	25	13	0.160	0	0.374	0	1
IPM gesamt	*49*	*0*	*0.143*	*0*	*0.354*	*0*	*1*

Von proprietär zu Open-Source - Umstellung der kommunalen GDI bei der Stadt Reutlingen

Simon Kondic, Linus Lambrecht, Melanie Machts

Zusammenfassung

Die Stadt Reutlingen, eine Großstadt in Baden-Württemberg, stellte ihre kommunale Geodateninfrastruktur (GDI) von proprietärer auf Open-Source Software um. Ziel war eine größere digitale Souveränität, Lizenzkostenreduktion und Unabhängigkeit von kommerziellen Anbietern. Im Vortrag wird die Vorgehensweise der Systemmigration und der Aufbau der neuen Open-Source GDI vorgestellt.

Key Facts Stadt Reutlingen

Reutlingen ist eine der neun Großstädte Baden-Württembergs mit 118.000 Einwohnern. Die Stadtverwaltung beschäftigt mehr als 2.600 Mitarbeiterinnen und Mitarbeiter. Die Hochschulstadt und frühere Reichsstadt ist die größte Stadt und zugleich Kreisstadt im Landkreis Reutlingen.

Die Stadt gehört zur Region Neckar-Alb und zur Metropolregion Stuttgart. Nahebei befindet sich die Mittelstadt Tübingen (12 Kilometer westlich), die nächste Großstadt ist Stuttgart (31 Kilometer nördlich). Mit der Spreuerhofstraße besitzt Reutlingen die engste Straße der Welt. 17,5 Prozent der Gemarkung Reutlingens gehören zum Biosphärengebiet Schwäbische Alb, weshalb sich die Stadt auch als das „Tor zur Schwäbischen Alb“ bezeichnet.

Key Facts Sachgebiet GeoInformationsSystem der Stadt Reutlingen

Das Sachgebiet ist dem Baudezernat, Amt für Stadtentwicklung und Vermessung, Abteilung Vermessung zugeordnet. Die Stadt Reutlingen ist Untere Vermessungsbehörde in Baden-Württemberg.

Das Geoinformationssystem wurde 1997 aufgebaut mit dem Liegenschaftskataster als erste Fachschale. Im Team sind fünf Personen. Wir sehen uns als stadtinterner „GIS-Dienstleister“ mit ca. 250.000 internen Zugriffen auf das Web-GIS Auskunftssystem im Jahr. Die Datenbank beinhaltet ca. 75 verschiedene Datenbankthemen. Ungefähr 29 Themen davon werden von den Fachämtern über GIS-Schnittstellen fortgeführt. 2020 wurde ein Geoportal im Internet für Bürger mit verschiedenen Themenkarten veröffentlicht. Die Umstellung der GDI auf Open-Source wurde von 2021 bis Ende 2024 durchgeführt.

Motivation zur Umstellung auf Open-Source-Software

Die GDI war ursprünglich stark von Autodesk abhängig, da das amtliche Liegenschaftskataster mit dieser Software geführt wurde. Mit dem Wechsel der ALKIS-Lösung zum Land Baden-Württemberg entfiel diese Abhängigkeit von der bestehenden GIS-Software. Dies eröffnete die Möglichkeit, eine alternative Software für den Aufbau einer modernen GDI zu suchen.

Bei der Auswahl der neuen Systeme stand im Vordergrund, die softwaretechnische Abhängigkeit zu reduzieren und digitale Souveränität zu gewährleisten. Gleichzeitig sollte die Migration eigenständig erfolgen können, ohne auf die Unterstützung eines bestimmten Herstellers oder Dienstleisters angewiesen zu sein. Ein weiteres wesentliches Ziel war die Umstellung auf eine kostengünstige und frei verfügbare GDI. Durch die konsequente Nutzung von Standards, insbesondere OGC-konformer Dienste, sollte eine zukunftssichere Infrastruktur geschaffen werden.

Systemmigration

Ausgangslage war eine überwiegend proprietäre Geodateninfrastruktur, die auf einer Oracle-Datenbank basierte und mit verschiedenen Autodesk-Programmen administriert wurde. Die Hauptkomponenten dieser Infrastruktur umfassten die Datenbank (Oracle/Topobase), deren Kartendarstellung im Web-GIS Auskunftssystem über Mapguide Maestro und Autodesk Infrastructure Studio erfolgte. Die

Datenaufbereitung und -pflege wurde mit Autodesk AutoCAD MAP 3D durchgeführt, womit auch das Einpflegen, Ändern und Löschen von Geoobjekten erfolgte. Als Kartenserver diente MapGuide Open-Source, während die aufbereiteten Daten den Nutzern über das Web-GIS TBview von der Firma Gerst Ingenieure GmbH bereitgestellt wurden.

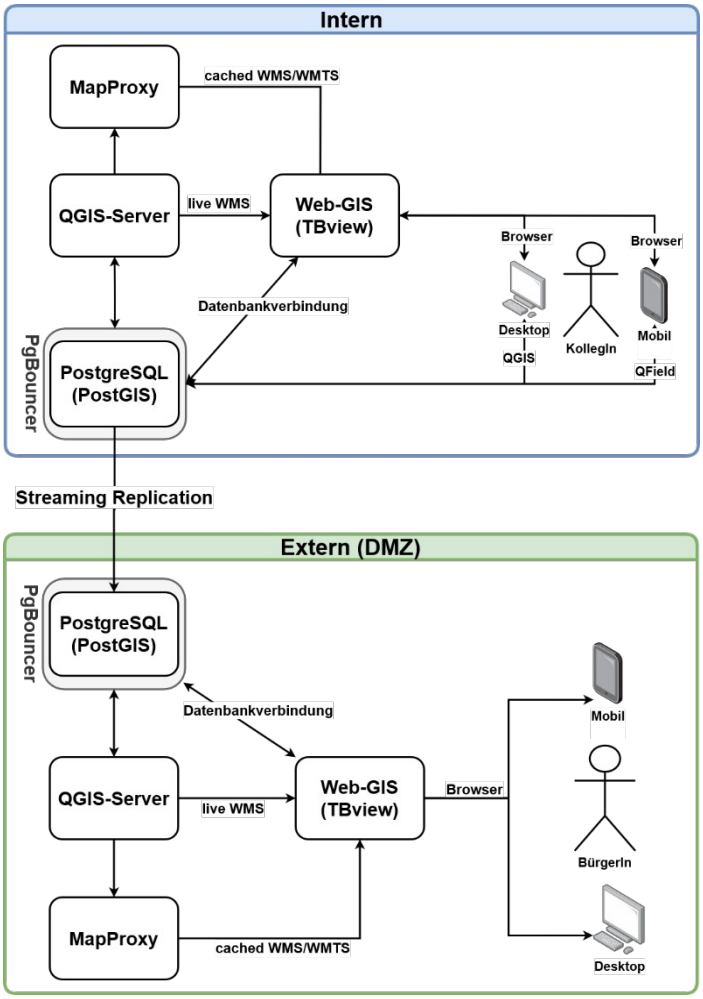


Abbildung 1: Aufbau der Open-Source-GDI (Techstack)

Im Zuge der Umstellung wurden drei zentrale Komponenten durch Open-Source-Produkte ersetzt: Die bisherige Datenbank wurde durch PostgreSQL (mit PgBouncer für das connection-pooling) abgelöst, die Datenaufbereitung und -fortführung erfolgt mit QGIS, und als neuer Kartenserver kommt QGIS-Server zum Einsatz. Die Migration der GDI wurde schrittweise und themenbezogen durchgeführt, wobei sie sich an den einzelnen Fachschalen der Datenbank orientierte:

Zunächst wurde die Datenmigration von Oracle nach PostgreSQL/PostGIS durchgeführt. Dieser Prozess war der Schwerpunkt einer Bachelorthesis [1] in den Jahren 2021/2022. Zunächst wurden die vorhandenen Daten in Oracle gesichtet und analysiert, um deren Struktur zu verstehen und die relevanten Anforderungen für die Migration zu identifizieren – beispielsweise hinsichtlich der Topologie, Datentypen und unterschiedlicher Syntax. Anschließend wurde ein möglichst automatisierter Migrationsprozess erarbeitet, für den die Software FME von Safe Software Inc. und Ora2Pg genutzt wurde.

Nach der erfolgreichen Migration wurden die bisherigen Themen aus MapGuide in QGIS nachgebildet. Dies konnte sowohl durch Fachkräfte als auch durch Auszubildende durchgeführt werden. Da die

Symbole ursprünglich als Autodesk-Blockelemente vorlagen, mussten sie mit Inkscape neu als SVG-Dateien erstellt werden.

Anschließend erfolgte die Publikation der Themen mit QGIS-Server und MapProxy über WMS/WMTS. QGIS-Projekte konnten direkt über den Server bereitgestellt werden, während MapProxy für gecachte Dienste genutzt wird. Die neu aufbereiteten Themen wurden daraufhin in das Web-GIS TBview integriert, indem sowohl die PostgreSQL-Datenbanken als auch die WMS- und WMTS-Dienste eingebunden wurden.

Im letzten Schritt wurden die Projekte umgestellt, die für die Datenerfassung von den Fachämtern verwendet werden. Für komplexe Themen wurde auf QGIS bzw. QField für die mobile Erfassung gewechselt. Einfache Themen werden direkt in TBview erfasst und wurden hier entsprechend umgestellt. Dies betraf beispielsweise Anwendungen zur Erfassung von Straßenschäden, Leerständen im Gewerbebereich oder Fahrradabstellanlagen.

Für die Veröffentlichung der Geodaten über das Geoportal der Stadt im Internet wurde eine baugleiche GDI in einer DMZ (Demilitarized Zone) aufgebaut. Die Synchronisierung der Datenbank zwischen der internen GDI und der DMZ-GDI erfolgt über Streaming Replication. Dabei wird die PostgreSQL-Datenbank kontinuierlich mit den aktuellen Änderungen aus der internen Umgebung aktualisiert.

Fazit:

Durch die Umstellung der GDI auf Open-Source-Komponenten wurde eine moderne und zukunftssichere Infrastruktur geschaffen, die den heutigen Standards entspricht. OGC-konforme Geodatendienste ermöglichen es nun, diese auch in weitere interne GIS/CAD-Systeme einzubinden, wodurch die Geodaten systemübergreifend genutzt werden können. In einem nächsten Schritt sollen die Geodatendienste auch für eine externe Weitergabe konfiguriert werden, um sie zukünftig im Internet veröffentlichen zu können. Mit dem Umstieg auf freie Open-Source-Software konnten im Fachbereich GeoInformationsSystem der Stadt Reutlingen die laufenden Kosten für Softwarelizenzen und -wartung deutlich reduziert werden.

Kontakt zum Autor:

Simon Kondic, Linus Lambrecht und Melanie Machts
Stadt Reutlingen
Marktplatz 22
72760 Reutlingen
07121/303-2507
gis@reutlingen.de

Literatur:

[1] Machts, Melanie: Migration einer Oracle Geodatenbank in ein Open-Source-Geodatenbanksystem bei der Stadt Reutlingen, Stuttgart, 2022

Entwicklung eines LLM-basierten Assistenten für die Suche nach Geodaten

Simeon Wetzel[1], Matthes Rieke[2]

[1]Professur für Geoinformatik, TU Dresden, [2]52°North Spatial Information Research GmbH

In Geodaten- und Forschungsdateninfrastrukturen (GDI/FDI) basiert die Datensuche primär auf Metadaten. Diese Suchfunktionalität wird hierbei über Metadatenkataloge oder Open Data Portale wie CKAN[1] oder GeoNetwork[2] realisiert. Diese Dienste bieten eine Volltextsuche zur Identifikation relevanter Metadatensätze, die durch Suchfacetten und Filter ergänzt werden kann [1]. Die implementierte Volltextsuche beruht dabei auf einem lexikalischen Ansatz, der auf die exakte Übereinstimmung von Wörtern oder Phrasen zwischen Suchanfragen und Metadatensätzen basiert [2].

Die lexikalische Suche weist jedoch bedeutende Einschränkungen auf:

- Suchende müssen die spezifische Terminologie der Metadaten kennen und die richtigen Suchbegriffe verwenden, um relevante Ergebnisse zu erhalten. Diese Fachkenntnis ist besonders in interdisziplinären Portalen nicht selbstverständlich.
- Sprachliche Variationen in den Suchanfragen – etwa durch Tippfehler, Abkürzungen, Synonyme oder Akronyme – können die Ergebnisqualität erheblich beeinträchtigen (vgl. „vocabualary mismatch problem“ [3]).

Zur Bewältigung dieser Herausforderungen existieren verschiedene Lösungsansätze. Die Integration kontrollierter Vokabulare oder Ontologien ermöglicht beispielsweise das Vorschlagen verwandter Suchbegriffe wie Synonyme im Rahmen der Volltextsuche [4]. Zudem eröffnen aktuelle Entwicklungen im Bereich neuronaler Netze und Sprachmodelle neue Möglichkeiten, die Semantik von Suchanfragen und -ergebnissen zu erfassen und deren Relevanz präziser zu bestimmen [5]. Entsprechende Anwendungen im Kontext der (Geo-)Metadatensuche wurden bereits entwickelt [6].

Allerdings garantiert selbst die erfolgreiche Identifikation relevanter Metadatensätze nicht, dass die zugehörigen Datensätze den Anforderungen der Suchenden entsprechen. Mögliche Ursachen hierfür sind:

- Unzureichende Repräsentation der Datensatzinhalte und -eigenschaften in den Metadaten
- Unvollständige oder ungenaue Metadaten
- Spezifische Nutzeranforderungen, die in den verfügbaren Metadaten nicht erfasst sind

Um diese verbleibenden Herausforderungen anzugehen, wurde ein Framework, das auf einem Large Language Model (LLM) basiert, entwickelt. Es implementiert einen innovativen, dialogorientierten Suchansatz. Das System ist dabei bewusst für die Integration verschiedener Open-Source-Sprachmodelle ausgelegt. Neben proprietären Modellen können insbesondere freie LLMs wie Llama[3] oder Mixtral[4] eingebunden werden. Das System kombiniert einen Chatbot für die Interaktion in natürlicher Sprache mit einem semantischen Suchindex für Geodaten und deren Metadaten. Das Framework ist in einem leicht konfigurierbaren Webserver gekapselt, der eine OpenAPI-basierte REST-API bereitstellt. Diese API bietet mehrere Endpunkte für Chatbot-Inferenzen und für die Verwaltung verbundener (Meta-)Datenressourcen. Abbildung 1 zeigt einen Überblick der Systemarchitektur.

1 https://ckan.org/
2 https://geonetwork-opensource.org/
3 https://www.llama.com/
4 https://mistral.ai/news/mixtral-of-experts/

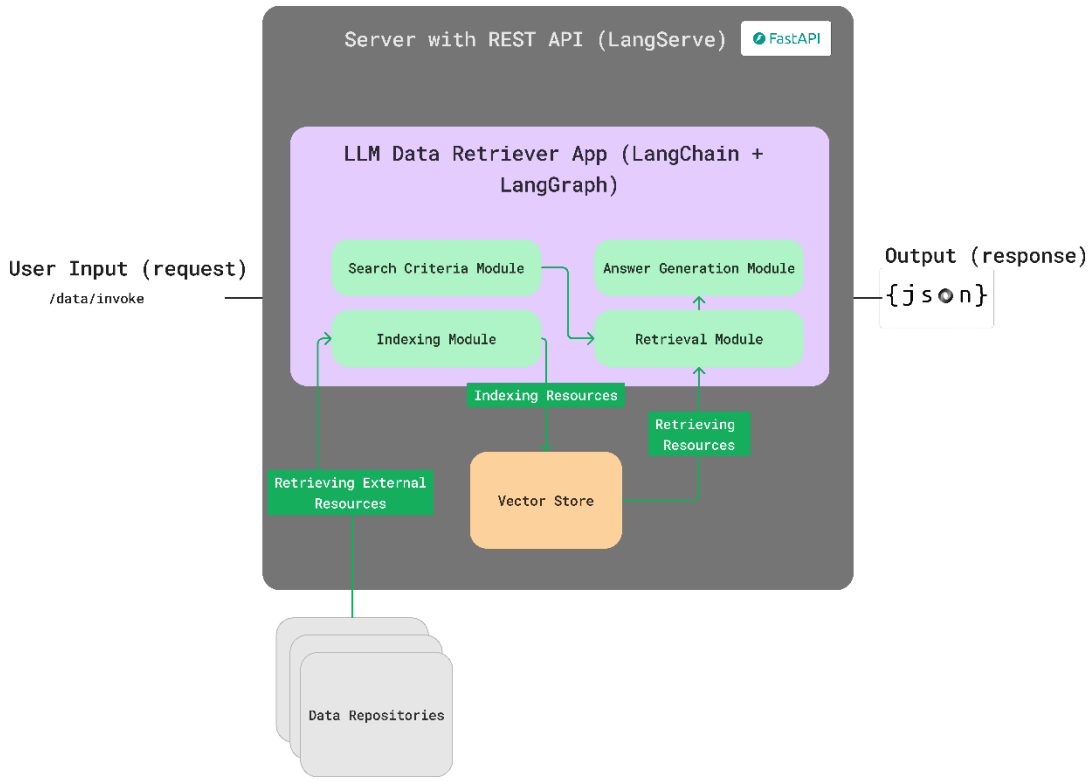


Abbildung 1: Architekturübersicht des entwickelten LLM-Frameworks.

Anders als bei konventionellen Suchsystemen ermöglicht dieser Ansatz eine präzise Erfassung der Nutzeranforderungen durch interaktiven Dialog. Der Chatbot kann gezielt Rückfragen stellen, um Suchintentionen zu klären und spezifische Anforderungen zu konkretisieren. Durch die Integration eines speziellen Suchindex, der für die semantische Indexierung (Vektordatenbank) optimiert ist, können sowohl die eigentlichen Geodaten als auch deren Metadaten gespeichert und durchsucht werden. Dies ermöglicht eine semantische Suche nach den in den Daten vorhandenen Attributen oder Features.

Abbildung 2 zeigt einen Demo-Client, der das entwickelte LLM-Framework nutzt. Eine Beispielsuche nach „Gebäuden in Dresden“ zeigt die Verarbeitung von Anfragen in natürlicher Sprache, sowie die Möglichkeit, das LLM zur Interpretation der gefundenen Ergebnisse zu nutzen.

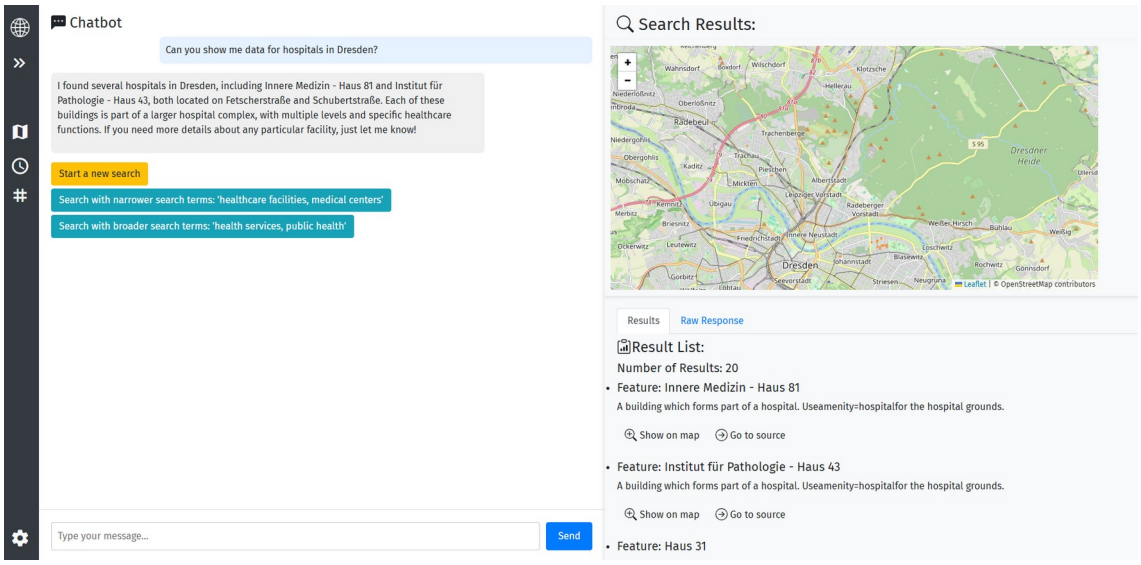


Abbildung 2: Chatbot-Client zum Testen des Suchassistenten.

Ein konkretes Beispiel verdeutlicht die Vorteile dieses Ansatzes: Bei der Suche nach "Krankenhäusern mit Notaufnahme" werden nicht nur Datensätze gefunden, die diese Begriffe in ihren Metadaten enthalten. Vielmehr können auch die in den Geodaten selbst vorhandenen Attribute wie "emergency=yes" oder "healthcare=hospital" für die Suche genutzt werden. Der Chatbot kann dabei gezielt nach weiteren Anforderungen fragen, etwa ob bestimmte Fachabteilungen benötigt werden. Diese Informationen werden dann mit den tatsächlich in den Daten vorhandenen Attributen abgeglichen.

Dies ermöglicht es, auch implizite oder komplexe Zusammenhänge zwischen Suchanfragen und verfügbaren Datensätzen zu erkennen und somit die Diskrepanz zwischen Nutzeranforderungen und Datenverfügbarkeit zu minimieren.

Kontakt zum Autor:

Simeon Wetzel
Technische Universität Dresden
Professur für Geoinformatik
Helmholtzstraße 10, 01069 Dresden
simeon.wetzel@tu-dresden.de

References**:**

[2] Formal, T., Piwowarski, B., & Clinchant, S. (2022). Match Your Words! A Study of Lexical Matching in Neural Information Retrieval. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13186 LNCS*, 120–127. https://doi.org/10.1007/978-3-030-99739-7_14

[3] Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The Vocabulary Problem in Human-System Communication. *Communications of the ACM*, *30*(11), 964–971. https://doi.org/10.1145/32206.32212

[1] Hervey, T., Lafia, S., & Kuhn, W. (2020). Search Facets and Ranking in Geospatial Dataset Search. *Leibniz International Proceedings in Informatics, LIPIcs*, *177*(5), 1–5. https://doi.org/10.4230/LIPIcs.GIScience.2021.I.5

[4] Jiang, S., Hagelien, T. F., Natvig, M., & Li, J. (2019). Ontology-Based Semantic Search for Open Government Data. *Proceedings - 13th IEEE International Conference on Semantic Computing, ICSC 2019*, 7–15. https://doi.org/10.1109/ICOSC.2019.8665522

[5] Thakur, N., Reimers, N., Rücklé, A., Srivastava, A., & Gurevych, I. (2021). *BEIR: A Heterogenous Benchmark for Zero-shot Evaluation of Information Retrieval Models*. *NeurIPS*. http://arxiv.org/abs/2104.08663

[6] Wetzel, S., & Mäs, S. (2024). Context-Aware Search for Environmental Data Using Dense Retrieval. *ISPRS International Journal of Geo-Information*, *13*(11), 380. https://doi.org/10.3390/ijgi13110380

Was wäre, wenn wir Algorithmen demokratisieren? Neue kollaborative Infrastrukturen für Modelle und Simulationen in digitalen Stadtzwillingen

Rico H Herzog[1]

[1] City Science Lab, HafenCity Universität Hamburg
DOI: 10.5281/zenodo.14615602

Zusammenfassung: Neue kollaborative Infrastrukturen für urbane digitale Zwillinge können eine bessere Teilhabe und Zugang zur Simulationsentwicklung schaffen. Am Beispiel von drei Open-Source-Tools aus dem „Connected Urban Twins“-Projekt wird demonstriert, wie Szenarien für Stadtentwicklungsprozesse gemeinsam entwickelt, analysiert und zugänglich gemacht werden können. Ziel ist eine nachhaltigere und inklusivere Planung, die durch pluralistische und agonistische Ansätze technologisch unterstützt wird.

Schlüsselwörter: Urban Digital Twin, Co-Modelling, Multi-Modeling, Kollaboration

Einführung

Als Teil der Antwort auf die vielfältigen Herausforderungen des 21. Jahrhunderts widmen sich immer mehr Städte und Kommunen der Entwicklung und dem Einsatz von Urbanen Digitalen Zwillingen (UDZ). Neben einer großen Bandbreite an Einsatzmöglichkeiten im Bereich der Visualisierung und Informationsbereitstellung steht vor allem die Erstellung von "Was wäre, wenn?"-Szenarien im Vordergrund. Die Auswirkungen von Planungen sollen, so das Versprechen, durch unterschiedlichste Analyse- und Simulationsalgorithmen bereits in einem frühen Stadium berechnet, analysiert und bewertet werden. Damit sollen nicht nur Effizienzgewinne ermöglicht, sondern auch die inhaltliche Qualität der Planung im Sinne einer nachhaltigen und integrierten Stadtentwicklung erhöht werden. Gleichzeitig existiert neben unterschiedlichsten proprietären Lösungen derzeit keine offene und öffentliche Infrastruktur, die die systematische Erstellung, Bereitstellung und Nutzung von Simulationsmodellen und Algorithmen unterstützt.

Dieses Paper präsentiert drei miteinander verschränkte FOSS-Tools, die im City Science Lab der HafenCity Universität Hamburg im Rahmen des "Connected Urban Twins"-Projektes entwickelt, getestet und angewandt wurden. Dazu zählt eine Urban Model Platform zum Bereitstellen und Ausführen von dezentralen Algorithmen (https://github.com/citysciencelab/urban-model-platform), ein Scenario Explorer Tool zur Erstellung, Verwaltung und zum Teilen von urbanen Szenarien zwischen diversen Stakeholdern (https://github.com/citysciencelab/materportal-scenario-explorer) und ein Urban Model Builder zur kollaborativen Neu- und Weiterentwicklung von Simulationsmodellen basierend auf urbanen Daten (https://github.com/citysciencelab/urban-model-builder). Alle Tools sind über offene Standards wie die OGC API Processes verknüpft, bauen auf quelloffenen Projekten wie dem Masterportal oder der Open-Source Simulation Library des InsightMaker auf und ermöglichen die offene und kollaborative Erstellung von Simulationsmodellen sowie deren Nutzung in der Planungspraxis.

Kern der Tools ist die praktische Umsetzung und Überführung des Multi-modeling-Ansatzes in die Stadtentwicklung: Nur durch eine vielfältige und offene digitale Repräsentation von städtischen Prozessen durch eine Bandbreite an Expert:innen, zivilgesellschaftlichen Akteuren und interessierten Bürgern können blinde Flecken von Simulationsmodellen erkannt und abgeschwächt werden. Dafür ist sowohl eine offene Umgebung zur gemeinsamen Entwicklung, Diskussion der enthaltenen Annahmen und Bereitstellung von Simulationsmodellen notwendig, als auch eine offene Plattform zum Teilen unterschiedlichster Algorithmen. Darüber hinaus sollen Stakeholder die resultierenden Szenarien mitein-

ander teilen und organisieren können. Damit sind Grundsteine dafür gelegt, mit den Tools planerische und demokratischere Prozesse zu unterstützen.

Das Paper ist im Folgenden in fünf Abschnitte unterteilt: Zunächst wird der theoretische Hintergrund einer Demokratisierung von Algorithmen und der Multi-Modellierung im Zusammenhang mit Stadtentwicklungsprozessen und Smart-City-Entwicklungen dargelegt. Anschließend werden die drei miteinander verknüpften Tools und deren Funktionalitäten vorgestellt und diskutiert.

Hintergrund

Städte sind komplexe adaptive Systeme. Sie quantitativ zu modellieren und dadurch aussagekräftige „Was wäre wenn?"-Szenarien berechnen zu können, bedeutet zwangsläufig eine Auseinandersetzung mit deren zugrundeliegenden Charakteristika. Viele urbane Systeme, die von menschlichem Verhalten abhängen, sind hochgradig non-linear, betrachterabhängig und chaotisch. Beispielsweise hat die Corona-Pandemie einer breiten Öffentlichkeit eindrücklich vorgeführt, inwiefern die Voraussagen zu Infektionszahlen selbst Einfluss auf das Geschehen haben können. Simulationsmodelle von sozialen Prozessen sind damit nicht als bloße Abbilder der Realität, sondern als deren Mitgestalter zu begreifen [1].

Da die Simulationsmodelle eines urbanen digitalen Zwillings erheblichen Einfluss auf die Entscheidungsfindung in Stadtverwaltungen und Lokalpolitik haben können, ist deren Ausgestaltung und Integration nicht zuletzt eine politische und demokratietheoretisch relevante Fragestellung. Welche Aspekte werden abgebildet? Welche Annahmen werden getroffen? Wer entwickelt die Modelle? Wie transparent ist deren Funktionsweise? Hierbei lassen sich Bezüge zu Robert Dahls Konzept der pluralistischen Demokratie, der Habermas'schen Idee von deliberativer Demokratie und dem von Chantal Mouffe geprägten Agonismus herstellen [2].

Im wissenschaftlichen Diskurs wird seit einiger Zeit ein Ansatz proklamiert, der sich unter dem Stichwort „Multi-Modeling" subsumieren lässt. Nachdem in der Vergangenheit monolithische, hochgradig detaillierte Simulationsmodelle ihre Versprechen von präzisen Vorhersagen nur bedingt einlösen konnten, wird nun eine Diversität an Modellen zur Exploration möglicher Zukünfte in den Vordergrund gestellt [3,4]. Diese Modelle können dann – so die Annahme – im Sinne des Pluralismus diverse Weltanschauungen abbilden und zu einem besseren Umgang mit Unsicherheit und städtischer Komplexität beitragen. Gleichzeitig besteht die Möglichkeit zum „Counter-Modeling": Dem aktiven Erstellen eines alternativen Modells für das gleiche System unter anderen Annahmen [5].

Übertragen auf die Simulationsinfrastruktur eines urbanen digitalen Zwillings stellen wir im Folgenden Ansätze vor, wie die Ideen eines demokratischeren Umgangs mit Algorithmen und die der Multi-Modellierung praktisch umgesetzt und in integrierte Stadtentwicklungsprozesse eingebunden werden können. Zunächst beschreiben wir eine „Urban Model Platform" als grundlegende offene Infrastruktur zum Teilen und Verknüpfen von Modellen. Anschließend widmen wir uns der deliberativen und agonistischen Erstellung und Aushandlung von Simulationsmodellen mithilfe des „Urban Model Builder" Tools. Zuletzt beschreiben wir, wie im „Szenario Explorer" diverse Szenarien erstellt und geteilt werden können.

Urban Model Platform

Dreh- und Angelpunkt der Simulationsinfrastruktur ist eine Offene Urbane Plattform für Algorithmen und (Simulations-)modelle. Die zugrundeliegende Prämisse ist dabei, dass bestehende Prozesse nicht zwangsläufig verändert werden müssen, sondern ein „System der Systeme" das Teilen und Verknüpfen von digitalen Ressourcen durch offene Standards ermöglicht [6]. Da bereits in vielen Smart-City-Domänen bestimmte Algorithmen und Simulationsmodelle eingesetzt werden – beispielsweise Wind, Lärm, Mobilität oder Bevölkerungsprognosen – scheint es hier sinnvoll, auf den bereits existierenden Modellen aufzubauen und diese im Sinne eines modularen UDZ für bestimmte Anwendungsfälle miteinander zu kombinieren und Schritt für Schritt zu erweitern [7].

Die Urban Model Platform (UMP) wurde speziell für diesen Zweck in Zusammenarbeit des City Science Labs und des Landesbetriebs für Geoinformation und Vermessung entwickelt und ist seit Beginn 2025 im produktiven Betrieb der Freien und Hansestadt Hamburg. Die Idee dahinter ist, dass eine Vielzahl an unterschiedlichsten Servern an die UMP angebunden werden können, die Algorithmen und Simulationsmodelle mittels verschiedenster organisationsabhängiger Technologie-Stacks berechnen. Nutzer:innen kann so Zugriff auf eine Reihe an Modellen gewährt werden, die aktuell in verschiedenen Silos verwendet werden. Sämtliche Algorithmen werden über den offenen OGC-API-Processes-Standard bereitgestellt, sodass eine nahtlose Integration in diverse Frontends und Clients ermöglicht wird.

Technisch besteht die UMP aus vier Komponenten:

- Eine **Flask API**, die die Endpunkte für Metadaten und die Ausführung von verschiedenen Modellen konform zur OGC API Processes bereitstellt
- Eine **PostGIS**-Datenbank, die ausgeführte Szenarien, Ergebnisse und Metadaten speichert
- Eine **GeoServer**-Instanz, die die resultierenden Geodaten von entsprechend konfigurierten Modellen über gängige offene Standards bereitstellt
- Eine **Keycloak**-Instanz, die Authentifizierung und Autorisierung von Usern zur Ausführung bestimmter Modelle oder zur Abfrage deren Ergebnisse ermöglicht

Durch das Vorhandensein einer UMP können Simulationsmodelle für bestimmte Personen(-gruppen) freigegeben werden und agnostisch vom entsprechenden Frontend ausgeführt werden. Strukturell wird so eine modulare und offene Umgebung für verschiedenste Modelle geschaffen, die beliebig erweiterbar und ersetzbar ist – ein Gegenmodell zu einem proprietären, in sich abgeschlossenen digitalen Zwilling.

Urban Model Builder

Viele städtische Prozesse – insbesondere solche, die soziale Dynamiken abbilden – können zwar quantitativ modelliert werden, sind jedoch lokalspezifisch und besitzen nicht dieselbe Allgemeingültigkeit wie bspw. physikalische Gleichungen der Strömungsdynamik. Oftmals muss mit Annahmen operiert werden, die einen großen Einfluss auf die Resultate der Modellierung haben. Um diese Annahmen transparent zu machen und zwischen Stakeholdern zu verhandeln, kann auf Co-Modellierungsansätze zurückgegriffen werden. Dabei werden konzeptionelle Zusammenhänge und Prozesse in Workshops diskutiert und deren entsprechende digitale Abbilder mittels geeigneter Simulationssoftware erschaffen [4]. Bereits existente Tools hierfür sind überwiegend proprietär, zumeist nicht direkt kollaborativ nutzbar und selten web-basiert.

Der Urban Model Builder ist ein FOSS-Tool, welches diese Lücke füllt und den Co-Modellierungsprozess in Form einer browserbasierten Webapp bestmöglich unterstützen soll. Nutzer:innen können sich registrieren und gelangen nach dem erfolgreichen Login zu einer Übersichtsseite, auf der alle von ihnen erstellen und mit ihnen geteilten Modelle sichtbar sind. Jedes Modell selbst kann je nach Belieben auf einem unendlich großen Canvas bearbeitet werden (s. Abbildung 1). Dafür stehen verschiedene Primitives zur Verfügung, die die Systemmodellierung nach dem System-Dynamics-Paradigma, der agentenbasierten Modellierung oder einer Kombination ermöglichen. Im Hintergrund werden die Elemente in einer Postgres-Datenbank gespeichert und mit einer offenen JavaScript-Simulationsbibliothek verknüpft [8]. Erwähnenswert ist auch die Möglichkeit, über die OGC-API-Features-Schnittstelle Daten direkt von urbanen Datenplattformen abzufragen und in das Simulationsmodell zu integrieren. Damit ist die direkte Verbindung von Datensätzen und Simulationsmodellen für „Was wäre wenn?“-Szenarien geschaffen.

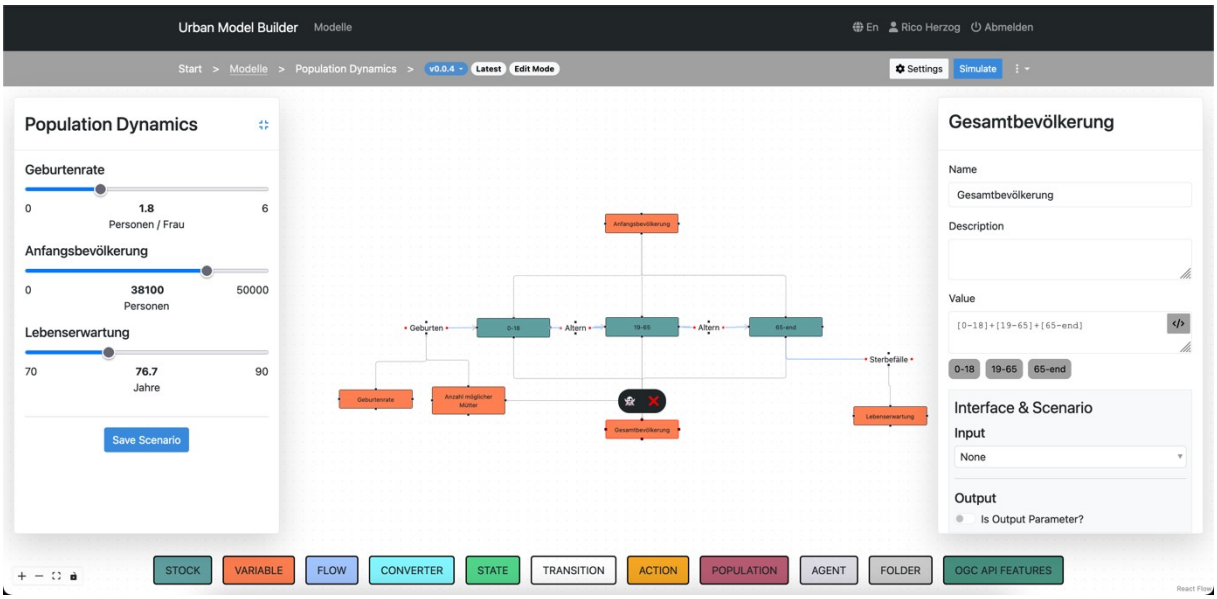


Abbildung 1: Mit dem Urban Model Builder können generische Modelle mit dem System-Dynamics- und Agent-based-Modelling-Paradigma kollaborativ erstellt werden

Ferner besteht die Möglichkeit, Variablen und Inputs des Modells als Parameter zu deklarieren, die direkt im Interface einstellbar sind und somit von Nutzer:innen angepasst werden können. Mit einem Klick auf „Simulieren“ werden im Browser zuvor definierte Outputs über einen definierten Simulationszeitraum berechnet und entweder in Form eines Graphen, eines Scatter-Plots oder einer Karte (für einzelne Agenten) visualisiert (s. Abbildung 2).

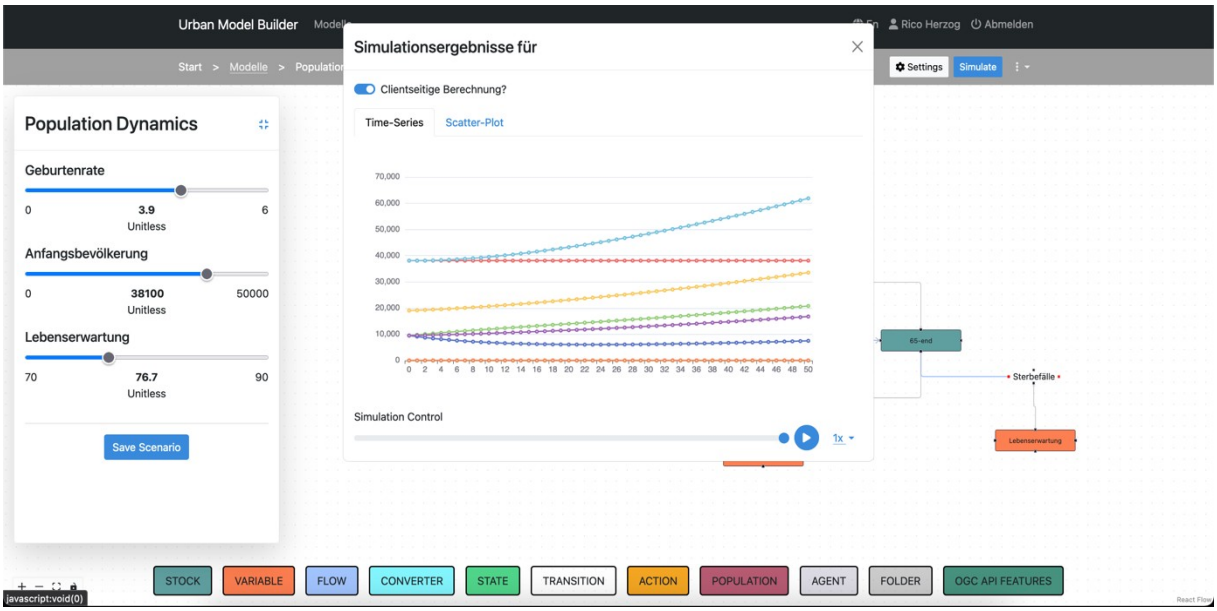


Abbildung 2: Variierende Parametereinstellungen produzieren unterschiedliche Outputs, die direkt angezeigt werden können.

Für eine kollaborative Entwicklung können Modelle mit anderen Usern geteilt werden und durch WebSockets in Verbindung mit einer Autosave-Funktion in Echtzeit gemeinsam bearbeitet werden. Verschiedene Rollen und Rechte unterstützen diverse Möglichkeiten der Kollaboration wie bspw. das alleinige Anzeigen, die Mitbearbeitung oder das weitere Teilen und Veröffentlichen. Die Veröffentlichung eines Modells ermöglicht auch nicht-registrierten Nutzern das Anzeigen und Ausführen von Modellen im Browser und bietet gleichzeitig die Möglichkeit, das Modell über eine OGC-API-Processes-Schnittstelle bereitzustellen, sodass eine Anbindung an die UMP stattfinden kann. Mit der Veröffentlichung und Bereitstellung eines Modells geht eine gleichzeitige Versionierung einher, die für Persistenz und Reproduzierbarkeit der Ergebnisse erforderlich ist. Nutzer:innen können unterschiedliche Entwurfsversionen anlegen sowie beliebig viele „Major"- und „Minor"-Versionen ihres Modells veröffentlichen.

In Stadtentwicklungsprozessen kann der Urban Model Builder damit in diversen Anwendungsfällen eingesetzt werden. Interessengruppen können gemeinsam digitale Abbildungen von unterschiedlichsten urbanen Systemen wie bspw. Energie, Verkehr oder Bevölkerungsdynamiken co-modellieren, welche durch die Anbindung an die UMP geteilt, verknüpft und in unterschiedlichsten Anwendungen integriert werden können. Gleichzeitig besteht die Möglichkeit und Offenheit zum Klonen und „Counter Modeling" bereits veröffentlichter Simulationsalgorithmen.

Scenario Explorer

Neben dem Modellierungsprozess ist vor allem die Nutzung unterschiedlichster Simulationsmodelle in einem urbanen digitalen Zwilling für eine evidenzbasierte Entscheidungsfindung relevant. In der Praxis werden hierzu aktuell eine Vielzahl an Gutachten erstellt, die dann ihren Weg über etliche Seiten statischer PDF-Dokumente in einen sektoral fragmentierten und oftmals undurchsichtigen Entscheidungsfindungsprozess finden. Eine Integration der Planungsprognosen und Ergebnisdaten oder eine effiziente (Teil-)Automatisierung bestimmter Prüfungen findet ausgesprochen selten statt.

Um die auf der UMP verfügbaren Modelle nutzen zu können und eine möglichst breite Zielgruppe anzusprechen, entwickelten wir den Scenario Explorer als Addons zum Masterportal 3. Er dient als Client für unterschiedlichste Simulationsalgorithmen, die dynamisch ausgeführt und deren Ergebnisse angezeigt und untereinander geteilt werden können (s. Abbildung 3).

Startet man das Addon, erscheint eine Seitenleiste, die verschiedene Funktionalitäten bedient:

1. **Startseite** mit zuletzt verwendeten Simulationsmodellen, Szenarien sowie einer Hilfesektion
2. **Modell-Übersicht** auf der alle Simulationsmodelle angezeigt werden, für die der Nutzer autorisiert wurde. Nicht authentifizierte Nutzer:innen des Addons sehen nur solche Modelle, die entsprechend offen konfiguriert sind. Nutzer:innen können von hier aus Simulationen starten.

3. **Szenario-Übersicht** auf der alle Szenarienergebnisse aufgeführt werden, die ein Nutzer ausgeführt hat oder die mit einem Nutzer geteilt wurden. Von hier gelangt man zur Ergebnisanzeige und den Details des jeweiligen Szenarios
4. **Ensemble-Übersicht** auf der alle berechneten Ensembles gelistet werden. Ensembles sind eine Bandbreite an Szenarien von einem oder mehreren Modellen, die mit unterschiedlichen Eingangsbedingungen automatisiert gestartet wurden
5. **Hilfeseite** mit FAQs und Tutorials zur Nutzung des Addons

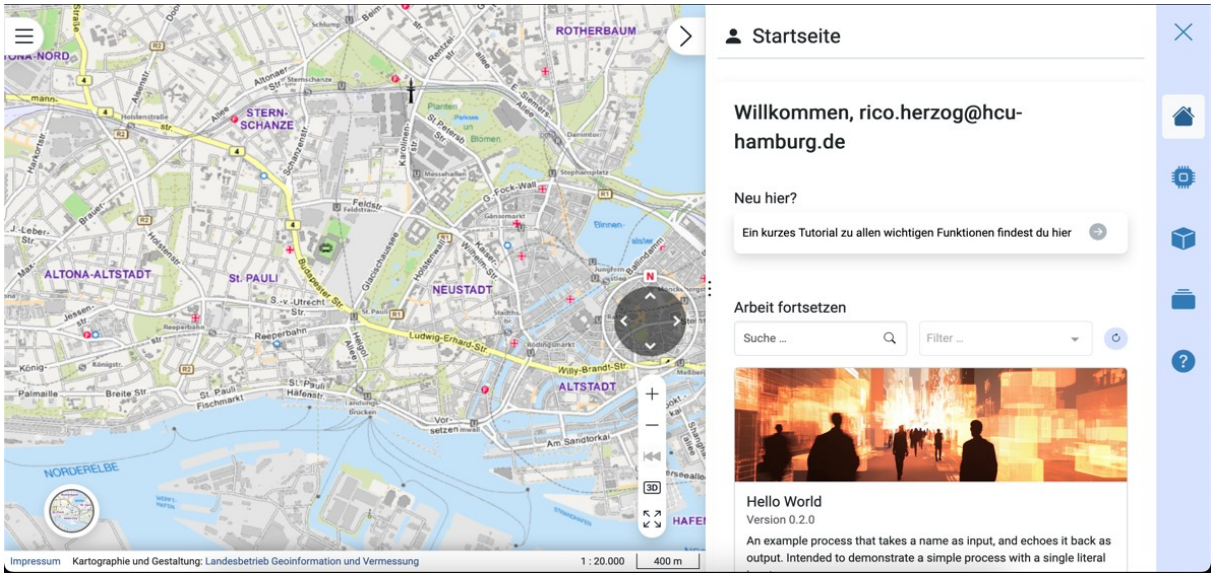


Abbildung 3: Der Szenario Explorer ist ein Masterportal 3-Addon, welches das Ausführen von Simulationsmodellen, Anzeigen und Teilen von Ergebnissen sowie die Ensemble-Modellierung unterstützt

Ein großer Vorteil des Scenario Explorers ist das dynamische Anzeigen und Ausführen der Simulationsmodelle. Unterstützt durch den offenen Standard der OGC API Processes können so mit wenig Konfigurationsaufwand neue Modelle über die UMP bereitgestellt werden und mit dem Scenario Explorer ausgeführt werden. Ergebnissedaten können anschließend dargestellt (s. Abbildung 4), in Ensembles gruppiert und mit anderen Nutzer:innen geteilt werden, die möglicherweise keinen Zugriff auf das zugrundeliegende Modell haben. Gleichzeitig können andere Funktionen und Addons des Masterportals genutzt werden, um die Ergebnisdatenlayer zu analysieren.

Diskussion und Fazit

Urbane digitale Zwillinge werden von einer zunehmenden Anzahl an Städten entwickelt, implementiert und erprobt. Dabei ist davon auszugehen, dass mit einer wachsenden Datenbasis und -verfügbarkeit auch immer mehr Algorithmen und Simulationsmodelle für die Erstellung von „Was wäre, wenn?“-Szenarien unterschiedlichster urbaner Prozesse eingesetzt werden. Die Ausgestaltung dieses Prozesses wirft gleichzeitig Fragen zu Repräsentation im Digitalen, Teilhabe, Machtstrukturen und neu geschaffenen Pfadabhängigkeiten auf.

Basierend auf dem Ansatz des Multi-Modellings [3] und digitaler Diversität im Abbilden von städtischen Prozessen soll mit einer offenen Infrastruktur zum Erstellen, Teilen und Nutzen von Simulationsmodellen ein Gegenpol zu proprietären Technologien geschaffen werden. Eine Urban Model Platform als

grundlegende und offene Infrastruktur kann analog zu einer Urban Data Platform einen transparenten Zugangspunkt zu Algorithmen schaffen, der es einer Bandbreite an Akteuren erlaubt, ihre (Simulations-)Modelle zur Verfügung zu stellen. Damit wird die technische Grundlage für eine pluralistische Repräsentation von städtischen Systemen geschaffen und Lock-In-Effekte werden durch offene Standards wir die OGC API Processes reduziert. Gleichzeitig können unterschiedlichste Modelle und Paradigmen verwendet werden, um ein urbanes System abzubilden. Beispielsweise können sowohl mit einem proprietären KI-Modell, als auch mit einem städtischen Agent-based-Model Verkehrsmengen vorhergesagt werden.

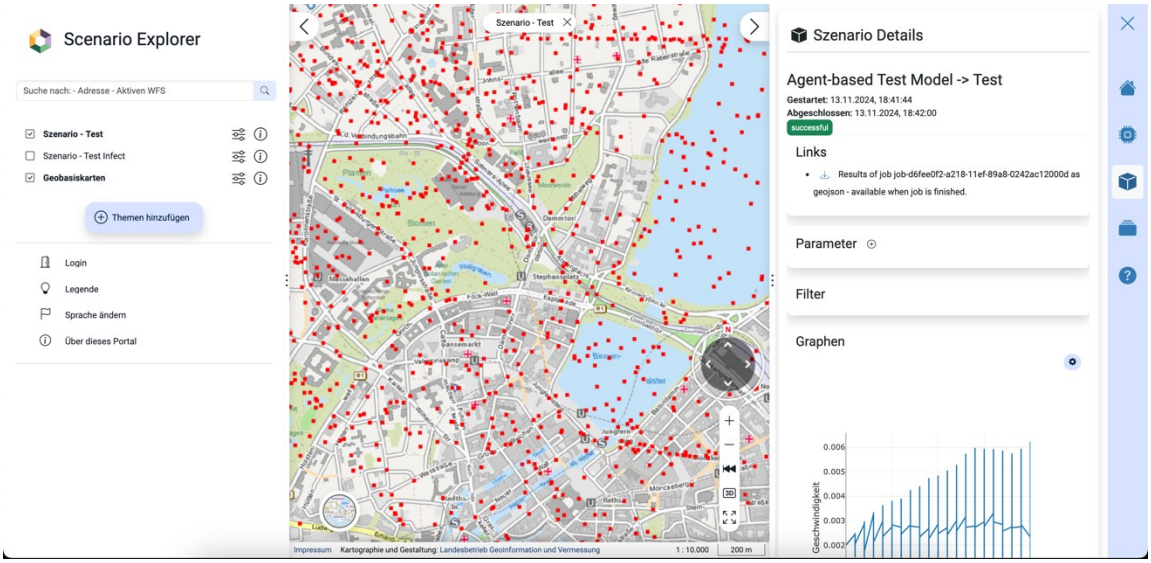


Abbildung 4: Die Szenario-Detailseite zeigt resultierende Geodaten an und visualisiert diese auch als konfigurierbaren Graph. Nutzer:innen können ihre Szenarien-Ergebnisse hierüber für andere Nutzer:innen freigeben.

Um deliberative Prozesse zu unterstützen, in denen unterschiedliche Stakeholder zu einem gemeinsamen Verständnis kommen oder durch einen agonistischen Wettkampf der Ideen verschiedene normative Vorstellungen verhandeln, kann der Urban Model Builder eine wichtige Rolle im Co- und Counter-Modelling einnehmen. Durch eine Full-Stack FOSS-Infrastruktur zum kollaborativen Erstellen, Teilen und Versionieren von System Dynamics und Agent-based Modellen ist die Möglichkeit geschaffen, Transparenz über die Struktur und Annahmen der Modelle herzustellen und Gegenentwürfe zu verhandeln.

Nicht zuletzt ist durch den Scenario Explorer die Möglichkeit geschaffen, in einem breit genutzten und verfügbaren System wie dem Masterportal unterschiedlichste Szenarien zu erstellen, Ergebnisse zu analysieren und domänenübergreifend zu teilen.

Kontakt zum Autor:

Rico Herzog
City Science Lab, HafenCity Universität Hamburg
Henning-Voscherau-Platz 1
20457 Hamburg
rico.herzog@hcu-hamburg.de

Literatur und Quellen

[1] Herzog, R., & Probst, V. (2024). Guide to Model Land: A guide to ethical questions for modeling and simulation in urban digital twins. HafenCity University Hamburg. https://doi.org/10.34712/142.54

[2] Sætra, H. S., Borgebund, H., & Coeckelbergh, M. (2022). Avoid diluting democracy by algorithms. Nature Machine Intelligence, 4(10), 804–806. https://doi.org/10.1038/s42256-022-00537-w

[3] Batty, M. (2021). Multiple models. Environment and Planning B: Urban Analytics and City Science, 48(8), 2129–2132. https://doi.org/10.1177/23998083211051139

[4] Herzog, Rico. (2023). Exploring multi-modelling approaches in Hamburg, Germany's evolving digital urban twin infrastructure. 22nd International Conference on Modelling and Applied Simulation (MAS), Athen. https://doi.org/10.46354/i3m.2023.mas.001

[5] Greenberger, M., Crenson, M. A., & Crissey, B. L. (1976). Models in the policy process: Public decision making in the computer era. Russel Sage Foundation.

[6] DIN SPEC 91357:2017-12, Referenzarchitekturmodell Offene Urbane Plattform_(OUP); Text Englisch. (2017). Beuth Verlag GmbH. https://doi.org/10.31030/2780217

[7] Schubbe, N., Bodecker, M., & Moshrefzadeh, M. (2023). Urbane Digitale Zwillinge als Baukastensystem: Ein Konzept aus dem Projekt Connected Urban Twins (CUT). zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, 1/2023, 14–23. https://doi.org/10.12902/zfv-0417-2022

[8] Fortmann-Roe, S. (2014). Insight Maker: A general-purpose tool for web-based modeling & simulation. Simulation Modelling Practice and Theory, 47, 28–45. https://doi.org/10.1016/j.simpat.2014.03.013

Migration eines Auskunftssystems zu einer Open-Source Lösung mit QGIS

Victor Ali Lagoa

Die Geodatenabteilung der Stadtwerke München nutzte GE Smallworld für ihr Netzinformationssystem, was aufgrund steigender Komplexität und Nutzerzahl zu hohen Kosten führte. Vor fünf Jahren begann die Migration zu NIS-QGIS, einer kostensparenden, auf QGIS-Desktop und PyQGIS-Plugins basierenden Open-Source-Lösung. Der Vortrag liefert einen Überblick über die neue Anwendung, Anwendungsfälle, Herausforderungen und Erkenntnisse.

Seit vielen Jahren bot die Geodatenabteilung der Stadtwerke München GmbH ihren Nutzern in anderen Bereichen des Unternehmens (Wasser, Gas, Strom, etc.) ein Netzinformationssystem basierend auf GE Smallworld an. Diese Anwendung wurde in Zusammenarbeit mit einem Dienstleister unterstützt und betrieben, was erheblich zu den Gesamtkosten des Systems beitrug. Die unvermeidliche Zunahme der Komplexität, proportional zur ständig wachsenden Anzahl von Nutzern und Use-Cases, verbunden mit der konkurrierenden Natur der Interessen der verschiedenen Stakeholder, führte zu vielen Komplikationen, Ineffizienzen und letztendlich zu hohen Kosten.

Vor etwa fünf Jahren beschloss das Geodaten-Team, das System schrittweise zu modernisieren und auf eine Open-Source-Lösung zu migrieren. Eines der neuen Werkzeuge ist NIS-QGIS, eine Anwendung basierend auf QGIS-Desktop, PyQGIS-Plugins und einigen Back-End-Diensten, deren Ziel es war, die gesamte Funktionalität der alten Anwendung bereitzustellen sowie ein Werkzeug bereitzustellen, das einfacher und schneller aktualisiert und weiterentwickelt werden konnte und dabei Kosten senkt und „Time-to-Market“ erhöht. In diesem Vortrag geben wir einen Überblick über diese Anwendung, einige ihrer Hauptanwendungsfälle und Komponenten, Herausforderungen (sowohl gelöst als auch ungelöst) und die dabei gewonnenen Erkenntnisse.

Künstliche Intelligenz als Unterstützung in geografische Applikationen

Andrea Borghi, Marion Baumgartner

In dieser Präsentation zeigen wir die Ergebnisse von zwei POCs, welche wir mit dem Large Language Models (LLM) Open Source Python Framework “LangChain” durchgeführt haben. Mittels LLM (wie GPT) zeigen wir, wie eine geographische Anwendung durch künstliche Intelligenz benutzerfreundlicher und zugänglicher wird. Auch im Hintergrund arbeiten wir mit künstlicher Intelligenz, um die Anreicherung und die Indexierung von Metadaten zu verbessern und zu automatisieren.

Wir haben uns die Frage gestellt, wie man das Potenzial von Künstlicher Intelligenz in Form von Sprachmodellen wie GPT in der Geoinformatik nutzen kann, um geografische Anwendungen benutzerfreundlicher und auch für Laien zugänglich zu machen. Denn ist man nicht vertraut mit den geographischen Fachbegriffen oder der Applikation, treten oft Fragen auf wie:

Welchen Layer muss ich aktivieren, um die Daten auf der Karte zu sehen, die mich interessieren? Wo versteckt sich die Funktion, um etwas auf der Karte zu zeichnen?

Daraus ergaben sich drei Bereiche, die wir in diesem Vortrag erläutern möchten. Anhand von zwei verschiedenen Proof of Concepts (POC), welche wir mit zwei Kunden in diesen Bereichen umsetzen, werden funktionierende Beispiele gezeigt.

Zum einen wird der Einsatz von Large Language Models (LLM) für die Suche in geographischen Anwendungen gezeigt, zum anderen die Steuerung einer Anwendung über Text bzw. Sprache, das sogenannte Conversational User Interface (CUI). Eine dritte Anwendung ist die Verwendung von Sprachmodellen zur Unterstützung der Metadatengenerierung.

Den letzten Teil dieses Vortrags werden wir der Frage widmen: "Wie testet man eine Anwendung, die LLMs verwendet?" Eine Frage, die uns bei der Durchführung der oben genannten POCs immer wieder beschäftigt. Denn wie stellt man fest, ob sich der Aufwand lohnt und die Suche wirklich ein besseres Ergebnis liefert? Oder ob die erzeugten Metadaten brauchbar sind.

Metadaten für eine verantwortungsvolle und kritische Geodatenpraxis

Ester Scheck[1]

[1] Forschungsbereich Kartographie, TU Wien

DOI: 10.5281/zenodo.14616221

Zusammenfassung: Die Arbeit zeigt die Bedeutung des Entstehungskontextes für Datensätze auf und untersucht, welche Informationen für eine reflektierte und verantwortungsvolle Nutzung von Geodaten relevant sind. Dazu werden Dokumentations- und Reflexionsframeworks aus den Critical Data Studies sowie der Forschung zu Ethical AI und Maschinellem Lernen untersucht und in Bezug zum ISO-Standard für Metadaten geographischer Informationen gesetzt, da Kontextinformationen als Metadaten zur Verfügung gestellt werden können. Die Analyse offenbart zahlreiche Überschneidungen, aber auch inhaltliche Lücken und signifikante Unterschiede im Format der Frameworks und des Standards. In einem Anwendungsfall werden die Metadaten von drei offenen Datensätzen beispielhaft analysiert, um die Verfügbarkeit relevanter Kontextinformationen zu evaluieren. Die Ergebnisse zeigen, dass wesentliche Aspekte des Entstehungskontexts fehlen und dies den reflektierten Umgang mit Daten erschwert. Abschließend werden Potenziale und Bedarfe für die Integration und Weiterentwicklung von Kontextinformationen sowie deren Nutzbarmachung in Karten und GIS-Anwendungen skizziert.

Schlüsselwörter: Open Data, Metadaten, Critical Data Studies, Datenkontext, OSM

Die Bedeutung von Kontext im Umgang mit (Geo)-Daten

Zahlen sprechen nicht für sich selbst, stellen Catherine D'Ignazio und Lauren F. Klein in ihrem Buch *Data Feminism* fest [1] und berufen sich dabei u.a. auf Donna Haraways Theorie des situierten Wissens [2]. Als eins der sieben Prinzipien von *Data Feminism* fordern sie deshalb, in der Arbeit mit Daten den Kontext der Daten einzubeziehen. Dies lässt sich insbesondere auf die Arbeit mit offenen (Geo)-Daten beziehen. Dank der Open Data-Bewegung ist es mittlerweile möglich, zu verschiedensten Themen einen offenen Datensatz im Internet zu finden. Für datenbereitstellende Institutionen und Personen liegt der Fokus häufig auf dem Offenlegen von möglichst vielen Daten; für eine umfangreiche Dokumentation fehlen oft die Ressourcen [1, S. 155]. Offene Daten ermöglichen die Nutzung durch Personen mit diversen Hintergründen und Interessen, gleichzeitig hat sich durch die freie Verfügbarkeit die Distanz zwischen Datenproduzent*innen und Datennutzer*innen vergrößert: Der Austausch zwischen den beiden Parteien wird durch Verfügbarkeit von Datenportalen stark reduziert und offene Datensätze werden überwiegend von Personen genutzt, die die Daten nicht selbst gesammelt oder aufbereitet haben [3, S. 10; 4, S. 1]. Auch domänenspezifisches Wissen, wie im Kontext von Geodaten die dahinterliegenden Konzepte räumlicher Referenzierung, kann bei Nutzer*innen nicht vorausgesetzt werden [5, S. 2; 4, S. 1]. Obwohl also mehr und mehr Daten verfügbar gemacht werden, fehlt zunehmend der Kontext dieser Daten, der für eine „genaue, ethische Analyse" [1, S. 149, eigene Übersetzung] nötig ist. Dieser Kontext umfasst beispielsweise die Erhebungsmethoden, Definitionen, die hinter angewandten Klassifikationen stehen, die Ziele der Datenerhebung, wer die Daten erhoben hat, sowie Limitationen der Daten [6, S. 10].

Nur durch eine Betrachtung der Entstehung der Daten, der sozialen, kulturellen, historischen, institutionellen und materiellen Begebenheiten und den damit zusammenhängenden Machtverhältnissen, lassen sich Grenzen der Daten, mögliche Bias und eine daraus resultierende ethische Verantwortung in der Nutzung erkennen und verstehen [1, S. 152]. Fehlende Kontextinformationen können zu irreführenden Interpretationen und Analysen führen, insbesondere bei der Kombination heterogener Daten-

sätze oder Langzeitanalysen [7, S. 1513; 8]. Heather Krause verdeutlicht dies an einem Beispiel in dem das Tabellenformat eines Datensatzes fälschlicherweise Vergleichbarkeit suggeriere, obwohl die Daten aus sehr unterschiedlichsten Quellen und Kontexten stammten [9]. Ohne Kontextinformationen wird oft auch der Detailgrad von Daten überschätzt, etwa wenn Verarbeitungsprozesse wie Aggregationen nicht transparent gemacht werden [7, S. 1513; 8]. Datennutzer*innen könnten zudem annehmen, die Daten vollständig zu verstehen, obwohl abweichende Definitionen oder Konzepte – etwa beim Begriff ‚Wald' – genutzt wurden [3, S. 6]. Da Daten in unserer Gesellschaft zunehmend Grundlage politischer und wirtschaftlicher Entscheidungen sind, können Fehlinterpretationen im weiteren Verlauf zu nicht bedarfsgerechten Entscheidungen und Ressourcenverteilungen führen.

Im Bereich von Geodaten machen spezifische Herausforderungen Kontextinformationen besonders relevant. Geodätische Messungen sind stets mit Unsicherheiten behaftet, und Erfassungs- sowie Klassifizierungsmethoden, etwa bei der Bodenbedeckung, können variieren [3, S. 1–2]. Jede Repräsentation unserer Umwelt ist eine vereinfachte Abstraktion, die durch Modellierungsentscheidungen Informationsverluste und Unsicherheiten verursacht [3, S. 2]. Um mit der Variabilität zwischen gleichwertigen Darstellungen derselben realen Objekte und den Unstimmigkeiten in räumlichen Daten konstruktiv umzugehen, sind Kontextinformationen unverzichtbar [3, S. 2].

Während in der Forschung zu Maschinellem Lernen (ML) und Künstlicher Intelligenz (KI/AI) längst bekannt ist, dass verzerrte Trainingsdaten zu problematischen Ergebnissen führen können [10, S. 2; 11], ist das praktische Bewusstsein für die Bedeutung von Kontextinformationen in der Erstellung von Karten, GIS-Anwendungen und der Analyse von Geoinformationen noch wenig ausgeprägt. Jede Geovisualisierung basiert auf Daten - ob quantitativ oder qualitativ, heruntergeladen oder selbst erhoben. Neben den Karten, die auf Entscheidungen wie Projektion, visuelle Variablen, Generalisierungsgrad und Datenklassifikation beruhen, sind auch die verwendeten Ausgangsdaten als Ergebnis eines komplexen, teils bewussten, teils unbewussten Entscheidungsprozesses zu verstehen und zu reflektieren. Das Modell der fundamentalen Informationstransformationen in der Kartographie von Robinson et al. [12, S. 18] beschreibt die Überführung von der geographischen Umwelt zu geographischen Informationen als erste von drei Transformationen im Kartierungsprozess. Jede dieser Transformationen hat laut Robinson et al. das Potenzial, das finale Produkt, die Karte, zu beeinflussen. Daher ist eine bewusste Auseinandersetzung mit dem Einfluss dieser Transformation und des Kontexts von Daten unverzichtbar.

In meiner Forschung untersuche ich deshalb die Relevanz von Kontextinformationen speziell für Geodaten und ihren Bezug zu Metadaten. Zunächst analysiere ich, welche Kontextinformationen in der Literatur häufig als relevant betrachtet werden. Als Grundlage dienen Dokumentations- bzw. Reflexionsframeworks aus den Critical Data Studies und der *Ethical AI*- und ML-Forschung. Da Kontextinformationen als Metadaten zur Verfügung gestellt werden können, betrachte ich die Überschneidung zu Metadatenstandards und vergleiche die Frameworks mit dem ISO-Metadatenstandard für Geodaten. Anschließend untersuche ich in einem kleinen Use Case drei Geodatensätze hinsichtlich der in den Metadaten bereitgestellten Informationen und ihrer Abdeckung der vorab identifizierten Kontextinformationselemente. Meine Arbeit überträgt damit einen zentralen Ansatz der Critical Data Studies auf die Kartographie und GIS und liefert Ansatzpunkte für die Weiterentwicklung einer verantwortungsvollen Geodatenpraxis.

Dokumentations- und Reflexionsframeworks für Kontextinformationen

Kontext, von Duden als ‚Zusammenhang' definiert [13], ist ein breit gefasstes und schwer abzugrenzendes Konzept. Welche Informationen in Bezug auf Datensätzen konkret dazu gehören ist umfassend und abschließend kaum zu definieren. Einen Ausgangspunkt können jedoch verschiedene Dokumentations- und Reflexionsframeworks für Datensätze bieten, die in den Critical Data Studies oder der Forschung zu *Ethical AI* und ML entwickelt wurden. Als Framework verstehe ich in diesem Zusammenhang einen strukturierten Leitfaden, um bestimmte Prozesse, Fragestellungen oder The-

men zu analysieren und zu dokumentieren. Beispiele dafür sind die von D'Ignazio und Klein [1, S. 168–169] erwähnten Ansätze *Data Biography*, *Datasheets for Datasets* und *Data User Guides*. Ergänzend wurden weitere Frameworks aus der Literatur identifiziert und in Tabelle 1 zusammengefasst. Nach einem konzeptionellen Vergleich dieser Frameworks untersuche ich inhaltliche Überschneidungen, wobei Aspekte, die nur in wenigen der Frameworks vorkommen, aus Platzgründen vernachlässigt werden.

Tabelle 1: Übersicht der Reflexions- und Dokumentationsframeworks, die verglichen werden

Titel	Autor*in oder Herausgeber*in	Art	Kommentar
Data Biography	Krause, Heather (2017, 2023) [9; 14]	Reflexionsframework	
Data User Guides	Western Pennsylvania Regional Data Center (o.J.) [15]	Dokumentationsframework	Template von 2015, aus Bedarf von datenveröffentlichender Stelle heraus entwickelt
CIVIC Data Library of Context	CIVIC Software Foundation (o.J.) [16]	Dokumentationsframework	2020 in einem Workshop entwickelt, nur Beispielmodell und wenige Hintergrundinformationen verfügbar
Datasheets for Datasets	Gebru et al. (2018), Microsoft Research [17]	Dokumentationsframework	Entwickelt für KI- und ML-Model-Datensätze
Aether Data Documentation	Microsoft's Aether Transparency Working Group (2022) [18]	Dokumentationsframework	Weiterentwicklung von Datasheets for Datasets
Data Nutrition Project	Holland et al. (2020), Chmielinski et al. (2020)[5]	Dokumentationsframework	Teilweise automatisiert, entwickelt für KI-Model-Datensätze
Data Cards	Pushkarna et al. (2022), Google Research [19]	Dokumentationsframework	Teilweise automatisiert, entwickelt für KI-Model-Datensätze

Das Konzept der Datenbiographie, entwickelt von der Datenwissenschaftlerin Heather Krause, richtet sich als einziges der untersuchten Frameworks explizit an Datennutzer*innen. Krause fordert, einen Datensatz genauso kritisch zu betrachten wie jede andere Informationsquelle. Zur Reflexion liefert sie fünf Leitfragen und Unterfragen, die dabei helfen sollen, Lücken, Verzerrungen, Falschinformationen und übersehene Details des Datensatzes aufzudecken [9]. Die anderen untersuchten Frameworks richten sich eher an Datenbereitsteller*innen zur Dokumentation eines Datensatzes und werden in dieser Arbeit deshalb als Dokumentationsframeworks kategorisiert. *Datasheets for Datasets* stellt jedoch ebenfalls eine Reflexion von Datenproduzent*innen im gesamten Datenerstellungs- und Datenmanagementprozess in den Vordergrund [17, S. 2]. Die Zielgruppen der Frameworks sind nicht immer explizit genannt, lassen sich jedoch zum Teil aus Formulierungen und Anwendungsbeispielen ableiten.

Vier der untersuchten Frameworks stammen aus der wissenschaftlichen Forschung, drei davon aus Unternehmen (Microsoft und Google), und wurden im Kontext von *Ethical AI*- und ML-Forschung entwickelt, um eine kritische Auseinandersetzung mit Trainingsdaten zu fördern. Der Großteil der dort behandelten Elemente lässt sich auch auf (offene) Datensätze allgemein übertragen. Speziell auf Geodaten ausgerichtet ist jedoch keines der Frameworks.

Abgesehen vom *Data User Guide*, welcher nur aus acht Aspekten mit einzelnen Erklärungen oder Fragen besteht, sind alle untersuchten Frameworks in sechs bis neun Sektionen unterteilt, die insgesamt bis zu 60 Fragen umfassen. Die *Data Biography* orientiert sich beispielsweise an den fünf Leitfragen:

wer, was, wo, wie und wozu [20]. Die Abschnitte in *Datasheets for Datasets* sind an den Lebenszyklus eines Datensatzes angelehnt und umfassen die Motivation, Zusammenstellung, Erhebungsprozess, Vorverarbeitung/Datenreinigung/ Labeling, Anwendungen, Distribution und Pflege [17]. In den Frameworks sind die Fragen teils sehr offen gestellt, teils geben konkrete Beispiele oder Single- bzw. Multiple-Choice-Antworten eine klare Richtung vor. Auch die Formate der Framework-Vorlagen variieren erheblich: von einem einfachen Textdokument [21; 22] über eine Google-Sheets-Vorlage [20], eine statische Website [23] und ein PDF-Dokument [24; 25; 21] bis hin zu einer webbasierten, interaktiven Nutzer*innen-Oberfläche [26]. Auch die Bereitstellung von Begleitmaterialien, Anwendungsbeispielen und Hintergrundinformationen unterscheidet sich zwischen den verschiedenen Frameworks stark.

Das *Dataset Nutrition Label* und die *Data Cards* wenden als einzige untersuchte Frameworks zum Teil automatisierte Prozesse an, etwa für Diagramme, Tabellen und einfache statistische Analysen, die in die Dokumentation integriert werden [10; 19]. Gebru et al. hingegen betonen, dass eine automatisierte Erstellung der Dokumentation dem Reflexionsziel entgegenwirken würde [17, S. 3].

Die Dokumentationsframeworks zielen auf die Bereitstellung von Informationen zu einem Datensatz ab. Nach einer breiten Definition von Metadaten als ‚Daten über Daten' [3, S. 6] handelt es sich somit um Metadaten. Dennoch wird in den Frameworks nur wenig bis gar kein Bezug zu Metadaten hergestellt. Spärliche Verweise umfassen die Definition des *Data Nutrition Labels* als ein „Standard Format für Metadaten-Kommunikation" [10, S. 8, eigene Übersetzung]. Die *Data Cards* und *Data User Guides* verstehen sich als ein über Metadaten hinausgehendes Konzept [19, S. 1777; 15] und das Modell der *CIVIC Data Library of Context* [23] soll ein Beispiel für kontextuelle Metadaten sein. Den bisher begrenzten Bezug der Frameworks zu Metadaten stelle ich dieser Arbeit im nachfolgenden Kapitel her.

Im inhaltlichen Vergleich der Frameworks sind heterogene Begriffsverwendungen und lückenhafte Erklärungen oder Beispiele eine Herausforderung. Es ist davon auszugehen, dass teilweise unterschiedliche Dinge mit der gleichen Bezeichnung gemeint sind oder unterschiedliche Bezeichnungen für denselben Aspekt verwendet werden (z.B. *data acquisition* und *data collection*, *data creator* und *data owner*, *lineage* und *provenance*, *dataset type*, *data type* und *type of instances*). Dies muss bei der Betrachtung der folgenden Erkenntnisse zu inhaltlichen Überschneidungen der Frameworks berücksichtigt werden.

Neben einer kurzen Beschreibung des Datensatzes wird in den sieben untersuchten Frameworks besonders häufig nach der Motivation und dem Zweck der Datenerhebung gefragt. Auch die Finanzierung ist ein mehrfach genannter Aspekt. Typische Metadaten wie der Datensatzname und das Veröffentlichungsdatum werden ebenfalls häufig abgefragt. Im Vergleich wurde sichtbar, wie viele verschiedene temporale Informationen neben dem Veröffentlichungsdatum in Bezug auf einen Datensatz erfasst werden können: Zeitraum der Datenerhebung, Zeitpunkt der Datenerstellung, erfasster Zeitraum oder auch das Aktualisierungsdatum. Auch in Bezug auf das ‚Wer' werden in den Frameworks unterschiedliche Rollen erfasst, die unter Umständen nicht immer klar voneinander getrennt werden können. Mindestens eine Rolle davon wird in den meisten Frameworks abgefragt: Datensatzersteller*in, Datensatzeigentümer*in, Datensatzveröffentlicher*in sowie ergänzend Datensatzpfleger*in. Dabei wird auch deutlich, wie sich der Detailgrad in den Inhalten der Frameworks unterscheidet, wenn teilweise mehrere dieser Rollen abgefragt werden.

Fast alle Frameworks fragen nach der Art der Daten oder der Features, z.B. Text, Bilder oder tabellarische Daten. Auch die Anzahl und der Inhalt der Features bzw. die enthaltenen Datenfelder sollen in einigen Frameworks vermerkt werden. Darüber hinaus ist in den meisten Frameworks die Pflege und Aktualisierung des Datensatzes ein Thema.

Außer dem kurz gefassten *Data User Guide* thematisieren alle Frameworks sensible Daten und menschliche Attribute, die im Datensatz enthalten sein können, und deshalb besonderer Vorsicht bedürfen. Konkrete Fragen dazu umfassen unter anderem, welche demografischen Gruppen beschrieben werden und wie diese identifiziert wurden, wie sensible menschliche Attribute erfasst wurden (z.B. ethnische Zugehörigkeit, sexuelle Orientierung, sozio-ökonomischer Status) und ob Individuen identifi-

zierbar oder vertrauliche Daten enthalten sind. Letzteres ist im Rahmen von offenen Daten vernachlässigbar, da Daten mit diesen Eigenschaften nicht veröffentlicht würden. Bei Datensätzen, die Personen repräsentieren, wird außerdem in vier Frameworks das Thema der Zustimmung zur Datensammlung behandelt. Vier der Frameworks fragen darüber hinaus nach Verfahren und Überlegungen zur ethischen Überprüfung der Datenerhebung und möglichen Auswirkungen. In einzelnen Fällen werden speziell Privatsphäre und die Betrachtung systemischer Ungleichheiten miteinbezogen.

In allen Frameworks ist der Erhebungs- oder Sammelprozess der Daten ein Thema, teilweise mit detaillierten Unterfragen zu den verwendeten Werkzeugen, Mechanismen, Technologien und Software sowie der Frage, wer die Daten erhoben hat. Teil davon ist in einigen Frameworks auch die Frage nach der Repräsentativität der Daten und möglichen Limitationen dadurch, sowie der Bezug zu oder die Abhängigkeit von anderen Datensätzen. Ebenso ist die weitere Verarbeitung der Daten im Pre-Processing und der Datenbereinigung ein in allen Frameworks präsenter Aspekt, mit unterschiedlichem Detailgrad. Die vier für ML-Prozesse entwickelten Frameworks betrachten darüber hinaus Prozesse zum Labeling von Features und Attributen.

Alle Dokumentationsframeworks (also exklusiv der Datenbiographie als Reflexionsframework) betrachten bisherige, beabsichtigte bzw. potenzielle und/oder ungeeignete Anwendungen der Datensätze. Damit zusammen hängen auch Nutzungsbedingungen und Lizenzen, die ebenfalls in fast allen Dokumentationsframeworks vorkommen. Auch die Frage des Zugangs und der Veröffentlichung findet sich in allen Frameworks außer im *Data User Guide*, der im Kontext der Bereitstellung von offenen Daten entwickelt wurde und wo dementsprechend ein offener Zugang vorausgesetzt werden kann.

Einige seltenere Elemente der verschiedenen Frameworks wurden hier nicht erwähnt und der ausgeführte Vergleich kann nur einen ersten Eindruck der Vorgehensweisen und Umfänge der Frameworks geben. Eine tiefergehende Analyse könnte weitere Erkenntnisse zu konkreten Unterschieden und verschiedenen Fokussen bringen, übersteigt jedoch das Ausmaß dieses Papers.

Kontextinformationen und Metadaten

Wie bereits erwähnt, bezieht sich keines der Frameworks ausführlich auf bestehende Metadaten und Standards dafür. Zudem fehlt der Bezug zu Geodaten. Daher konzentriere ich mich im folgenden Abschnitt – nach einer kurzen Einführung in das Thema der Metadaten und die aktuelle Forschung dazu – auf einen Vergleich der in den sieben untersuchten Frameworks häufig angesprochenen Aspekten mit dem ISO-Metadatenstandard für Geodaten.

Grundsätzlich sollen Metadaten drei Hauptzwecke für potenzielle Nutzer*innen erfüllen: Datenquellen auffinden, die Nutzbarkeit evaluieren und verschiedene Datensätze miteinander integrieren [3, S. 6; 27, S. 133; 28, S. 1]. Insbesondere zu den letzten beiden Zwecken werden in der aktuellen GIS-Forschung jedoch aufgrund mangelhafter Nutzer*innenfreundlichkeit und Verständlichkeit von Metadaten Zweifel formuliert [4; 29]. Auch die Automatisierung der Erstellung und Aktualisierung von Metadaten wird in aktuellen Arbeiten behandelt und nutzt unter anderem Technologien wie ML, Natural Language Processing und Data Mining [30; 7; 31; 28, S. 1; 32]. Der Bedarf von Metadaten für eine kritische und reflektive Auseinandersetzung mit Datensätzen und ihren Bias – und damit die Limitationen von Automatisierung, wie Gebru et al. betonen [17, S. 3] – wird jedoch aktuell weniger thematisiert.

Form und Inhalt von Metadaten zu Datensätzen werden in verschiedenen Standards, Richtlinien und Metadatenschemas festgelegt. Eine bedeutende Institution hierfür ist die *Dublin Core Metadata Initiative* (DCMI), die bereits 1995 erstmals eine Sammlung von Kernelementen für Metadaten vielfältiger digitaler Ressourcen veröffentlichte und bis heute Spezifikationen erarbeitet [33; 34]. Bezüglich Konventionen für Metadaten speziell von Geodaten sei auf Standards der Internationalen Organisation für Normung (ISO), die *Infrastructure for Spatial Information in Europe* (INSPIRE) der Europäischen Kommission sowie die Arbeit des *Open Geospatial Consortium* (OGC) hingewiesen [29, S. 2]. In dieser Arbeit betrachte ich aufgrund des geographischen Fokus die ISO-Standards 19115-1:2014 (Geoinforma-

tion – Metadaten – Teil 1: Grundsätze) [35] sowie 19157-1:2023 (Geoinformation – Datenqualität – Teil 1: Allgemeine Anforderungen) [36]. Mein Vergleich der in der INSPIRE-Verordnung vorgeschriebenen Metadatenelemente [37] mit den ISO-definierten Elementen hat gezeigt, dass ersteres verallgemeinernd als Teilmenge des ISO-Standards betrachtet werden kann, und auch OGC bezieht sich auf den ISO-Standard [38]. Neben den Standards bilden Metadatenprofile und Vokabulare für Datenportale (z.B. GeoDCAT-AP) eine Grundlage für die Integration und Darstellung von Metadaten in Datenportalen und orientieren sich dabei häufig an dem ISO-Standard oder INSPIRE-Richtlinien [39, S. 1]. Sie werden in dieser Arbeit jedoch nicht weiter behandelt.

ISO 19115-1:2014 definiert zwölf Metadaten-Klassen, wobei die Klasse Datenqualität in einem eigenen Standard (ISO 19157-1:2023) beschrieben wird [35, S. 9]. Einige Klassen und Attribute innerhalb des ISO-Standards sind verpflichtend, die meisten jedoch freiwillig oder nur unter bestimmten Bedingungen erforderlich. Die Umsetzung und Kodierung des Metadatenmodells wird in ISO 19139:2019 (Geoinformation XML-Schema-Implementierung) spezifiziert.

Neben einem inhaltlichen Vergleich der in den beiden ISO-Standards definierten Metadatenelementen (ISO 19115-1:2024 und 19157-1:2023) mit den häufig in den Frameworks genannten Kontextinformationen werden auch einige strukturelle Unterschiede deutlich. Sichtbar wird ebenso, dass der ISO-Standard ein sehr umfassendes Modell liefert, denn er deckt viele Aspekte der Frameworks ab, sowie weitere geodatenbezogene Komponenten, die in den Frameworks nicht enthalten sind.

Wie bereits angesprochen fragen alle Frameworks nach dem Erhebungsprozess der Daten und Schritten der Vorverarbeitung bzw. Datenbereinigung, teilweise mit detaillierten Fragen zu den verwendeten Methoden, Werkzeugen oder beteiligten Personen. Der ISO-Standard sieht vor, die Informationen zu den verwendeten Quellen und Produktionsprozessen in dem Paket *LI_Lineage* festzuhalten. Dieses Paket kann sowohl eine Textbeschreibung als auch separate Quellen und Prozessschritte in eigenen Objekten mit verschiedenen Attributen enthalten. Der ISO-Standard bietet damit ebenfalls einen Behälter für Kontextinformationen hinsichtlich des Produktions- und Verarbeitungsprozesses, wenn auch weniger konkret als die Frameworks.

Derselbe strukturelle Unterschied zeigt sich in weiteren Aspekten. Ein Beispiel ist das Attribut *credit* in der *MD_Identification* des ISO-Standards, das genutzt werden kann, um die Personen oder Organisationen anzuerkennen, die zum Datensatz beigetragen haben. Obwohl dieses Attribut verwendet werden kann, um Informationen zu den konkret an der Datenerhebung beteiligten Personen bereitzustellen, fragt der ISO-Standard im Vergleich zu einigen Frameworks nicht explizit danach. Auch bei den Rollen der Datensatzersteller*in, Datensatzeigentümer*in, Datensatzveröffentlicher*in sowie ergänzend der Datensatzpfleger*in zeigt sich, dass der ISO-Standard ein Modell für die Bereitstellung von Informationen vorgibt (in diesem Fall *MD_Identification.pointOfContact* mit verschiedenen Rollendefinitionen in *CI_Responsibility*), jedoch nicht explizit Fragen formuliert, auf die diese Informationen eine Antwort liefern sollen. Die Frameworks als Leitfäden verfolgen dabei einen anderen Ansatz. Ähnlich ist es mit der Klasse *CI_Date*, wo über den *dateType* die bereits erwähnten Datentypen wie Veröffentlichungsdatum, Zeitpunkt der Datenerstellung, Aktualisierungsdatum und noch mehr definiert werden können. Um einen Eindruck vom Umfang des ISO-Standards zu bekommen, reicht es nicht aus, nur die Klassen und ihre Attribute zu betrachten, sondern auch die für einige Attribute definierten Codelisten, die einen Eindruck davon vermitteln, welche Informationen dort dokumentiert werden können.

Folgende der in den Frameworks häufig behandelten Metadatenelemente tauchen ebenfalls im ISO-Standard auf: die Motivation bzw. die Intention, mit der der Datensatz erstellt wurde (*MD_Idenfiticatio-n.purpose*), die Pflege und Aktualisierungsregelungen zum Datensatz (*MD_Identification.resource-Maintenance*), Zugang und Nutzungsbedingungen (ausführlicher als in den meisten Frameworks, definiert in *MD_Identification.resourceConstraints*) sowie die Art der Daten und die Datenfelder (*MD_Con-tentInformation*). Die Finanzierung wird nicht als spezielles Attribut bereitgestellt, lässt sich aber ebenfalls durch die Rollendefinition innerhalb der Klasse *CI_Responsibility* als Partei, die finanzielle Unterstützung für den Datensatz bereitstellt, angeben. Anzumerken ist außerdem, dass das Metadatenmo-

dell von ISO auf verschiedene Granularitäten anwendbar ist, und Metadatenelemente auch für einzelne Teilmengen oder Attribute im Datensatz angegeben werden können, um detailliertere Informationen auf Ebene der Dateninhalte bereitzustellen.

Ein deutlicher inhaltlicher Unterschied zwischen den Frameworks und dem ISO-Standard zeigt sich bei den Kontextinformationen zu Datensätzen, die Personen beschreiben, und daraus resultierenden Herausforderungen, Limitationen und möglichen Verzerrungen, z.B. zu sensiblen Daten, Repräsentativität oder angewandten sozialen Kategorien. Auch ethische Überlegungen und Bedenken sind im ISO-Standard nicht konkret enthalten. Eine mögliche Erklärung könnte darin liegen, dass Geodaten seltener personenbezogen sind. INSPIRE im speziellen bezieht sich auf umweltbezogene Datensätze. Grundsätzlich können jedoch auch Geodaten Menschen beschreiben und können Daten über Personengruppen einen räumlichen Bezug haben, da Geodaten im weitesten Sinne als „Daten mit einem impliziten oder expliziten Ortsbezug zur Erde“ [36, NA 7] definiert werden. Das Fehlen von entsprechenden Klassen und Attributen zu diesem Thema lässt sich daher als relevante Lücke im ISO-Standard sehen.

Der ISO-Standard enthält ebenso weniger Informationen zu vorgesehenen, bisherigen und ungeeigneten Anwendungen des Datensatzes. *MD_Identification.resourceSpecificUsage* sieht lediglich die Angabe bisheriger und aktueller Anwendungen des Datensatzes durch verschiedene Nutzer*innen vor [35, S. 41].

Einige Kontextinformationen sind im ISO-Standard als Metadatenelemente definiert, kommen jedoch nicht, weniger detailliert oder nur in wenigen Frameworks vor. Dazu gehören beispielsweise das Datenformat, die Sprache des Datensatzes, Informationen zu den Metadaten, zugehörige Ressourcen, die zeitliche Auflösung oder die Datenqualität. Besonders relevant sind in diesem Zusammenhang jedoch die für Geodaten spezifischen Metadatenelemente, die aufgrund des fehlenden Geo-Fokus in den Reflexions- und Dokumentationsframeworks nicht behandelt werden. Hierzu zählen die räumliche Repräsentation, die räumliche Auflösung, die räumliche Ausdehnung, der Maßstab, das Koordinatenreferenzsystem sowie einige der Datenqualitätsindikatoren.

Der ISO-Standard 19157, der sich mit der Beschreibung der Datenqualität von räumlichen Daten befasst, definiert dafür fünf Gruppen an Qualitätselementen: Vollständigkeit (einschließlich überschüssiger und fehlender Daten), logische Konsistenz (konzeptionell, domänenbezogen, im Format, sowie topologisch), positionelle Genauigkeit (absolut, relativ und im Raster), zeitliche Qualität (Genauigkeit, Konsistenz, Gültigkeit) sowie thematische Qualität (Klassifizierung, Genauigkeit quantitativer und nicht-quantitativer Attribute). Einzelne Aspekte von Datenqualität werden in vier der Frameworks angesprochen, nehmen dort jedoch deutlich weniger Raum ein und sind weniger ausdifferenziert.

Der Vergleich zwischen den untersuchten Frameworks und dem ISO-Standard zeigt deutliche Unterschiede in Struktur und Format. Während die Frameworks als Leitfäden überwiegend konkrete Fragen stellen oder Punkte zum Ausfüllen anbieten, häufig in offener Textform, beschreibt der ISO-Standard in einem Modell, wie welche Informationen dokumentiert werden können und legt dort deutlich mehr Format fest. Der ISO-Standard ist ein höchst formalisiertes Dokument, für fachfremde Personen schwer verständlich und nur gegen eine kostenpflichtige Lizenz zugänglich.

Inhaltlich gibt es jedoch zahlreiche Überschneidungen. Abgesehen von Lücken bezüglich ethischer Überlegungen, Herausforderungen und Einschränkungen sowie bei Daten über Personen, deckt der ISO-Standard einen Großteil der in den Frameworks angesprochenen Kontextinformationen ab und hat somit großes Potenzial für die Bereitstellung von Kontextinformationen. Offensichtlich wird dies in der Veröffentlichung von offenen Daten, mit Metadaten oft angelehnt an den ISO-Standard, jedoch selten umfassend genutzt. Wie anfangs ausgeführt, fehlt häufig der Kontext von offenen Daten. Existenz und Umfang des Standards allein garantieren nicht, dass die entsprechenden Metadaten erstellt und veröffentlicht werden [40, S. 2]. Einige Studien zeigen, dass die Metadatenqualität in Datenportalen häufig unzureichend ist und viele Metadaten fehlen [40; 29, S. 2]. Besonders kritisch ist der Mangel an Informationen zur Datenqualität und der Provenienz, also den Quellen und dem Produktions-

prozess [7, S. 1511–1512; 41, S. 314; 42]. Auch die Lesbarkeit und damit Nutzbarkeit von Metadaten wird in Studien kritisiert [7, S. 1513; 27, S. 134; 28; 43, S. 1].

Use Case

Beispielhaft untersuche ich die Verfügbarkeit von Kontextinformationen in Metadaten anhand eines Use Cases und drei offenen Datensätzen. Als Anwendungsfall betrachte ich die Webkarte ‚Grätzlfarben' [44], an deren Umsetzung in meiner Forschungsgruppe ich maßgeblich beteiligt war. Die Kartenanwendung stellt die Verteilung der Flächennutzung in Wien dar und ermöglicht es Nutzer*innen, basierend auf diesen Daten eine Postkarte mit der Flächennutzung in einer bestimmten Nachbarschaft (auf Wienerisch: Grätzl) zu erstellen (Abb. 1). Die Anwendung basiert auf der Open-Source-Anwendung ‚*Kiezcolors*' der Open Data Informationsstelle Berlin [45].



Abb 1: Screenshot der ‚Grätzlfarben' Webanwendung [44]

Für die Wiener Umsetzung ‚Grätzlfarben' wird bisher der offene Datensatz zur Realnutzungskartierung der Stadt Wien [46] verwendet. Benötigt werden für die Anwendung Vektordaten mit flächendeckenden Informationen zur Flächennutzung für Wien und ggf. darüber hinaus. Diese Informationen können u.a. auch aus dem Österreichischen Kataster [47] und OpenStreetMap (OSM) [48] extrahiert werden. Der Use Case wurde gewählt, weil er einen praxisnahen Anwendungsfall darstellt, bei dem die Fitness-For-Use von verschiedenen Datensätze verglichen werden muss. Eine Limitation dieses Anwendungsfalls ist jedoch, dass keine Daten zu Personen verwendet werden, deren Reflexion einen wesentlichen Bestandteil in den Frameworks darstellt. Entsprechend gibt es in den Metadaten zu den Datensätzen keine Informationen zu diesen Aspekten.

Die einzelnen untersuchten Datensätze mit Verweisen zu den Metadaten sowie zu zusätzlicher Dokumentation, die für die Untersuchung herangezogen wurde, sind in Tabelle 2 aufgelistet.

Tabelle 2: Untersuchte Datensätze mit Verweisen zu dazugehöriger Dokumentation

Datensatz	Metadaten	Zusätzlich untersuchte Dokumentation
Realnutzungskartierung	[46]	Produktbeschreibung [49]
Kataster Stichtagsdaten	Einzelne Einträge für verschiedene Dateiformate mit teilweise unterschiedlichen Metadaten: [50–52]	Produktbeschreibung [53] Schnittstellenbeschreibungen [54; 55]
OpenStreetMap	Nicht verfügbar	OSM-Wiki [56]

OSM als Volunteered Geographic Information Datenbank verfügt über keine klassischen Metadaten. Aufgrund der Heterogenität der Beitragenden und Datenerhebungsprozesse sind viele der genannten Metadatenelemente nicht auf den gesamten Datensatz anwendbar. Das OSM-Wiki bietet jedoch einige allgemeine Kontextinformationen zum Projekt, etwa zur Motivation, Finanzierung, Anwendungen, Zugang, Lizenzierung, Datenstruktur und Definitionen bestimmter Datenfelder (*Tags*). Jedoch muss hier die Granularität bedacht werden, denn diese Informationen lassen sich nicht auf einzelne Datenerhebungen übertragen – so bleibt beispielsweise unklar (es sei denn Kommentare in *Changesets* oder Hinweise bei Features weisen darauf hin), welche Intention hinter einer konkreten Datenerfassung steht, ob die im Wiki beschriebene Definition eines *Tags* eingehalten wurde, oder ob Kartendaten von Unternehmen finanziert und beigetragen wurden. Einige Metainformationen zu den Daten, wie Zeitpunkt der Erstellung, letzte Aktualisierung, Datenerfasser*in oder verwendete Tools, können in manchen Fällen durch Analyse der einzelnen Features und *Changesets* extrahiert werden. Allerdings bleibt fraglich, wie erkenntnisreich solche fragmentierten Analysen für die Arbeit der Datennutzer*innen sind. Darüber hinaus bieten wissenschaftliche Untersuchungen Einblicke in allgemeinere Trends, Verzerrungen und Lücken im OSM-Datensatz [57, S. 1].

Die Realnutzungskartierung der Stadt Wien und die Katasterdaten des Bundesamtes für Eich- und Vermessungswesen (BEV) sind über das österreichische Open-Data-Portal *data.gv.at* zugänglich. Für den Katasterdatensatz beziehe ich mich jedoch auf den umfangreicheren Eintrag im eigenen Datenportal des BEV, dessen Metadaten auf dem ISO-Standard beruhen. Auf *data.gv.at* wird die Metadatenstruktur ‚OGD Austria Metadata 2.6' verwendet, welche sich ebenfalls auf INSPIRE und ISO bezieht [58]. Produktbeschreibungen und für den Kataster verfügbare Schnittstellenbeschreibungen wurden in der Analyse zusätzlich zu den Metadaten in den Datenportaleinträgen herangezogen.

Beide Datensätze enthalten grundlegende Informationen zum Inhalt des Datensatzes, der Lizenz und den Nutzungsbedingungen. Hinweise zu Aktualisierungen sind in beiden Fällen allgemein und eher unspezifisch gehalten. Für beide Datensätze wird die Institution oder eine Organisationseinheit als datenverantwortliche Stelle mit einer Kontaktmöglichkeit angegeben, eine Differenzierung in verschiedene Rollen beteiligt an der Datenerhebung oder -verarbeitung gibt es jedoch nicht. Die untersuchten Dokumente enthalten keine Information zum Zweck oder der Motivation der Datenerhebung. Auch domänenspezifische Hintergründe oder Definitionen für die angewandte Klassifikation in verschiedene Nutzungsarten sind nur begrenzt verfügbar. In den Schnittstellenbeschreibungen des Katasterdatensatzes gibt es zum Teil Ausführungen zum Dateiformat *GeoPackage* oder *Shapefile*, es ist jedoch davon auszugehen, dass die Informationen für Datennutzer*innen ohne Erfahrung mit Geodaten nicht ausreichend sind. Ähnlich ist es mit den im BEV-Datenportal angegebenen Schlagworten, welche ohne fachspezifisches Wissen schwer verständlich sind. Die im Datensatz enthaltenen Attribute sind in den Schnittstellenbeschreibungen des BEV umfassend aufgeschlüsselt, während für die Realnutzungskartierung nur einzelne Attribute zur Nutzung beschrieben werden.

Der räumliche Umfang wird in beiden Fällen als Text und als geographische *Bounding Box* angegeben und ist somit sowohl maschinen- als auch menschenlesbar. Das Koordinatenbezugssystem wird nur in den Metadaten des Katasters kommuniziert (abhängig vom Dateiformat und der räumlichen Ausdehnung). Bei der Realnutzungskartierung der Stadt Wien ist eine Internetrecherche nötig, um herauszu-

finden, dass alle offenen Geodatensätze der Stadt Wien in EPSG 31256 (MGI / Datum Austria, Gauß-Krüger Zone Ost) veröffentlicht werden [59].

Die Angaben zu Quellen, Erhebungs-, Verarbeitungs- und Validierungsprozessen sind bei beiden Datensätzen oberflächlich. Detaillierte Produktions- oder Herkunftsangaben bleiben aus, obwohl sie aufgrund der Integration verschiedener Datenquellen aufschlussreich sein könnten. Die Datenqualität wird bei beiden Datensätzen thematisiert, jedoch nicht umfassend untersucht. Beim Kataster sind die detaillierteren Informationen verfügbar, insbesondere zur positionellen Genauigkeit und zur Vollständigkeit. Bei der Realnutzungskartierung fehlen darüber hinaus Informationen zum Veröffentlichungs-, Aktualisierungs- und/oder Erstellungsdatum. Einzig das Aktualisierungsdatum der Metadaten ist verfügbar. Beim Kataster hingegen ist sowohl das Erstellungsdatum als auch das Datum der Publikation und der letzten Überarbeitung angegeben.

Zu intendierten, bisherigen und ungeeigneten Datennutzungen gibt es begrenzte Informationen: in der Produktbeschreibung des Katasters wird erwähnt, für welche Anwendungszwecke der Stichtag-Datensatz geeignet ist (z.B. Übersichten, großflächige Planungen, Informationslayer in GIS) und wofür nicht (eigentumsrelevante Aspekte), während im Datenportaleintrag der Realnutzungskartierung auf von Nutzer*innen eingereichte Anwendungen verwiesen wird.

Die Katastererhebung, einschließlich der Benützungsarten und Nutzungen, ist durch Verordnungen [60; 61] geregelt, sodass eine vertiefte Recherche in den damit verbundenen Dokumenten weitere Kontextinformationen liefern könnte, etwa zu den angewandten Klassifikationen und Mindestgrößen von Elementen für ihre Erfassung. Im Sinne einer Datenbiographie, die eine umfassende Untersuchung des Datenkontexts durch die Datennutzer*innen erfordert, wäre die Betrachtung dieser Dokumente notwendig. Da ich mich in dieser Arbeit jedoch auf das Potenzial und den Status Quo von Metadaten konzentriere, wurde diese weiterführende Recherche nicht weiterverfolgt.

Die bereitgestellten Metadaten der beiden offenen Verwaltungsdatensätze sind insgesamt begrenzt und decken viele der in den Frameworks und ISO-Standards genannten Aspekte nur unvollständig oder oberflächlich ab. Der Katasterdatensatz enthält mehr Informationen als die Realnutzungskartierung, doch besonders in den Bereichen Provenienz, Motivation, Nutzung und Qualität gibt es in beiden Fällen Lücken. Obwohl der angewendete Use Case nicht repräsentativ ist, stimmen diese Erkenntnisse mit den bereits angesprochenen Forschungsergebnissen aus der Literatur überein. Zudem wird durch die Gegenüberstellung die besondere Herausforderung bei der Bereitstellung und Extraktion von Kontextinformationen für crowd-sourced Daten wie OSM deutlich.

Für eine Fitness-For-Use Analyse für ‚Grätzlfarben' müssen selbstverständlich neben den Kontextinformationen auch die Anforderungen der Anwendung berücksichtigt und Tests mit den verschiedenen Datensätzen durchgeführt werden. Aus diesem Grund kann an dieser Stelle noch keine Entscheidung für den am besten geeigneten Datensatz getroffen werden kann. Der Use Case hat jedoch gezeigt, welche Informationen aus den Metadaten extrahiert werden können und welche Kontextinformationen unbekannt bleiben, sodass diese entweder weiter recherchiert oder ihre Unsicherheit reflektiert werden sollte.

Diskussion und Ausblick

Das Wissen über den Entstehungskontext eines Datensatzes ist von großer Bedeutung für seine Interpretation und Analyse. Anhaltspunkte für solche Kontextinformationen finden sich in Reflexions- bzw. Dokumentationsframeworks und Metadaten. Der Vergleich verschiedener Frameworks und dem ISO-Metadatenstandard für geographische Daten zeigt viele Überschneidungen, aber auch unterschiedliche Schwerpunkte. Inhaltlich deckt der ISO-Standard bereits viele der in den Frameworks angesprochenen Kontextinformationen ab, jedoch fehlen ethische Überlegungen, Herausforderungen und Informationen bei Daten, die Menschen beschreiben. Im Gegenzug beinhalten ISO 19115:1-2024 und 19157-1:2023 einige über die Frameworks hinausgehenden geodatenbezogene Attribute und detail-

lierte Aspekte der Datenqualität. Vor allem in diesen Unterschieden können sich die verschiedenen Formate gegenseitig ergänzen.

Inhaltlich sollte tiefergehend untersucht werden, inwiefern die für eine reflektierte, verantwortungsvolle und kritische Geodatenpraxis erforderliche Kontextinformationen damit abgedeckt sind und wo noch Lücken liegen. Auch Aspekte, die nur in wenigen der Frameworks erwähnt werden und deshalb in dieser Arbeit vernachlässigt wurden, sollten weiter betrachtet werden, wie die Frage des ‚Wer' und der damit einhergehenden Perspektive [62, S. 12].

Aufgrund der verschiedenen Zielsetzungen – die Frameworks sollen als Leitfaden und Vorlage dienen, während die ISO-Standards Modelle vorschlagen – unterscheiden sich die Formate und die Anwendung. Potenziale könnten sich aus der Nutzbarmachung der ISO-Metadatenelemente im Sinne eines geospezifischen Reflexions- oder Dokumentationsframeworks oder einer Checkliste mit standardisierten Fragen ergeben. Eine besondere Herausforderung stellt die Anwendung der Frameworks und eines Metadatenstandards für crowd-sourced Daten wie OSM dar, wo alternative Ansätze für die Bereitstellung, Extraktion und Reflexion von Kontextinformationen notwendig sind.

Trotz bereits bestehender Frameworks und umfangreicher Standards, wie ISO 19115:1-2014, stellt der begrenzte Zugang zu Kontextinformationen von Datensätzen eine Herausforderung für die Reflexion eines Datenkontexts dar. Zu beleuchten bleiben deshalb Gründe für das Fehlen vieler Informationen in den Metadaten offener Daten und Möglichkeiten, um dem entgegenzusteuern. Neben der Förderung eines höheren Bewusstseins für die Relevanz von Kontextinformationen eines Datensatzes sollten auch technische Ansatzpunkte wie die Entwicklung eines erweiterten Metadatenprofils, die Integration umfassenderer Kontextinformationen in Metadateneditoren und Geodateninfrastrukturen sowie Möglichkeiten und Grenzen der Automatisierung untersucht werden. Dafür ist zunächst eine tiefergehende Auseinandersetzung mit bestehenden Arbeiten und Tools nötig. Auch die Nutzer*innenfreundlichkeit und Verständlichkeit der Metadaten müssen berücksichtigt werden.

Parallel führt der Anspruch an eine verantwortungsvolle und reflexive Geodatenpraxis zu den Fragen, wie der Kontext eines Datensatzes von Nutzer*innen reflektiert werden kann, wenn die zur Verfügung gestellten Informationen lückenhaft sind und wer die Verantwortung der Erstellung und Veröffentlichung dieser Kontextinformationen tragen sollte. Zu betrachten sind dabei nicht nur Potenziale für Datennutzer*innen sondern auch für Datenproduzent*innen sowie weitere beteiligte Parteien.

Obwohl der Zugang zu Kontextinformationen und ein Leitfaden für die Reflexion für eine verantwortungsvolle Geodatenpraxis wichtig sind, stehen sie nur am Anfang einer GIS-Analyse oder eines Kartographieprojekts. Daher bleibt zu erforschen, wie die Erkenntnisse aus einer Reflexion über den Entstehungskontext von Daten den weiteren Prozess und das Endprodukt beeinflussen und verbessern können. Vor dem Hintergrund zunehmender KI-generierter Inhalte spielen dabei auch Transparenz und Vertrauen eine wichtige Rolle.

Abschließend sollte nicht nur die Auseinandersetzung mit dem Kontext einzelner Datensätze gefördert, sondern auch die Strukturen der Wissensproduktion kritisch hinterfragt werden [1, S. 171–172]: Welche Daten werden aufgrund von Machtungleichheiten nicht erhoben? Welche Datensätze werden aufgrund bestimmter Interessen nicht veröffentlicht? Wessen Wissen wird unterdrückt? Eine verantwortungsvolle Geodatenpraxis erfordert nicht nur die Reflexion über den Kontext bestimmter Daten, sondern auch die kritische Auseinandersetzung mit den zugrunde liegenden Strukturen.

Kontakt zur Autorin:

Ester Scheck
TU Wien
Wiedner Hauptstraße 8-10 / E120-6, 1040 Wien, Österreich
ester.scheck@geo.tuwien.ac.at

Literatur und Quellen

[1] D'Ignazio, C.; Klein, L.F.: Data Feminism. The MIT Press, Cambridge, 2020.

[2] Haraway, D.: Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *In:* Feminist Studies 14 (1988), Heft 3, S. 575. https://doi.org/10.2307/3178066.

[3] Comber, A.; Fisher, P.; Wadsworth, R.: User-focused metadata for spatial data, geographical information and data quality assessments. In: Wachowicz, M.; Bodum, L. (Hrsg.): Proceedings of the 10th AGILE International Conference on Geoinformation Science, Aalborg University, 2007.

[4] Devillers, R.; Bédard, Y.; Gervais, M. et al.: How to Improve Geospatial Data Usability: From Metadata to Quality-Aware GIS Community. In: AGILE (Hrsg.): AGILE Pre-Conference Workshop, 2007.

[5] Chmielinski, K.S.; Newman, S.; Taylor, M. et al.: The Dataset Nutrition Label (2nd Gen): Leveraging Context to Mitigate Harms in Artificial Intelligence. In: NeurIPS (Hrsg.): Workshop on Dataset Curation and Security, 2020.

[6] D'Ignazio, C.: Creative data literacy. *In:* Information Design Journal 23 (2022), Heft 1, S. 6-18. https://doi.org/10.1075/idj.23.1.03dig.

[7] Fischer, J.; Egli, L.; Groth, J. et al.: Approaches and tools for user-driven provenance and data quality information in spatial data infrastructures. *In:* International Journal of Digital Earth 16 (2023), Heft 1, S. 1510-1529. https://doi.org/10.1080/17538947.2023.2198778.

[8] Whitfield, P.H.: Why the Provenance of Data Matters: Assessing Fitness for Purpose for Environmental Data. *In:* Canadian Water Resources Journal / Revue canadienne des ressources hydriques 37 (2012), Heft 1, S. 23-36. https://doi.org/10.4296/cwrj3701866.

[9] Krause, H.: Data Biographies: Getting to Know Your Data, 2017, https://gijn.org/stories/data-biographies-getting-to-know-your-data/ [Zugriff am: 17.10.2024].

[10] Holland, S.; Hosny, A.; Newman, S. et al.: The Dataset Nutrition Label – A Framework to Drive Higher Data Quality Standards. In: Hallinan, D.; Leenes, R.; Gutwirth, S. et al. (Hrsg.): Data Protection and Privacy. Hart Publishing, 2020.

[11] O'Neil, C.: Weapons of math destruction – How big data increases inequality and threatens democracy. Penguin Books, London, 2016.

[12] Robinson, A.H.; Morrison, J.L.; Muehrcke, P.C. et al.: Elements of cartography. Wiley, New York, NY, 1995.

[13] Duden: Kontext, 2018, https://www.duden.de/rechtschreibung/Kontext [Zugriff am: 20.12.2024].

[14] Krause, H.: An Introduction to the Data Biography, 2023, https://weallcount.com/2019/01/21/an-introduction-to-the-data-biography/ [Zugriff am: 02.10.2024].

[15] Western Pennsylvania Regional Data Center: Data User Guides, n.d., https://www.wprdc.org/en/data-user-guides [Zugriff am: 16.12.2024].

[16] Civic Software Foundation: CIVIC Structured Context Program, 2020, https://civicsoftwarefoundation.org/posts/structured-context-data-feminism [Zugriff am: 17.12.2024].

[17] Gebru, T.; Morgenstern, J.; Vecchione, B. et al.: Datasheets for Datasets, 24. März 2018.

[18] Microsoft Research: Data Documentation - Microsoft Research, 2022, https://www.microsoft.com/en-us/research/project/datasheets-for-datasets/ [Zugriff am: 09.10.2024].

[19] Pushkarna, M.; Zaldivar, A.; Kjartansson, O.: Data Cards: Purposeful and Transparent Dataset Documentation for Responsible AI. In: ACM (Hrsg.): 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul Republic of Korea, 2022, S. 1776-1826.

[20] We All Count: Data Biography 2023 Template Detailed, 2023, https://docs.google.com/spreadsheets/d/1z5qlMasRLBaagd35nua5hVqH5VQZPTt2U_CQkspwlfM/edit?gid=0#gid=0 [Zugriff am: 02.10.2024].

[21] Pushkarna, M.; Zaldivar, A.; Kjartansson, O.: Data Cards Playbook Templates, 2023, https://github.com/PAIR-code/datacardsplaybook/tree/main/templates [Zugriff am: 18.12.2024].

[22] Western Pennsylvania Regional Data Center: Data-User-Guides – GitHub, 2015, https://github.com/rgradeck/Data-User-Guides/ [Zugriff am: 16.12.2024].

[23] Civic Software Foundation: CIVIC Data Library of Context – Structured Context Demo, https://www.civicdatalibrary.org/context [Zugriff am: 11.10.2024].

[24] Gebru, T.; Morgenstern, J.; Vecchione, B. et al.: Datasheets for Datasets Template, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE4t8QB [Zugriff am: 09.10.2024].

[25] Microsoft's Aether Transparency Working Group: Aether Data Documentation Template (Draft 08/25/2022), 2022, https://www.microsoft.com/en-us/research/uploads/prod/2022/07/aether-datadoc-082522.pdf [Zugriff am: 09.10.2024].

[26] Data Nutrition Project: The Dataset Nutrition Label, https://labelmaker.datanutrition.org/ [Zugriff am: 03.10.2024].

[27] Henzen, C.; Mäs, S.; Bernard, L.: Provenance Information in Geodata Infrastructures. In: Vandenbroucke, D.; Bucher, B.; Crompvoets, J. (Hrsg.): Geographic Information Science at the Heart of Europe. Springer International Publishing, Cham, 2013, S. 133-151.

[28] Kalantari, M.; Syahrudin, S.; Rajabifard, A. et al.: Spatial Metadata Usability Evaluation. *In:* ISPRS International Journal of Geo-Information 9 (2020), Heft 7. https://doi.org/10.3390/ijgi9070463.

[29] Ziaimatin, H.; Nili, A.; Barros, A.: Reducing Consumer Uncertainty: Towards an Ontology for Geospatial User-Centric Metadata. *In:* ISPRS International Journal of Geo-Information 9 (2020), Heft 8. https://doi.org/10.3390/ijgi9080488.

[30] Batcheller, J.K.; Gittings, B.M.; Dunfey, R.I.: A Method for Automating Geospatial Dataset Metadata. *In:* Future Internet 1 (2009), Heft 1, S. 28-46. https://doi.org/10.3390/fi1010028.

[31] Jochim, C.: Natural Language Processing and Information Retrieval Methods for Intellectual Property Analysis. Stuttgart, Universität Stuttgart, Dr. phil., 2014.

[32] Wagner, M.; Henzen, C.; Müller-Pfefferkorn, R.: A Research Data Infrastructure Component for the Automated Metadata and Data Quality Extraction to Foster the Provision of FAIR Data in Earth System Sciences. *In:* AGILE: GIScience Series 2 (2021), S. 1-7. https://doi.org/10.5194/agile-giss-2-41-2021.

[33] DCMI: About DCMI, 2025, https://www.dublincore.org/about/ [Zugriff am: 19.01.2025].

[34] DCMI: DCMI History, 2025, https://www.dublincore.org/about/history/ [Zugriff am: 19.01.2025].

[35] International Organization for Standardization: ISO 19115-1:2014 – Geographic information — Metadata. Part 1: Fundamentals, 2014.

[36] International Organization for Standardization: ISO 19157-1:2023 – Geographic information — Data quality. Part 1: General requirements, 2023.

[37] European Commission: Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata – 1205/2008. European Commission, 2008.

[38] Open Geospatial Consortium: OGC e-Learning - Metadata - Introduction, 2024, http://opengeospatial.github.io/e-learning/metadata/text/main.html [Zugriff am: 14.01.2025].

[39] Perego A.; Cetl V.; Friis-Christensen A. et al.: GeoDCAT-AP: Representing geographic metadata by using the "DCAT application profile for data portals in Europe", 2017.

[40] Quarati, A.; Martino, M. de; Rosim, S.: Geospatial Open Data Usage and Metadata Quality. *In:* ISPRS International Journal of Geo-Information 10 (2021), Heft 1. https://doi.org/10.3390/ijgi10010030.

[41] Martino, M. de; Rosim, S.; Quarati, A.: Hydrographic Datasets in Open Government Data Portals: Mitigation of Reusability Issues Through Provenance Documentation. In: Garoufallou, E.; Falluchi, F.; de Luca, E.W. (Hrsg.): Metadata and Semantic Research – 13th International Conference, MTSR. Springer, 2019, S. 307-319.

[42] Yang, X.; Blower, J.D.; Bastin, L. et al.: An integrated view of data quality in Earth observation. *In:* Philosophical transactions. Series A, Mathematical, physical, and engineering sciences 371 (2013). https://doi.org/10.1098/rsta.2012.0072.

[43] Lush, V.; Bastin, L.; Lumsden, J.: Developing a GEO label: providing the GIS community with quality metadata visualisation tools. In: GISRUK (Hrsg.): Proceedings of the 21st GIS research UK conference, 2013.

[44] Forschungsbereich Kartographie, TU Wien: Grätzlfarben. Forschungsbereich Kartographie, TU Wien, 2024, https://cartolab.at/graetzlfarben/ [Zugriff am: 19.12.2024].

[45] ODIS: Kiezcolors, 2023, https://kiezcolors.odis-berlin.de/ [Zugriff am: 19.12.2024].

[46] Stadt Wien: Realnutzungskartierung ab 2007 Wien - data.gv.at, 2024, https://www.data.gv.at/katalog/de/dataset/stadt-wien_realnutzungskartierungab200708wien [Zugriff am: 19.12.2024].

[47] Bundesamt für Eich- und Vermessungswesen: Österreichischer Kataster, 2024, https://kataster.bev.gv.at/ [Zugriff am: 19.12.2024].

[48] OpenStreetMap: OpenStreetMap, 2024, https://www.openstreetmap.org/ [Zugriff am: 19.12.2024].

[49] Stadt Wien: Realnutzungskartierung - Flächennutzung im Stadtgebiet, https://www.wien.gv.at/stadtentwicklung/grundlagen/stadtforschung/siedlungsentwicklung/realnutzungskartierung/ [Zugriff am: 19.12.2024].

[50] Bundesamt für Eich- und Vermessungswesen: Kataster Grafik Grundstücksverzeichnis GPKG Stichtag 01.10.2024 – Datenkatalog BEV, 2024, https://data.bev.gv.at/geonetwork/srv/ger/catalog.search#/metadata/fae0e053-b22a-4bbd-a3d3-442fc632a816 [Zugriff am: 19.12.2024].

[51] Bundesamt für Eich- und Vermessungswesen: Kataster Grafik INSPIRE Stichtag 01.10.2024 – Datenkatalog BEV, 2024, https://data.bev.gv.at/geonetwork/srv/ger/catalog.search#/metadata/908ec148-91e9-4a36-b6c0-4ee52f6e5a50 [Zugriff am: 19.12.2024].

[52] Bundesamt für Eich- und Vermessungswesen: Kataster Grafik SHP Stichtag 01.10.2024 – Datenkatalog BEV. Bundesamt für Eich- und Vermessungswesen, 2024, https://data.bev.gv.at/geonetwork/srv/ger/catalog.search#/metadata/3d6a823f-08ba-41a8-b838-a841d5b82a59 [Zugriff am: 19.12.2024].

[53] Bundesamt für Eich- und Vermessungswesen: Kataster Stichtagsdaten – Informationen zum Produkt. Bundesamt für Eich- und Vermessungswesen, 2024, https://www.bev.gv.at/Services/Produkte/Kataster-und-Verzeichnisse/Kataster-Stichtagsdaten.html [Zugriff am: 19.12.2024].

[54] Bundesamt für Eich- und Vermessungswesen: Digitale Katastralmappe und Grundstücksdaten Stichtagsdaten GPKG – Schnittstellenbeschreibung. Version 1.0, 26. Juli 2023.

[55] Bundesamt für Eich- und Vermessungswesen: Katastralmappe SHP – Schnittstellenbeschreibung. Version 2.9, 4. Dezember 2024.

[56] OpenStreetMap: OpenStreetMap Wiki, 2024, https://wiki.openstreetmap.org/ [Zugriff am: 19.12.2024].

[57] Grinberger, A.Y.; Minghini, M.; Yeboah, G. et al.: Bridges and Barriers: An Exploration of Engagements of the Research Community with the OpenStreetMap Community. *In:* ISPRS International Journal of Geo-Information 11 (2022), Heft 1, S. 1-29. https://doi.org/10.3390/ijgi11010054.

[58] OGD Österreich: Metadaten data.gv.at, 2021, https://go.gv.at/ogdmetade [Zugriff am: 19.01.2025].

[59] Stadt Wien: OGD-Daten Download - Stadtvermessung Wien, https://www.wien.gv.at/stadtentwicklung/stadtvermessung/geodaten/viewer/geodatendownload.html [Zugriff am: 20.12.2024].

[60] Bundesamt für Eich- und Vermessungswesen: Information zur Neuordnung der BANU im Kataster, 1. April 2014.

[61] Republik Österreich: Verordnung des Bundesministers für Wirtschaft, Familie und Jugend über die Angabe und Definition der Benützungsarten und Nutzungen im Grenzkataster – Benützungsarten-Nutzungen-Verordnung – BANU – V, BGBl. II Nr. 116/2010. Republik Österreich, 2010.

[62] Burns, A.; On, T.; Lee, C. et al.: Making the Invisible Visible: Risks and Benefits of Disclosing Metadata in Visualization. In: IEEE (Hrsg.): Workshop on Visualization for Social Good (VIS4Good), New Orleans, 2021, S. 11-15.

Das Projekt GEOrg bei den SWM: Mit Freier Software zur konzernweiten Geodatenplattform

Nina Röckelein [1], Benedikt Seidl [1]

[1] Stadtwerke München GmbH (SWM)

Zusammenfassung: Mit freier Software aus dem QGIS-Universum bauen wir bei den Stadtwerken München eine konzernweite Geodatenplattform. Wir erzählen von unseren Erfahrungen bei der Integration verschiedener Anwendungen und der Etablierung von Open-Source Lösungen im Unternehmen.

Schlüsselwörter: QGIS, QWC2, PostGIS, Open Source

Wir berichten euch aus dem 2024 gestarteten Projekt „GEOrg“, welches zentrale Infrastrukturkomponenten für viele bereits selbstständig arbeitenden QGIS-Nutzer:innen im Unternehmen bereitstellen soll.

Ausgangslage

Bei den Stadtwerken München GmbH wird intensiv mit Geodaten gearbeitet. Dabei gibt es neben den zentral verwalteten Netzinfrastrukturdaten verschiedene andere Abteilungen, die Geodaten dezentral mit QGIS verwalten. Konkrete Beispiel für solche dezentralen Datenbestände sind die Daten aus der Erkundung von neuen Geothermie-Standorten oder aus der Überprüfung von Trinkwasserschutzgebieten.

Die Datenspeicherung erfolgte bisher bei vielen Anwender:innen auf geteilten Laufwerken. Dies war bisher die einfachste Lösung einer Zusammenarbeit, führt aber zu hohen Aufwänden. So müssen beispielsweise Bearbeitungskonflikte manuell vermieden werden und Berechtigungen sind auf Dateisystemebene zu verwalten.

Zusätzlich besteht bei fast allen QGIS-Anwender:innen auch Bedarf, ihre Projekte in einer einfachen Webansicht mit berechtigten (internen) Nutzer:innen zu teilen. Andere suchen Möglichkeiten zur mobilen Erfassung der Daten im Feld.

Unsere Lösung

Das Projekt „GEOrg“ soll die nötige Infrastruktur bereitstellen, um Datenverwaltung aus dem Dateisystem in die Datenbank zu heben. Die Anwender:innen behalten dabei die Hoheit über ihre Daten und können über das im Konzern etablierte Identitätsmanagement steuern, wer lesenden oder schreibenden Zugriff auf ihrer Datenbank bekommen soll.

Die zentrale Geodatenplattform besteht im Kern aus einem PostGIS-Datenbankcluster, einem S3-kompatibelen Storagesystem, sowie einer Instanz des QGIS Web Clients (QWC2) im Intranet. Daneben wurde eine eigene QFieldCloud Instanz aufgesetzt, um einigen Anwender:innen die Felderfassung von Daten mittels QField-App zu ermöglichen.

Integriert werden die verschiedenen Komponenten durch ein eigenentwickeltes QGIS-Plugin spezifisch für die SWM, welches die verschiedenen Komponenten automatisch in der Oberfläche von QGIS einbindet. So kann ein langwieriges Setup vermieden werden. Außerdem ermöglichen wir über die Plugin-Oberfläche das Teilen einer Projektdatei auf der internen QWC2-Instanz.

Das schlüsselbasierte Storage-System ist vorwiegend nötig, um die Veröffentlichung von QGIS-Projekten im QWC2 bereitzustellen. Perspektivisch wollen wir darüber auch die Synchronisierung von QGIS-Projektdateien zwischen mehreren Nutzer:innen ermöglichen.

Freie und Offene Software im Unternehmen

Durch den Aufbau der Datenplattform versprechen wir uns eine starke Vereinfachung in der Arbeit mit Geodaten bei den SWM. Mit dem Einsatz von QGIS und seinen Umfeldanwendungen setzen wir auf ein reifes Ökosystem mit einer aktiven Community. Dadurch sind wir nicht von einem einzelnen Hersteller abhängig und finden für jeden Anwendungsfall eine passende Lösung. Die Offenheit der Anwendungen erlaubt uns außerdem, eigene Anpassungen an der eingesetzten Software vorzunehmen. Diese reichen von kleinen Änderungen an der Benutzeroberfläche bis hin zur vollständigen Entwicklung eigener Features.

Der Einsatz von freien und quelloffenen Softwareprodukten ist bei den SWM im GIS-Bereich bereits gelebte Praxis. Neu ist in unserem Fall, dass wir verallgemeinerbare Eigenentwicklungen als Pull Request an die Community zurücktragen wollen. Damit beschreiten wir neue Wege im Unternehmen. Unsere erste erfolgte Einreichung dieser Art ist eine Weiterentwicklung des Druckdialogs für den QGIS Web Client. [1] Als die Anforderung dazu intern aufkam, war schnell klar, dass wir dieses Feature nicht nur für uns entwickeln. Die Neuentwicklung wird nun Teil des kommenden QWC2-Releases sein.

Das Einreichen von Anpassungen an das Ursprungsprojekt bietet uns mehrere Vorteile: Zum einen unterstützen wir so die QGIS-Community und tragen dazu bei, ein für uns wichtiges Open Source Projekt voranzubringen. Außerdem halten wir auch unseren Instandhaltungsaufwand für eigene Anpassungen an der Software gering. Daher wollen wir diesen Ansatz auch in Zukunft weiterverfolgen.

Wir freuen uns auf Vernetzung, wenn ihr an unseren Erfahrungen interessiert seid oder ähnlichen Themen arbeitet.

Kontakt zu den Autoren:

Benedikt Seidl
SWM Services GmbH
Seidl.Benedikt1@swm.de

Nina Röckelein
SWM Services GmbH
Roeckelein.Nina@swm.de

Literatur und Quellen:

[1] Pull Request im QWC2: https://github.com/qgis/qwc2/pull/340

Open Data und KI im Einsatz: Geodaten für alle, nicht nur für Profis?

Lisa Stubert, Klemens Maget

Im Rahmen der aktuellen Diskussion zur Nutzung von KI-Sprachmodellen untersuchen wir in einem Proof of Concept, wie KI im Zusammenspiel mit offenen Geodaten eingesetzt werden kann. Kann KI dabei unterstützen, frei verfügbare Daten besser aufzufinden, verständlicher zu machen, ihre Relevanz für spezifische Projekte zu bewerten oder ganz neue Ideen zu entwickeln? Oder eröffnet sie gar die Möglichkeit, neue Nutzendengruppen außerhalb der traditionellen Geodaten-Expert:innenkreise anzusprechen?

Für ein Forschungsprojekt die Folgen des Klimawandels im städtischen Raum verstehen?

Für die Arbeit im Stadtplanungsamt die Verkehrsinfrastruktur auf einer Karte präsentieren?

Für eine Kunstinstallation die Geometrien aller Berliner Bezirksregionen auf ein Poster drucken?

Für all diese und viele weitere Vorhaben benötigen wir... Geodaten! Das Land Berlin verfügt zum Glück über eine standardisierte Geodateninfrastruktur und über einen enormen Schatz an offen verfügbaren räumlichen Daten, die für eine Vielzahl von Personen und Vorhaben relevant sein können. Der Zugang zu Geodaten oder auch das Verständnis der Daten ist dabei jedoch Expert:innen mit Wissen über Geodatenstandards, Verwaltungsjargon und fachspezifischen Vorwissen vorbehalten.

Kann KI helfen diese Daten schneller zu finden und zu verstehen? Aus dieser Fragestellung heraus wurde der GeoExplorer entwickelt. Er ist ein Teil einer Machbarkeitsanalyse rund um Open Data und KI. Wie gut das funktioniert und was wir bisher daraus gelernt haben, stellen wir in diesem Impulsvortrag vor.

Als Open Data Informationsstelle Berlin, ein Projekt der gemeinnützigen Technologiestiftung Berlin erproben wir seit 2018 mit prototypischen Open Source Anwendungen neue Technologien und innovative Ansätze und zeigen das Potential offener Geodaten auf.



MOSIDI - Homogenisierung von offenen Daten für die kommunale Planung

Sebastian Meier[1], Qasem Safariallahkheili[1], Leonard Higi[1]

[1] Fachhochschule Potsdam

Zusammenfassung: Das Forschungsvorhaben InNoWest entwickelt mit dem System MOSIDI ein homogenisiertes Datenschema und benutzerfreundliche Interfaces, die es Kommunen ermöglichen, ohne technische Expertise datenbasierte Entscheidungen zu treffen und interkommunale Zusammenarbeit zu fördern. Dieses Paper zeigt, wie kommunale Akteur:innen von diesem System profitieren können.

Schlüsselwörter: Web GIS, Open Data, Öffentliche Verwaltung

Einleitung: Die zunehmende Digitalisierung eröffnet der kommunalen Verwaltung neue Möglichkeiten, Prozesse zu optimieren und daten- und damit evidenzbasierte Entscheidungen zu treffen. Gleichzeitig stehen viele Kommunen vor der Herausforderung, die Heterogenität und Komplexität der verfügbaren Datenquellen zu bewältigen. Unterschiedliche Datenformate, -strukturen und -auflösungen erschweren die Integration und Nutzung dieser Daten. Insbesondere kleine und mittlere Kommunen verfügen häufig nicht über die notwendigen Ressourcen, um diese Aufgabe zu meistern.

Das Forschungsvorhaben InNoWest adressiert diese Problematik durch die Entwicklung eines homogenisierten Datenschemas und zugehörigen APIs und Interfaces für kommunale Planungsszenarien, genannt *MOSIDI*, kurz für Modulare Open Source Infrastruktur für Dateninteraktion. Ziel ist es, eine flexible Datenbasis zu schaffen, die es Kommunen ermöglicht, ohne datenwissenschaftliche Expertise evidenzbasierte Entscheidungen zu treffen und interkommunale Zusammenarbeit zu fördern. Dieser konkrete Bedarf hat sich aus der nutzerzentrierten Arbeit mit regionalen Stakeholder:innen ergeben, die im Rahmen verschiedener Beteiligungsformate wie Workshops eingebunden wurden. Dieses Paper versteht sich dabei als Case Study mit einem Fokus auf die technische Umsetzung.

Open Data in der Anwendung: Die Anzahl offener Datensätze (Open Data) der öffentlichen Verwaltung nimmt kontinuierlich zu. Initiativen des Bundes wie GovData[1] der Föderalen IT-Kooperation (FITKO) oder das Geoportal[2] des Bundesamtes für Kartographie und Geodäsie (BKG) in Kooperation mit den Kommunen, Ländern und Behörden als Bereitsteller:innen, sowie europäische Bewegungen (wie z.B. INSPIRE[3] oder DCAT-AP[4]) fördern die Bereitstellung und Nutzung von offenen Verwaltungsdaten. Diese Entwicklungen bieten einer Vielzahl von Akteur:innen neue Möglichkeiten, auf umfangreiche und vielfältige Datenquellen zuzugreifen.

Für die Nutzung dieser Daten stehen bereits eine Vielzahl an offenen Werkzeugen zur Verfügung. Ein Beispiel für eine offenes Werkzeug, welches gemeinsam mit Verwaltugnsakteur:innen entwickelt wurde und geringe datentechnische Vorkenntnisse voraussetzt und damit einen barrierearmen Zugang zu Daten schafft ist das Masterportal[5]. Wie andere Werkzeuge in dieser Sparte, ist das Masterportal primär auf das Darstellen von und Interagieren mit einzelnen Geodatensätzen fokussiert, bzw. das Überlagern mehrerer Datensätze. Werden darüber hinaus komplexere Dateninteraktionen notwendig, müssen speziell angepasste Funktionen integriert werden. Diese speziellen Funktionen sind dann in der Regel auf für die Anwendung spezifisch ausgewählte Daten und Attribute ausgelegt. Die Hamburger Instanz des Masterportals stellt z.B. eine Funktion bereit, um Einwohnerzahlen für beliebige Auswahlpolygone zu berechnen.

Selbstverständlich gibt es auch bereits eine Vielzahl an offenen Werkzeugen, die komplexere Dateninteraktionen mit einem oder über mehrere Datensätze hinweg erlauben, wie z.B. QGIS[6]. Doch verlangen diese Anwendungen wesentlich mehr Expertenwissen im Datenbereich. Dies stellt für viele kleine und mittlere Kommunen eine große Hürde dar, weil es häufig keine spezialisierten Stellen für diese Aufgaben gibt.

Überblick: MOSIDI zielt darauf ab, genau diese Lücke zwischen benutzerfreundlichen Werkzeugen für Nutzer:innen ohne datenwissenschaftliches Expertenwissen und dem Bedarf komplexer Dateninteraktionen zu schließen.

Um diese Formen von Dateninteraktionen zu ermöglichen ist es nicht ausreichend Analyse- und Visualisierungsmethoden mit entsprechenden Interfaces auf bestehende Datensätze anzuwenden, sondern eine strukturiere Homogenisierung von Daten ist notwendig. Wenn wir die deutsche Open Data Landschaft genauer betrachten, wird deutlich, dass wir es selbst innerhalb desselben Themas (z.B. Bevölkerungsdaten) mit diversen Formaten, Strukturen und Auflösungen zu tun haben, was eine direkte Integration oft erschwert oder unmöglich macht. Nur durch eine Vereinheitlichung können Daten aus verschiedenen Quellen effektiv zusammengeführt und analysiert werden. Die Homogenisierung findet entsprechend auf drei Ebenen statt: räumlich, zeitlich und auf Attributsebene (siehe Tabelle 1).

Raum-Referenz	**Zeit**	**Attribut**	**Wert**
Foreign Key: Geometry ID	Interval: yearly Time: 2019	z.B. Einwohnerzahl, Anzahl Krankenhäuser, etc.	z.B. 287101

Beispiel für einen Datensatz in der Indikatorentabelle (siehe unten).

Das vorgestellte System nutzt einen Technologie-Stack, der auf bewährten Open-Source-Technologien basiert. Die Datenbank als Basis ist PostgreSQL, einer relationalen Datenbank, die in Kombination mit PostGIS, einer Erweiterung für geografische Informationssysteme (GIS), eine effiziente Verwaltung und Analyse räumlicher Daten ermöglicht. Für den Import von Daten sowie die Bereitstellung einer flexiblen Daten-API wird Python eingesetzt. Das Frontend des Systems wurde mit Vue.js und MapLibre realisiert.

Homogenisierung: Alle importierten Daten werden in ein PostgreSQL / PostGIS Datenbank abgelegt. Das Schema besteht aus drei Tabellentypen:

Metadaten: Diese Tabelle enthält nach dem Standard DCAT-AP[4] notwendige Metadaten zu allen importierten Daten sowie import-relevante Informationen wie den Zeitpunkt des Imports und den Datenlieferanten. Sie ermöglicht es, die Herkunft und Qualität der Daten nachzuvollziehen, was besonders bei der Nutzung unterschiedlicher Quellen entscheidend ist.

Geometrien: In diesen Tabellen sind, die vom BKG bzw. anderen Behörden offiziell herausgegebenen Geometrien enthalten, wie etwa Gemeindegrenzen oder andere administrative Einheiten. Diese Geometrien dienen als räumliche Referenz, um die Daten eindeutig zuzuordnen und konsistent zu halten. Die Geometrien sind aufgebrochen in drei Tabellen, für Polygone, Linien und Punkte, um eine höhere Performanz zu erreichen.

Indikatoren: Alle importierten Daten werden in ihre grundlegenden Bestandteile zerlegt und nach Referenz-Geometrie, Zeit und Indikator organisiert. Diese Strukturierung gewährleistet eine hohe Flexibilität bei der Analyse und Kombination der Daten, da sie eine klare Zuordnung und Vergleichbarkeit ermöglicht.

Die Geometrie- und Indikatorentabellen sind mit einer Reihe an spezifischen Indizes optimiert, um schnelle Suchen, Filter und Berechnungen zu ermöglichen.

Der Import und die regelmäßige Aktualisierung der Daten werden durch speziell entwickelte Python-Skripte durchgeführt. Diese Skripte laden die Daten in festgelegten Intervallen aus den verschiedenen Quellsystemen und passen sie an die Anforderungen des MOSIDI-Datenmodells an. Dabei werden unterschiedlichste Formate, Auflösungen und Strukturen verarbeitet, um die Daten in ein homogenisiertes Schema zu überführen. Aufgrund der großen Heterogenität der Datenquellen ist dieser Prozess bisher nicht automatisiert und erfordert manuelle Anpassungen und Wartung der Skripte. Wie unter den Limitationen beschrieben, gilt dieser manuelle Aufwand als größter technischer Nachteil des Systems. Dennoch ist dieser Ansatz aktuell unverzichtbar, um die Konsistenz und Qualität der Daten zu gewährleisten und den Nutzer:innen eine performante und zuverlässige Datenbasis bereitzustellen.

An Stellen, an denen keine offiziellen BKG-Geometrien verwendet werden können, wie beispielsweise bei fein aufgelösten Wetterdaten des Deutschen Wetterdiensts (DWD), aggregieren wir die Daten derzeit auf die Ebene von Gemeinden. Dies bedeutet, dass beispielsweise die minimalen, maximalen und durchschnittlichen Niederschlagswerte für ein bestimmtes Gemeindegebiet berechnet und bereitgestellt werden. Dieser Schritt stellt eine bewusste Entscheidung dar, die räumliche Granularität zu verringern, um ein einfach zu nutzendes und praktisches System zu gewährleisten. Durch diese Aggregation wird eine bessere Handhabung und Analyse der Daten ermöglicht, insbesondere bei sehr detaillierten Informationen, die andernfalls schwer zu verarbeiten wären.

Dateninteraktion: Im homogenisierten Schema sind verschiedenste Geometrietypen und -granularitäten enthalten. Um die Interkation im Frontend für die Nutzer:innen zu vereinfachen, haben wir eine Interaktionsmatrix entwickelt (siehe Tabelle 2), welche während der Auswahl von Daten zu tragen kommt, um Nutzer:innen nur unseren Erachtens nach sinnvolle Kombinationen von Daten anzubieten und damit die Interaktion sinnvoll zu leiten.

	Polygon	Punkt	Linie
Polygon	Freie Berechnung	Zählen & Zusammenfassungen	Streckensumme & Zusammenfassung
Punkt		Nur für identische Punkte, dann freie Berechnung	Zählen & Zusammenfassung mit Buffer
Linie			Nur für identische Linien, dann freie Berechnung

Tabelle 2: Interaktionsmatrix (der aktuelle Prototyp fokussiert bisher nur die Polygon-Interaktionen, weitere Interaktionen folgen in den nächsten Versionen). Zusammenfassung steht für Maximum, Minimum, Durchschnitt, etc.

Ein Sonderfall sind nicht deckungsgleiche Geometrien, welche sich durch verschiedene Ordnungsprinzipien ergeben können, oder durch verschiedene Granularitäten, wie z.B. die Kombination von Daten auf Bundes- und Gemeindeebene. In diesem Fall bedienen wir uns den Erkenntnissen des CoGran[7] Projektes, welches die Interpolation von Daten verschiedener Polygone, basierend auf Einwohnerzahlen erlaubt.

Basierend auf den oben beschriebenen Dateninteraktionen, wird den Nutzer:innen ein relativ einfaches Interface angeboten, über welches Daten angezeigt, kombiniert, analysiert und visualisiert wer-

den können. Während die meisten Funktionen gängigen Web-GIS entsprechen, wollen wir zwei Funktionen hervorheben, die ihren Mehrwert speziell durch die Homogenisierung entfalten:

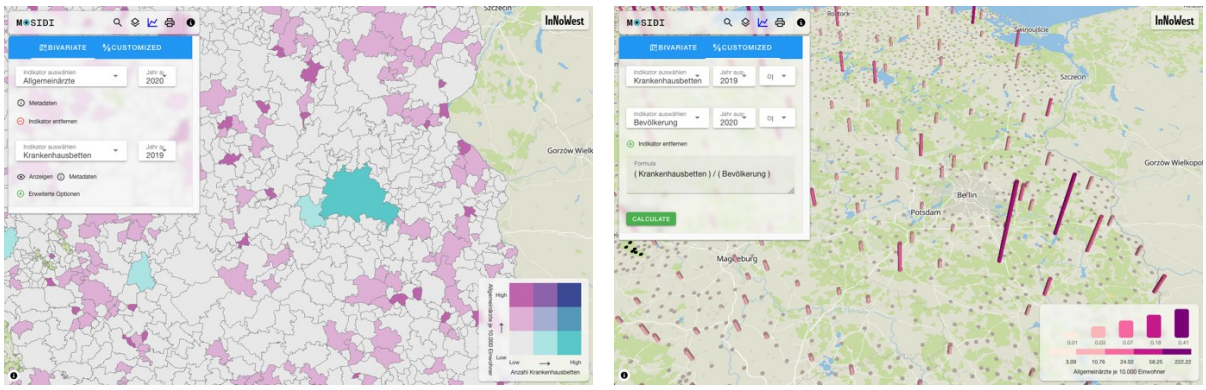


Abb. 1 (links) Multivariate Visualisierung und 2 (rechts) Formelbasierte Visualisierung

- *Multivariate Visualisierungen*

Durch ein benutzerfreundliches Interface können homogenisierte Datenquellen miteinander kombiniert werden(siehe Abb. 1), um multivariate Datenvisualisierungen zu erstellen. Das Interface ermöglicht es, verschiedene Datensätze zu selektieren und zu verbinden, indem unterschiedliche Variablen miteinander in Beziehung gesetzt werden. Diese Kombination führt zu komplexen Analysen, bei denen mehrere Dimensionen gleichzeitig betrachtet werden können. Auf diese Weise entstehen multivariate Visualisierungen, die es ermöglichen, Muster und Zusammenhänge zwischen verschiedenen Datenpunkten zu erkennen.

- *Formelbasierte Visualisierengen*

Das einfache Kombinieren von Daten stößt jedoch an Grenzen, insbesondere wenn es darum geht, komplexe Beziehungen zwischen unterschiedlichen Datenquellen zu modellieren. Um diese Einschränkungen zu überwinden, wurde ein Formel-Editor entwickelt (siehe Abb. 2), der es ermöglicht, verschiedene Indikatoren aus unterschiedlichen Datenquellen mithilfe mathematischer Formeln zu einem neuen, berechneten Indikator zu kombinieren. Dieser Ansatz erlaubt eine präzisere und flexiblere Modellierung von Zusammenhängen, da Nutzer:innen individuell definierte Formeln zur Berechnung von Indikatoren anwenden können. Die resultierenden Indikatoren können anschließend visualisiert werden.

Aus datentechnischer Perspektive stellen diese Funktionen keine grundlegenden Neuerungen dar, da ähnliche Prozesse bereits mit den richtigen Fachkenntnissen und Werkzeugen, wie zum Beispiel in QGIS, mit wenigen Klicks oder wenigen Zeilen Code umgesetzt werden können. Die wahre Innovation dieses Systems liegt jedoch in der Benutzerfreundlichkeit und Zugänglichkeit. Dank der bereits homogenisierten Daten können Nutzer:innen ohne tiefgehende Fachkenntnisse komplexe Datenkombinationen und Berechnungen über ein einfaches Interface durchführen. Dies senkt die Einstiegshürde und ermöglicht es auch weniger technisch versierten Anwender:innen, leistungsstarke Datenanalysen und Visualisierungen zu erstellen, die zuvor nur Fachleuten mit entsprechender Expertise vorbehalten waren. Das System erleichtert so die Nutzung komplexer Geodaten und fördert eine breitere Anwendung in verschiedenen Fachbereichen.

Limitierungen: Einige unserer Limitierungen sind bewusst gewählt und dienen der Fokussierung auf spezifische Bedarfe der kommunalen Verwaltung. Denn wir wollen kein umfängliches Web-GIS entwickeln und damit in Konkurrenz mit Systemen wie dem Masterportal treten, sondern unsere Entwicklungen als Ergänzung der bereits existierenden Technologien verstehen.

Eine der größten Limitierungen unseres Ansatzes liegt in der Homogenisierung von Daten, da diese ein kontinuierliches und manuelles Anpassen an Veränderungen der Daten erfordert. Neue Dateninte-grationen müssen manuell angelegt werden und das System kann entsprechend nicht automatisch auf Datenänderungen reagieren.

Aus der Perspektive von Linked Open Data ist auch das lokale Duplizieren bzw. Transformieren von Daten eine mögliche Limitierung. Diese Maßnahme ist jedoch notwendig, um eine performante Verar-beitung der Daten zu gewährleisten und komplexe Dateninteraktionen zu ermöglichen.

Nächste Schritte: In zukünftigen Arbeiten sollen mehrere Aspekte weiterentwickelt und erforscht wer-den, um die Funktionalität und Nutzbarkeit des Systems zu verbessern. Ein zentraler Ansatzpunkt ist die Integration von erklärenden Komponenten zu komplexeren Berechnungsverfahren, etwa zur Inter-polation. Dies soll Nutzer:innen unterstützen, die entstehenden Ergebnisse besser zu verstehen und Ungenauigkeiten in den Analysen nachvollziehen zu können. Entsprechende Lösungen sollen dann auch durch Nutzerstudien weiter evaluiert werden.

Ein weiterer Fokus liegt auf der Entwicklung von Exportfunktionen, die spezifisch an den Bedarfen der kommunalen Verwaltung ausgerichtet sind, wie beispielsweise Berichts- und Dokumentationszwecke. Solche Exportmöglichkeiten könnten die direkte Übertragbarkeit der gewonnenen Erkenntnisse aus den Daten in bestehende Verwaltungskontexte erleichtern.

Zudem sollen die Import-Pipelines weiter vereinfacht und generalisiert werden, um eine effizientere und nutzerfreundlichere Integration neuer Datenquellen zu ermöglichen. Dies ist besonders relevant, um das System nachhaltig und langfristig nutz- und wartbar zu machen.

Ein besonders vielversprechender Bereich für zukünftige Forschung ist die Erprobung der Integration des Systems in größere bestehende Lösungen wie das Masterportal, Civitas Core[8] oder die Daten-plattform des DKSR[9]. Die Schnittstellen zu solchen Systemen könnten neue Möglichkeiten für Inter-operabilität und Synergien schaffen, die weit über die derzeitigen Anwendungsmöglichkeiten hinaus-gehen.

Fazit: Mit MOSIDI wurde ein innovatives System geschaffen, das die kommunale Planung durch eine standardisierte und benutzerfreundliche Datenplattform nachhaltig unterstützt. Die ausschließliche Nutzung offener Technologien und der partizipative Entwicklungsprozess stellen sicher, dass die Lö-sung nicht nur kosteneffizient, sondern auch flexibel und zukunftsfähig ist. Das Projekt leistet somit ei-nen wichtigen Beitrag zur Digitalisierung und Modernisierung der kommunalen Verwaltung.

Der aktuelle Prototyp kann hier eingesehen werden: https://tv4-geo.innowest-brandenburg.de

Acknowledgements: Der Bedarf für das oben beschriebene Werkzeug haben wir im Rahmen von Workshops mit Stakeholder:innen der öffentlichen Verwaltung in Nord-West Brandenburg identifiziert. Wir möchten uns an dieser Stelle herzlich bei allen Teilnehmer:innen bedanken.

Das Verbundvorhaben InNoWest wird im Rahmen der Bund-Länder-Initiative „Innovative Hochschule“ gefördert (Förderkennzeichen 13IHS230A).

Kontakt zum Autor:

Prof. Dr. Sebastian Meier
Interaction Design Lab
Fachhochschule Potsdam
Kiepenheuerallee 5, 14469 Potsdam
sebastian.meier@fh-potsdam.de

Literatur und Quellen

[1] https://www.govdata.de

[2] https://www.geoportal.de

[3] https://inspire-geoportal.ec.europa.eu/srv/ger/catalog.search#/home

[4] https://ec.europa.eu/isa2/solutions/dcat-application-profile-data-portals-europe_en/

[5] https://www.masterportal.org/

[6] https://www.qgis.org/

[7] https://github.com/funkeinteraktiv/cogran

[8] https://www.civitasconnect.digital/civitas-core/

[9] https://www.dksr.city/

Der Digitale Zwilling - so wertvoll wie eine Karte im Maßstab 1:1

Matthias Daues

Digitale Zwillinge sind heute in aller Munde: Ein Allheilmittel für Probleme der räumlichen Modellierung. Aber stimmt das? Übersehen wir da nicht im Rausch des Machbaren die Frage nach dem Sinnvollen?

Ein "Rant" gegen die "Landkarte im Maßstab 1:1".

Digitale Zwillinge sind heute in aller Munde: Ein Allheilmittel für die Probleme der räumlichen Modellierung, der Datenverfügbarkeit, der Datendarstellung, der Erzeugung von Erkenntnissen über "das Wie vom Wo (und Wann)" aus der Ferne, der Nähe, von oben, unten und der Seite. Also etwas, ohne das wir eigentlich nicht Denken, Planen und Entscheiden können. Ach was, nicht Leben können! So der Eindruck, der beim Lesen einer Geo-gefärbten LinkedIn-Timeline entstehen könnte.

Aber mal ehrlich, stimmt das eigentlich? Der zwanghafte Trieb nach mehr Realitätstreue beschert uns - abgesehen vom Wettrüsten in der Unterhaltungselektronik - eben auch immer mehr vollmundige Versprechungen über Smart Cities, Smart Buildings, BIM, Management jeglicher Form der Landnutzung oder -beobachtung; und das alles sei ab sofort wie durch Magie für alle verfügbar und immer wahr und immer aktuell und immer richtig.

In der von konkretem Detailwissen unbelasteten Höhenluft von Managementmeetings erzeugen diese Versprechen und Ahnungen ein erleichtertes Aufatmen. Endlich kann ein Tool angeschafft werden, mit dem lästige Probleme ganz einfach gelöst werden können. Am Besten mit KI drin, dann wird es noch besser.

Aber wie ist es denn in Wirklichkeit? Übersehen wir im Rausch des Machbaren nicht gern die Frage nach dem Sinnvollen? Neigen wir nicht oft genug dazu, zu vergessen, dass die Herausforderung beim Modellieren von Informationszusammenhängen darin besteht, nur das Wichtige zu behalten und das Unwichtige zu vergessen; die Kunst also vor allem eine des Weglassens ist? Der "Digitale Zwilling" erscheint da wie das Versprechen einer Landkarte im Maßstab 1:1. Auch wenn alle das haben wollten, würde es doch niemandem nützen (mal abgesehen davon, dass wir sie ja schon haben).

Lasst uns also mehr Reduktion und Abstraktion wagen. Lasst uns mutig weglassen. Und lasst uns aufhören, etwas Zwilling zu nennen, was keiner ist und vor allem keiner sein sollte!

Routing, aber mehr explorativ statt automatisch

Katharina Rasch

Bei langen Zugreisen quer durch Europa ist auch immer ein bisschen der Weg das Ziel. Beim Planen solcher Reisen ist für mich der von Routingtools vorgeschlagene Reiseplan oft nur ein Ausgangspunkt. Ich möchte easy den Plan anpassen können, Verbindungen hin und her schieben, schauen, wie sich das auswirkt; vielleicht hier eine Übernachtung einlegen, oder dort; vielleicht einen Umweg über eine Stadt, in der ich noch nie war. Genau dafür baue ich eine UI und von der möchte ich euch gern erzählen.

Beim Planen langer Zugreisen quer durch Europa (z.B. Hamburg -> Sizilien) sind oft andere Kriterien wichtiger als die reine Reisezeit oder die Anzahl der Umstiege. Ich möchte eine bequeme Verbindung, mit genug Puffer, insbesondere bei Anbieterwechsel. Wenn eine Übernachtung unterwegs nötig ist, dann gern irgendwo wo ich noch nie war (aber dann gern dort mit Ankunft nicht so spät, damit ich noch Zeit habe, mich da umzuschauen). Aha, ein Reiseweg (nicht unbedingt der schnellste) führt über Zürich? Da kenn' ich jemanden, wo ich übernachten kann! Ach, ich muss eh in Rom umsteigen? Mittagspause auf der Piazza, das klingt super! Kurz gesagt: bei solchen Reisen ist auch immer ein bisschen der Weg das Ziel.

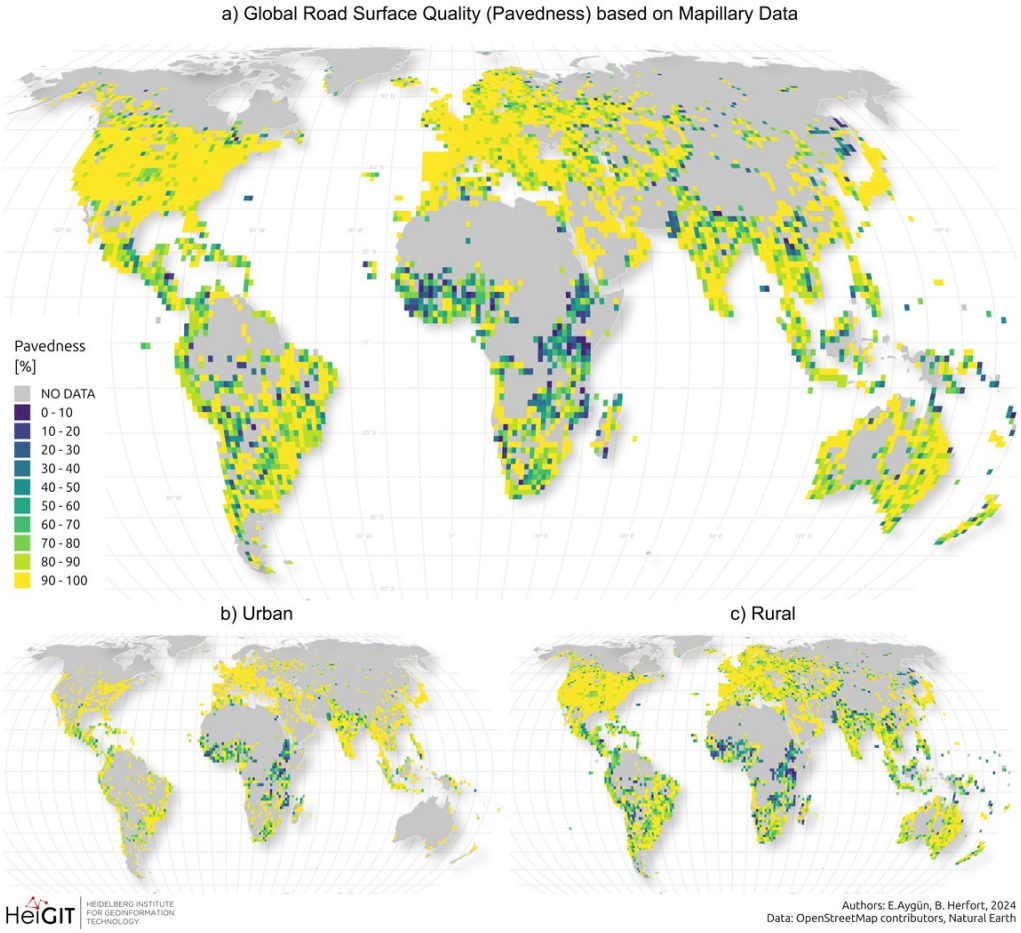
All die genannten Kriterien sind sehr persönlich und ergeben sich teilweise erst während des Planungsprozesses. Existierende Routingtools erlauben mir zwar, meine Suchanfrage zu ändern wenn mir noch was eingefallen ist. Aber was ich eigentlich will ist, mit den Routingvorschlägen zu interagieren, sie zu verändern wie es für mich passt. Verbindungen hin und her schieben zu können, per drag&drop ausprobieren, was wäre, wenn ich diese Mittagspause in Rom einlegen würde, wie würde sich der Rest der Reise verschieben? Auf einer Karte explorieren zu können, welche Reisewege es überhaupt gibt und mir einfach mal 'nen Umweg reinklicken, weil ich hab' Bock und ich war da noch nie.

In meinem Projekt Trans-Europa-Planer (gefördert vom prototypefund) baue ich eine solches User Interface und untersuche verschiedene UX-Mechanismen zum interaktiven Erforschen, Hin- und Herschieben, Vergleichen von Reiserouten. Und ich untersuche, inwiefern ein solches Interface den Nutzer*innen tatsächlich dabei hilft, ihre eigenen, auf sie persönlich zugeschnittenen Zugreisen zu planen. Was dabei herausgekommen ist, darüber möchte ich gern in meinem Vortrag berichten.

Anreicherung von Straßendaten mittels Deep-Learning-Methoden und Mapillary Bildern

Benjamin Herfort, Sukanya Randhawa

Wir haben einen globalen Datensatz zu Straßeneigenschaften ("befestigt" oder "unbefestigt") mithilfe von KI-Methoden erstellt, basierend auf 105 Millionen Bildern von Mapillary . Unser Ansatz kombiniert SWIN-Transformer-Vorhersagen des Straßenbelags mit einer CLIP-Filterung minderwertiger Bilder. Der Datensatz deckt 36 % des weltweiten Straßennetzes ab, wobei die Abdeckung in Städten höher ist. Er kann in der Stadtplanung, im Katastrophenmanagement, in der Logistik genutzt werden.



Wir haben einen offenen Datensatz mit globaler Abdeckung über Straßeneigenschaften (“befestigt” oder “unbefestigt”) veröffentlicht, der aus 105 Millionen Bildern von der weltweit größten crowdsourcing-basierten Street-View-Plattform Mapillary abgeleitet wurde. Wir nutzen einen hybriden Deep-Learning-Ansatz, der SWIN-Transformer nutzt, um die Straßenbedeckung vorherzusagen und mit einer CLIP-und-DL-Segmentierung zur Filterung von Bildern niedriger Qualität kombiniert ist. Die Vorhersagen zur Straßendecke sind mit OpenStreetMap (OSM) Straßendaten abgeglichen und zusammengeführt.

Unsere Ergebnisse liefern Einblicke in die räumliche Verteilung der Mapillary-Daten und der Straßenbedeckung auf Kontinent- und Länderebene. Im Vergleich zu OSM, erweitert unser Datensatz die Ver-

fügbarkeit globaler Informationen über Straßenbedeckung um über 3 Millionen Kilometer und repräsentiert nun etwa 36 % der Gesamtlänge des weltweiten Straßennetzes.

Die meisten Regionen zeigen eine mäßige bis hohe Abdeckung mit befestigten Straßen (60-80 %), aber in bestimmten Gebieten Afrikas und Asiens wurden signifikant niedrigere Werte festgestellt. Stadtgebiete tendieren zu einer nahezu vollständigen “Pavedness”, während ländliche Regionen eine größere Variabilität aufweisen.

Die Validierung des Modells gegenüber OSM-Daten erreichte F1-Scores für gepflasterte Straßen zwischen 91 und 97 % über die Kontinente hinweg. Aufbauend auf der Arbeit von Mapillary und deren Beitragenden sowie der Anreicherung von OSM-Straßenattributen bietet unsere Arbeit wertvolle Erkenntnisse für Anwender in der Stadtplanung, im Katastrophenmanagement, in der Logistikoptimierung und unterstützt verschiedene Nachhaltigkeitsziele (SDGs).

In unserer Analyse haben wir bisher ausschließlich Daten von Mapillary genutzt. Je nach Anwendungsregion könnte es interessant sein, in der Zukunft weitere Street-View-Platformen zu betrachten. Gerade Alternativen wie Panoramax (https://panoramax.fr/) haben in den vergangenen Monaten vorwiegend in Europa an Beliebtheit gewonnen.

G3W-SUITE: das OS-Framework für die Veröffentlichung und Verwaltung von QGIS-Projekte

Walter Lorenzetti, Antonello Andrea

G3W-SUITE ist eine modulare Client-Server-Anwendung (basierend auf QGIS-Server) für die Verwaltung und Veröffentlichung von interaktiven kartographischen QGIS-Projekten verschiedener Art auf eine völlig unabhängige, einfache und schnelle Weise.

Die Anwendung, die auf GitHub unter der Mozilla Public Licence 2.0 veröffentlicht wurde, ist mit den QGIS LTR Versionen kompatibel und basiert auf einer starken Integration mit der QGIS API.

G3W-SUITE ist eine modulare Client-Server-Anwendung (basierend auf QGIS-Server) für die Verwaltung und Veröffentlichung von interaktiven kartographischen QGIS-Projekten verschiedener Art auf eine völlig unabhängige, einfache und schnelle Weise.

Der Zugriff auf die Verwaltung, die Konsultation der Projekte, die Editierfunktionen und die Nutzung der verschiedenen Module basieren auf einem hierarchischen System von Benutzerprofilen, die bearbeitet und modifiziert werden können.

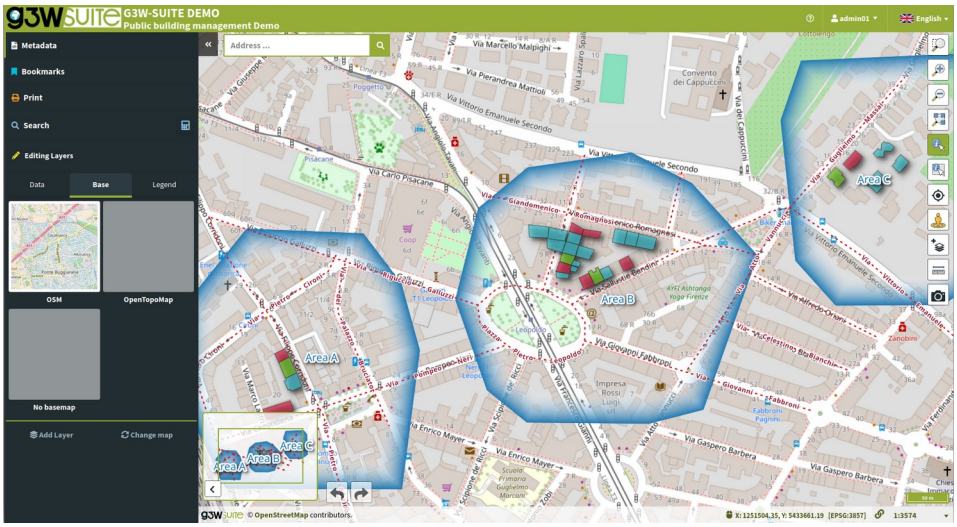
Die Suite setzt sich aus zwei Hauptkomponenten zusammen: G3W-ADMIN (basierend auf Django und Python) als Webadministrations-schnittstelle und G3W-CLIENT (basierend auf OpenLayer und Vue) als kartographischer Client, die über REST-APIs miteinander kommunizieren.

Die Anwendung, die auf GitHub unter der Mozilla Public Licence 2.0 veröffentlicht wurde, ist mit den QGIS LTR Versionen kompatibel und basiert auf einer starken Integration mit der QGIS API.

Diese Präsentation wird einen kurzen Überblick über die Geschichte der Anwendung und Einblicke in die wichtigsten Projektentwicklungen des letzten Jahres geben.

Die Entwicklungen betrafen sowohl die Verwaltungs- und Managementkomponente der WebGis-Dienste, als auch die Aspekte der Interaktion mit Webkarten und deren Inhalten, sowie die Aspekte und Funktionen im Zusammenhang mit dem Online-Editing mittels der Integration mit der QGIS API.

Eine spezifische Entwicklung betrifft die Integration mit der QGIS Processing API, um die Analysemodelle, die in QGIS über den ModelDesigner erstellt wurden, in einer Webumgebung auszuführen.



Der Wuppertaler Weg vom Geoportal zum Digitalen Zwilling

Stefan Sander

Die Stadt Wuppertal realisiert in einem bis Ende 2026 laufenden Förderprojekt einen Urbanen Digitalen Zwilling (UDZ), den "DigiTal Zwilling". In diesem Vortrag werden sie softwaretechnischen Herausforderungen und Lösungsansätze bei der Entwicklung der Anwendungskomponenten zum DigiTal Zwilling vorgestellt. Insbesondere entsteht eine generische Anwendung, die den Erwartungen an einen UDZ gerecht werden soll, aber auch die klassischen Anwendungsfälle eines kommunalen Geoportals bedienen soll.

Der DigiTal Zwilling ist ein „System of Systems“. Insbesondere umfasst er eine Reihe an Open Source Web-Applikationen, die den DigiTal Zwilling für die Nutzerinnen und Nutzer erschließen. Diese Applikationen werden unter https://github.com/digital-twin-wuppertal-live projektbegleitend publiziert. Das Flaggschiff dieser Applikationen wird eine generische Anwendung sein, in der man sich seine Arbeitsumgebung aus allen Daten, Diensten und Funktionen des DigiTal Zwillings frei zusammenstellen kann. Diese Anwendung wird das bisher geführte Wuppertaler Geoportal ablösen, sie muss also auch die z. T. gesetzlichen Anforderungen an die Publikation von OGC-konformen Web-Services erfüllen. Darüber hinaus werden einige Anwendungen für breitere Themenkomplexe (Fachzwillinge) und eine Vielzahl themenspezifischer Applikationen (Teilzwillinge) bereitgestellt. Im Vortrag werden die folgenden softwaretechnischen Aspekte thematisiert:

- Der DigiTal Zwilling als "System of Systems"
- Aufbau eines Monorepos und Migration von Legacy-Anwendungen in das Monorepo
- Nahtlose Integration verschiedener Mapping Engines (CesiumJS, Leaflet, MapLibre) in einer Applikation
- Integration von Vektor-Layern und -Styles über WMS Capabilities
- Innovative Ideen für zentrale GUI-Elemente (Gazetteer, Laden von Ressourcen, Layerbuttons etc.)
- Hilfe-Overlay als neuer Weg zur Unterstützung der Anwender/innen

Flexibles Open Source Routing mit Valhalla

Christian Beiwinkel

Valhalla ist eine vielseitige Routing Engine für das OSM Ökosystem, die vor allem durch ihre Flexibilität und RAM-Effizienz glänzt. Dieser Vortrag setzt sich zusammen aus einer allgemeinen Einführung in Valhalla sowie Updates zu neuen und verbesserten Features.

Valhalla hat als jüngeres Projekt in der OSM-Routing Welt vor allem durch sein flexibles Routing Aufmerksamkeit auf sich gezogen. Neben der Auswahl mehrerer Fahrzeugprofile ist vor allem die lange Liste an Parametern, mit denen Anfragen einzelne Anfragen zusätzlich angepasst werden können, ein Alleinstellungsmerkmal.

In diesem Vortrag werden einige grundlegende Konzepte, die Valhalla besonders machen, erläutert. Zudem werden einige Möglichkeiten aufgezeigt, wie Valhalla dank seiner Flexibilität an besondere Bedürfnisse angepasst werden kann.

Zuletzt werden neue und verbesserte Features vorgestellt,

Versiegelungsanalyse zur bioklimatischen Bewertung von Stadtgebieten

Markus Metz, Lina Krisztian, Anika Weinmann, Victoria-Leandra Brunn, Markus Eichhorn

1 mundialis GmbH & Co. KG

Zusammenfassung

Für eine bioklimatische Bewertung des Stadtgebietes Düsseldorf wurde der Grad der Versiegelung erfasst, wobei verschiedene Arten von vollversiegelten, teilversiegelten und unversiegelten Flächen unterschieden wurden. Zusätzlich wurde hohe Vegetation über diesen Flächen erfasst. Für diese Klassifikation wurde ein neuronales Netz (U-Net) trainiert und auf die Jahre 2021, 2022 und 2023 angewendet.

Beschreibung

Im Rahmen einer bioklimatischen Bewertung aller Siedlungs- und Verkehrsflächen sowie Grün- und Freiflächen im Stadtgebiet Düsseldorf wurde der Grad der Versiegelung erfasst. Hierbei sollten verschiedene Arten von vollversiegelten, teilversiegelten und unversiegelten Flächen erfasst werden. Zusätzlich wurde hohe Vegetation über den verschiedenen Oberflächenklassen bestimmt. Für die Umsetzung wurde eine automatisierte, auf Methoden der Künstlichen Intelligenz (KI) gestützte

Erfassung und Klassifizierung entwickelt. Konkret wurden zwei U-Nets (CNNs) entwickelt und trainiert, einmal für die Oberflächen und einmal für hohe Vegetation. Diese Ergebnisse wurden anschließend kombiniert, um Informationen über den Versiegelungsgrad, hohe Vegetation und die Oberflächenklasse unter dieser hohen Vegetation zu erhalten. Für diese Analysen wurden digitiale Orthophotos, digitale Oberflächenmodelle und digitale Geländemodelle als Eingangsdaten verwendet. Um Veränderungen im Grad der Versiegelung abschätzen zu können, wurden diese Analysen für die Jahre 2021, 2022 und 2023 durchgeführt und miteinander verglichen.

Kontakt zum Autor:

Dr. Markus Metz
mundialis GmbH & Co. KG
Kölnstraße 99, 53111 Bonn
+49 (0)228 / 38 75 80 80
metz@mundialis.de

Historische Karten mit QGIS georeferenzieren

Niklas Alt, Johannes Wagner

Dank umfangreicher Digitalisierungskampagnen steht historisch Interessierten ein enormer Schatz an kartografischen Produkten aus mehr als zwei Jahrhunderten zur Verfügung. Mit dem QGIS-Georefe-rencer können diese einfach und niedrigschwellig für die Nutzung im beliebten Geoinformationssystem aufbereitet werden. Wir demonstrieren die wichtigsten Vorgehensweisen, weisen auf Fallstricke hin und geben wertvolle Produktivitätstipps für die Referenzierung.

Der Vortrag bietet anhand unterschiedlicher, historischer Karten eine Einführung in die Georeferenzie-rung mit QGIS. Von der großmaßstäblichen Katasterkarte, über die verbreiteten Topografischen Kar-ten im Maßstab 1:25.000 (sog. "Meßtischblätter") bis hin zu den kleinmaßstäblichen Übersichtskarten. Jeder Maßstab bringt dabei seine eigenen Herausforderungen mit sich.

Alle Georeferenzierungen basieren auf dem Prinzip die Pixel-Koordinaten der historischen Karte mit Kontrollpunkten zu verknüpfen, um so historische Karte und moderne Geodaten in Beziehung zuein-ander setzen. Dies kann über die Identifizierung anhand von Gebäuden, Straßen, Kanälen u.v.m. er-folgen. Die so gesetzten Punkte (Ground control points, GCP) kommen jedoch mit Fallstricken: die kartografischen Prinzipien der Generealsieriung und Verdrängung bleiben hierbei häufig unberücksich-tigt. Als Alternative bietet sich für einige Karten die Nutzung des aufgedruckten Koordinatengitters an. Auch für diese Methode braucht es ein gewisses Maß an Hintergrundwissen, so sind z.B. die histori-schen geografischen

Koordinaten nicht mit modernen geografischen Koordinaten gleichzusetzen.

Je nach Alter und Zustand der Karte, sowie der zur Verfügung stehenden Zeit, kommen verschiedene Algorithmen zur Georeferenzierung in Frage (Polynominal 1-3, Thin Plate Spline u.a.). Welcher Algo-rithmus gewählt wird hängt davon ab, wie viele Kontrollpunkte gesetzt wurden, in welchem Zustand die Karte ist (Falzen etc.) und ob etwaige Messfehler ausgeglichen werden sollen.

Zuletzt soll es auch um Möglichkeiten gehen den händischen Prozess der Referenzierung mittels Sta-pelverarbeitung zu beschleunigen und reproduzierbar zu machen.

Inhalte:

- Nutzung manueller Ground Control Points (GCP)
- Nutzung von historischen Koordinatengittern
- Tipps und Kniffe
- Was eignet sich zur Referenzierung?
- Bedeutung von (historischen) Koordinatenreferenzsystem
- Möglichkeiten der Automatisierung im Batch-Verfahren

Masterportal - Liegenschaftsauskunft ONLIKA 4.0 mit Keycloak und BundID

Laura Meierkort, Erik Jost

grit GmbH

Zusammenfassung

Im Rahmen dieses Vortrags wird die „Online-Auskunft für die Daten des Liegenschaftskatasters“ Thüringen - kurz ONLIKA 4.0 - vorgestellt. Das automatisierte Abrufverfahren wurde vom Thüringer Landesamt für Bodenmanagement und Geoinformation (TLBG) in Zusammenarbeit mit grit entwickelt. Das Rollen- und Rechtekonzept wird durch die Authentifizierung via BundID und Keycloak abgebildet. Basis für die ONIKA Oberfläche ist die Open Source-Lösung Masterportal.

Schlüsselwörter: Masterportal, BundID, Keycloak, Liegenschaftskataster, Kataster, ALKIS, GeoInfoDok 7.1.2, Produktproxy, DHK, NAS, eSiegel, Service Konto, SAML, OpenID-Connect

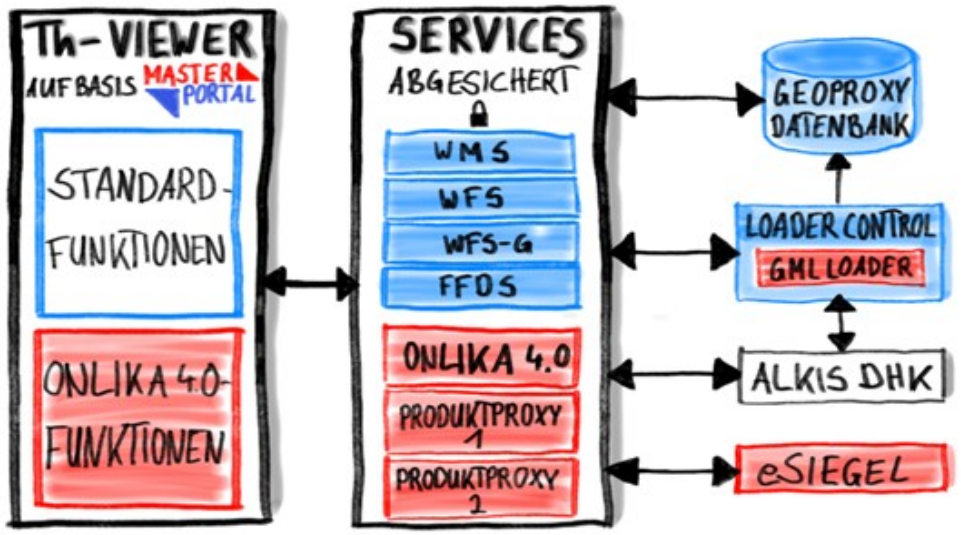


Abb. 1: Der Aufbau von ONLIKA 4.0. Die Oberfläche basiert auf dem Open-Source-Framework Masterportal

Mit der fortschreitenden Digitalisierung und dem steigenden Bedarf an präziser Geodatenverarbeitung stehen öffentliche Verwaltungen und Unternehmen vor der Herausforderung, ihre Daten sicher und effizient zu managen. Dies gilt besonders für die Verarbeitung und Beauskunftung sensibler Liegenschaftskataster Daten. Das automatisierte Abrufverfahren “Online-Auskunft für die Daten des Liegenschaftskatasters” Thüringen - kurz ONLIKA 4.0 - ist ein gutes Beispiel dafür, wie moderne Technologie und höchste Sicherheitsstandards kombiniert werden können, um diese Anforderungen zu erfüllen. Der Betrieb von ONLIKA 4.0 erfolgt zukünftig im brandneuen und innovativen RZ+ des Landesrechenzentrums Thüringen (TLRZ).

Optimierungen, Funktionalitäten und der Einsatz von Open-Source-Technologie

Ein zentrales Ergebnis der Weiterentwicklung von ONLIKA 4.0 ist die umfassende Modernisierung der Benutzeroberfläche. Diese neue Oberfläche wurde entwickelt, um den heutigen Anforderungen an Webanwendungen gerecht zu werden. Ein Fokus lag dabei auf der Optimierung der Benutzerführung, die es den Anwendern ermöglicht, Beauskunftungen effizienter abzurufen und einfacher mit den Geodaten zu interagieren.

Die Modernisierung der Oberfläche basiert auf dem Open-Source-Framework Masterportal. Da Masterportal deutschlandweit von einer Vielzahl von Nutzern eingesetzt und kontinuierlich weiterentwickelt wird, profitiert die Liegenschaftskatasterauskunft Thüringen hiervon. Somit bleibt ONLIKA 4.0 eine flexible und zukunftssichere Lösung, die sich an die sich verändernden Bedürfnisse ihrer Nutzer anpassen kann.

Neben der neuen Oberfläche bringt ONLIKA 4.0 auch neue Funktionen, die die Arbeit mit Geodaten deutlich vereinfachen: So können Flurstücke nun direkt durch einen Klick im Kartenfenster ausgewählt werden. Weiterhin wird das Masterportal Höhenprofil Addon in der ONLIKA Oberfläche nachgenutzt.

Der Produktproxy bleibt ein zentrales Kommunikationselement. Er dient als Schnittstelle zur ALKIS-Datenhaltungskomponente (DHK) und zu den DHK-Diensten. Die DHK umfasst den Produktdienst, der für die AdV-Produkte und die landesspezifischen ALKIS-Produkte zuständig ist sowie die DHK-Dienste, welche für die NAS-Abgaben verantwortlich sind. Der Produktproxy ermöglicht ebenfalls die Kommunikation mit dem eSiegel-Server zur Signierung von ALKIS-Produkten mit einem qualifizierten eSiegel nach eIDAS. An der Softwarekomponente Produktproxy selbst wurden keine Änderungen vorgenommen, was die Kontinuität und Stabilität der bewährten Prozesse sicherstellt.

Sicherheit und Zukunftsfähigkeit

Trotz der zahlreichen funktionalen Erweiterungen und technologischen Neuerungen bleibt die Sicherheit ein zentraler Aspekt von ONLIKA 4.0. Die Authentifizierung erfolgt über das Thüringer ID-Mercury, das in der aktuellen Version bereits die direkte Integration mit der BundID ermöglicht und damit die Grundlage für eine sichere Identitätsverwaltung auf Bundesebene schafft. Der Austausch sicherheitskritischer Informationen zwischen den verschiedenen Diensten und dem Backend erfolgt über bewährte Protokolle wie SAML in der Schnittstellenversion 2 und der Authentifizierungsschicht OpenID-Connect.

Da alle Instanzen in getrennten Containern betrieben werden und keine gemischte Nutzung mit nicht abgesicherten Viewern wie dem Th-Viewer oder BORIS-Viewer erfolgt, wird die Sicherheit der Plattform erhöht. Dieser Ansatz stellt sicher, dass nur authentifizierte und autorisierte Benutzer auf die Liegenschaftsdaten und -dienste zugreifen können.

Vom Labor in die Testphase

In einer Testphase testen eine Gruppe aus internen Mitarbeitenden des TLBG sowie vier öffentlich bestellte Vermessungsingenieure. Das Feedback aus dieser Testphase fließt direkt in die Entwicklung. Das finale Produkt wird mehr als 3000 Thüringer Anwendern, darunter allen Landesangestellten, ÖBVIs, aber auch berechtigten Anwendern, wie Banken und Kommunen, zur Verfügung stehen.

Potenzial für andere Verwaltungen: Anpassungsfähige und skalierbare Lösung

ONLIKA 4.0 ist ein zentraler Bestandteil der Geodateninfrastruktur Thüringens und bietet auch anderen Verwaltungen die Möglichkeit, ihre ALKIS-Daten in der GeoInfoDok 7.1.2 Version zu beauskunften. Dank seiner modularen und offenen Architektur, die den Einsatz von Masterportal und standardisierten

OGC Web-Services integriert, lässt sich ONLIKA 4.0 an die spezifischen Anforderungen anderer Bundesländer oder kommunaler Verwaltungen anpassen. Die Offenheit des Systems ermöglicht es zudem, die in Thüringen gesammelten Erfahrungen zu teilen und weiterzuentwickeln, sodass andere Verwaltungen von einem gemeinsamen Wissens- und Lösungs-Ökosystem profitieren können.

Fazit: Ein Meilenstein in der Geodaten Beauskunftung

ONLIKA 4.0 ist mehr als nur ein Upgrade – es ist ein Meilenstein in der Weiterentwicklung der Geodateninfrastruktur Thüringens. Durch die Kombination aus modernster Technologie, Open-Source-Lösungen und einer sicheren, skalierbaren Geodateninfrastruktur bietet ONLIKA 4.0 eine zukunftssichere Plattform für die Verwaltung und Verarbeitung von ALKIS-Daten sowie Geodaten. Die erfolgreiche Zusammenarbeit zwischen dem TLBG, TMIL, grit und dem TLRZ hat gezeigt, dass technologische Innovation und hohe Sicherheitsstandards Hand in Hand gehen können, um eine leistungsfähige und nachhaltige Geodatenlösung zu schaffen, die den wachsenden Anforderungen der digitalen Welt gerecht wird.

Kontakt zum Autor:

Laura Meierkort
grit GmbH
Landwehrstraße 143, 59368 Werne
+49 23 89 98 27 0
meierkort@grit.de

Erik Jost
grit GmbH
Landwehrstraße 143, 59368 Werne
+49 23 89 98 27 0
jost@grit.de

Transitous - Freies Public Transport Routing

Volker Krause

Während es im OSM Umfeld eine Reihe Routing-Dienste für den Individualverkehr gibt, fehlte bisher ein freier und unabhängig betriebener Routing-Dienst für den öffentlichen Personenverkehr, der nicht auf einen engen räumlichen Bereich begrenzt ist. Transitous schließt diese Lücke.

Routing-Dienste für den öffentlichen Personenverkehr werden häufig von Verkehrsanbietern betrieben, was einen inheränten Interessenkonflikt mit sich bringt. Unabhängig betriebene Dienste wie die Open Trip Planner Installationen in Herrenberg oder Norwegen zeigen, was im Bereich intermodales Routing ohne diesen Konflikt möglich wird.

In beiden Fällen hat man räumliche Grenzen zu berücksichtigen, was für Nutzende den Umgang mit zahllosen verschiedenen Apps und Webseiten bedeutet.

Mit zunehmender Verfügbarkeit freier Fahrplandaten und Verbesserungen der freien Routing-Software MOTIS kam eine anbieterunabhängige und räumlich unbegrenzte Alternative letztes Jahr in greifbare Nähe und die Arbeit an Transitous begann.

Zum Stand der Einreichung verarbeitet Transitous 1300 GTFS und GTFS-RT (Echtzeit-)Fahrplandatensätze aus 48 Ländern und wird von mehreren Apps genutzt, Trend stetig steigend.

Dieser Vortrag stellt vor, was Transitous leistet, wie es aufgebaut ist, wie sich weitere Gebiete bzw. Anbieter hinzufügen lassen und es sich in eigenen Anwendungen nutzen lässt.

Ermittlung von Flächenversiegelung als komplexe Aufgabe (Projekt SEAL)

Dr. Peter Lorkowski

Niedersachsen Wasser Kooperations- und Dienstleistungsgesellschaft mbH

Hintergrund

Die Niedersachsen Wasser erfasst als Tochtergesellschaft des OOWV (Oldenburgisch-Ostfriesischer Wasserverband) den Versiegelungsgrad für diverse Gemeinden. Neben herkömmlichen Methoden wie GIS-Analyse und Luftbildinterpretation soll dabei auch zunehmend das Potential von Künstlicher Intelligenz (KI) bzw. Machine Learning (ML) genutzt werden. Die Vielzahl und Heterogenität der für diesen Prozess in Frage kommenden Datenquellen und Methoden stellt dabei eine besondere Herausforderung dar. Nach anfänglich im Rahmen eines Proof of Concept durchaus vielversprechenden Ergebnissen mit ML müssen diese nun auf einer breiteren Ebene und im Hinblick auf deren Eignung zur tatsächlichen Gebührenermittlung evaluiert werden [4]. Eine sorgsame Auswahl und Kombination von Daten und Methoden sowie eine realistische Einschätzung der jeweils erreichbaren Aussagekraft soll verwertbare Ergebnisse liefern.

Für die konsistente Organisation und somit den kontrollierten Ablauf der hier erforderlichen Prozesskette sind umfassende Modellabstraktionen und Architekturentscheidungen notwendig [6]. Nur auf diese Weise kann die Aufgabe in ihrer zwangsläufigen Komplexität überhaupt bewältigt werden. Anforderungen der Nachvollziehbarkeit, Reproduzierbarkeit, Nachweisbarkeit und Prozess-Dokumentation sind die Treiber der hier vorgestellten Systemarchitektur.

Abstraktion von Daten und Prozessen

Die abflusswirksamen Flächen auf einem Grundstück als Grundlage zur Erhebung von Gebühren werden als Abfolge von Arbeitsschritten und Zwischenergebnissen ermittelt. Die in diesem Zusammenhang relevanten Daten und Prozesse lassen sich grundsätzlich vier Bereichen zuordnen, für die jeweils spezifische Eigenschaften gelten (siehe Abbildung 1):

- Der Bereich der externen Daten unterliegt nicht dem Einfluss der Systembetreiber und dient lediglich als Input-Quelle für alle weiteren Prozesse. Spezifikationen müssen somit übernommen werden. Wie in Abbildung 1 angedeutet, werden auch externe Daten durch einen Prozess generiert, dessen Ausgestaltung und Turnus (z. B. Bildflüge) wiederum meist nicht beeinflusst werden können.
- Interne Prozesse können gemäß der systemeigenen Spezifikationen konfiguriert, reproduziert, protokolliert und evaluiert werden. In diesem Bereich liegt die Gestaltungsfreiheit in Bezug auf Prozessoptimierung.
- Zwischendaten erfüllen als Kettenglieder innerhalb der gesamten Prozesskette eine zentrale Funktion. Sie enthalten den zu ihrer jeweiligen Erzeugung aufgewendeten Arbeitsaufwand (durch Mensch bzw. Maschine) als Mehrwert. Die Quantifizierung dieses Mehrwertes macht Aufwandsabschätzungen für andere, ähnliche Aufgaben möglich. Durch Aktualisierung von weiter vorne in der Prozesskette befindlichen Datenbeständen können die nachfolgenden Kettenglieder kaskadierend obsolet werden (siehe auch Abbildung 3).
- Schließlich stellen die kostenwirksamen Daten den eigentlichen Beweggrund zur Entwicklung des hier vorgestellten Systems dar. Diese Daten wirken sich direkt auf den Geschäftsprozess der Gebührenerhebung aus und müssen daher einem hohen Qualitätsanspruch gerecht werden. Sie sind jedoch für sich selbst genommen wertlos und unterliegen einer durch die Nachvollziehbarkeit ihrer Entstehung einzulösende Rechenschaftspflicht (Erklärbarkeit).

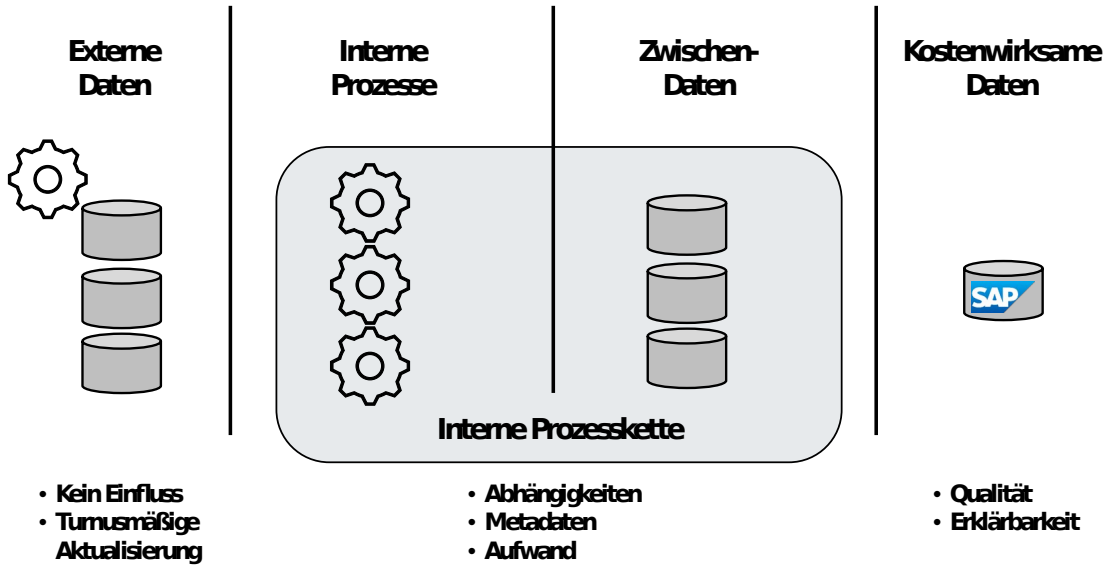


Abbildung 1: Kontexte von Daten und Prozessen nach Herkunft, Zweck und anderen Eigenschaften. Die interne Prozesskette (grauer Bereich) repräsentiert dabei eine Sequenz von Prozessschritten, die einheitlich beschreibbar und steuerbar sein sollen.

Die Ermittlung der kostenwirksamen Daten muss also stets als turnusmäßiger Aktualisierungsprozess statt als statischer Datenbestand gedacht werden [4]. Umso bedeutender wird deshalb die auf Automatisierung, Reproduzierbarkeit und Konfigurierbarkeit abzielende Beschreibung der jeweiligen Arbeitsschritte. Diese Metadaten müssen innerhalb des hier beschriebenen Systems einem definierten Schema folgen, um diesen Zweck erfüllen zu können. Dabei kann zwischen der Eigenschaften von Daten und den Eigenschaften von Prozessen unterschieden werden (siehe Abbildung 2). Während bei Daten eher beschreibende Eigenschaften wie raumzeitliches Bezugssystem, Struktur und Herkunft im Vordergrund stehen, werden bei einem Prozess diejenigen Daten als Metadaten gesehen, die der Methodenspezifikation, Steuerung und Konfiguration dienen.

Das KI-Modell (Layer von Neuronen-Gewichten [1]) kann in diesem Zusammenhang sowohl Daten als auch Steuerdaten zugeordnet werden. Es wird jedoch innerhalb der Prozesskette als Ergebnis eines aufwendigen Trainingsprozesses [1] mit Orthofotos und gelabelten bzw. Ground-Truth-Daten als Eingangsdaten erstellt. Obwohl dieses Modell dabei eher den Charakter einer Prozesssteuerung beim Inferencing statt von echten Eingabedaten hat, erscheint die Zuordnung zur Daten-Kategorie hier insgesamt sinnvoller (Abbildung 2). Auf diese Weise kann auch die Entstehungsgeschichte von KI-Modellen formal beschrieben werden.

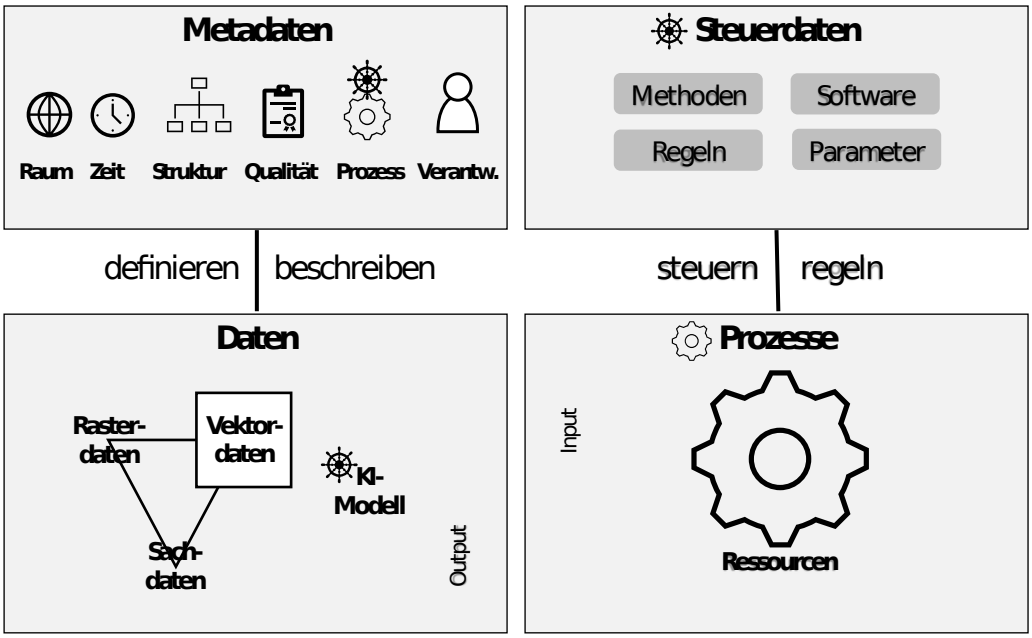


Abbildung 2: Beschreibung von Daten und Prozessen mittels Metadaten und Steuerdaten

Aufgrund der Input/Output-Beziehungen zwischen Eingabedaten, Prozessen und Ausgabedaten und der grundlegenden Abhängigkeitsstruktur ergibt sich für den Gesamtprozess eine Kaskade oder Daten-Pipeline (siehe Abbildung 3) mit einer schrittweisen Zunahme des informationellen Mehrwertes. Die Validierung der KI-Ergebnisse stellt innerhalb der Prozesskette die besondere Bedeutung einer iterativen Optimierung des KI-Prozesses zu [1].

Insgesamt dient die beschriebene Abstraktion innerhalb des Projektvorhabens folgenden Zwecken:

- Die Komplexität soll durch den so geschaffenen Ordnungsrahmen mental besser zu bewältigen sein [6].
- Eine effiziente Kommunikation mit Projektbeteiligten und Stakeholdern wird gefördert [5][6].
- Durch die Definition eines projektspezifischen Metadaten-Standard kann eine konsistente und qualitativ hochwertige Dokumentation gewährleistet werden, die den Nutzwert der Daten deutlich erhöhen kann.
- Die formale deklarative Beschreibung von Prozessschritten ist die Voraussetzung zur reproduzierbaren und konfigurierbaren Automation einer solchen Prozesskette.
- Das Erfassen des Aufwandes (Arbeitszeit von Personen, CPU-Zeit) eines Prozessschritts erlaubt eine Aufwandsschätzung bei Arbeitspaketen mit anderem Datenvolumen und erlaubt auch Aussagen über die Skalierbarkeit [2][3].
- Die strukturierte und standardisierte Beschreibung von Daten und Parametern ermöglicht eine bessere Kontrolle der Prozesskette [3].
- Arbeitsschritte und Prozessläufe können anforderungsgerecht dokumentiert und deren Zwischenergebnisse systematisch archiviert werden.

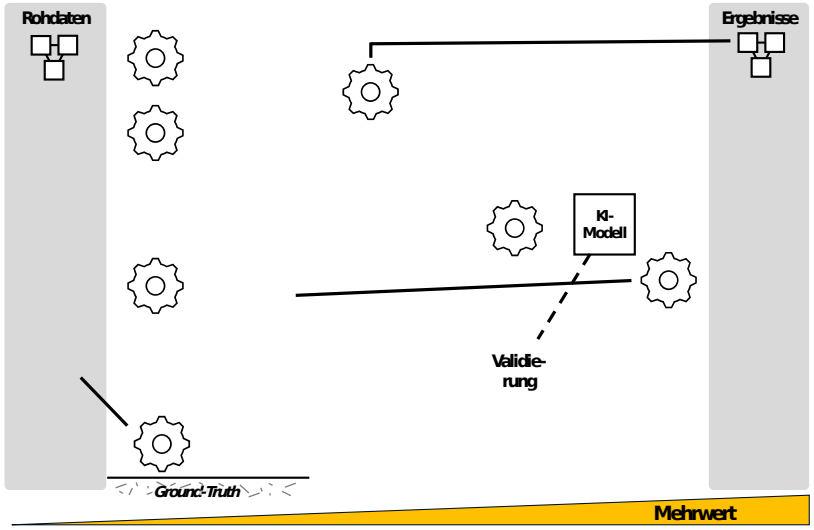


Abbildung 3: Prozesskette mit Daten-Elementen (Kästen) und Prozess-Elementen (Zahräder). Die Validierung dient dabei der iterativen Verbesserung des ML-Modells.

Umsetzungsstrategie

Das hier ausgeführte Beschreibungsschema der Prozesskette zur Ermittlung des Versiegelungsgrads setzt den Rahmen für einen systematischen und nachvollziehbaren Umgang mit der insgesamt sehr komplexen Datenverarbeitungsaufgabe. Eine von konkreten Produkten, Formaten, Programmiersprachen und Methoden abstrahierende Betrachtung minimiert Abhängigkeiten und erlaubt einen flexibleren Einsatz von Einzelkomponenten [5].

Als Datenquellen werden derzeit folgende genutzt bzw. sind vorgesehen:

- Orthofotos (z. T. TrueDOP, z. T. NIR-Kanal)
- DGM und DOM
- ALKIS-Daten (Tatsächliche Nutzung und Flurstücke)
- ATKIS-Daten
- Kanaldaten (Verbrauchsstellen-Positionen)
- Kundendaten aus SAP

Als Software kommen QGIS mit Geopackages und PostgreSQL / PostGIS zum Einsatz. Die Metadaten werden derzeit noch mit prototypisch MS Access verwaltet und sollen künftig in SQLite bzw. PostgreSQL geführt werden. Als Programmiersprache für GIS-Prozesse wird Python, für den ML-Bereich die Keras-Bibliothek verwendet. Über eine entsprechend gestaltete API-Definition soll der Austausch einzelner Komponenten grundsätzlich ohne größere Seiteneffekte ermöglicht werden [5][6].

Von der Nutzung spezifischer einschlägiger Tools und Standards zur Prozessautomatisierung und Metadatenverwaltung wurde zum jetzigen Zeitpunkt noch abgesehen. Die strategische Ausrichtung ist noch ausschließlich anforderungsgetrieben und auf die Vermeidung jeglicher technischer Abhängigkeiten ausgerichtet. Da aufgrund dieses Ansatzes mit Sicherheit zu erwarten ist, dass sich jetzt noch nicht bekannte Anforderungen ergeben werden, würde eine zu frühe Entscheidung zugunsten einer bestehenden Lösung u. U. mit hohen Anpassungsaufwänden verbunden, falls diese Lösung die Anforderung nicht erfüllt. Dies widerspräche der grundsätzlich auf Agilität ausgerichteten Projektstrategie [5].

Die Auswahl und Steuerung der gewählten Verarbeitungsmethoden von Datenbank-, GIS- und KI-Prozessen stellt angesichts der Vielzahl von Produkten und deren jeweiligen Konfigurationsmöglichkeiten bereits eine große fachliche Herausforderung dar. Diese Auswahl und Steuerung nicht jeweils über ein

GUI-Menü auszuwählen, sondern aus einer expliziten persistierten Konfiguration aus einer Datenbank heraus durchzuführen stellt ein Kernelement der Prozessautomatisierung dar. Für die konsistente Verwaltung, aber auch für die experimentelle Permutation von umfangreichen und ggf. hierarchisch organisierten Prozess-Parametern wird ab einem gewissen Komplexitätsgrad eine rekursiv innerhalb eines Komposite-Patterns [5] organisierte Speicherstruktur zum Einsatz kommen müssen [3]. Hier wird die besondere Herausforderung darin bestehen, auch bei Nutzung verschiedenster Softwarepakete und Berechnungsvarianten den jeweils verwendeten Parametersatz sowohl gegenüber Mensch als auch Maschine konsistent zu gestalten.

Zusammenfassung und Fazit

Die Ermittlung versiegelter Oberflächen stellt eine Herausforderung in Bezug auf das Daten- und Prozessmanagement unter Verwendung unterschiedlichster Datenquellen und -formate, Software-Produkte und Konfigurationsmöglichkeiten dar. Das hier vorgestellte Rahmenwerk spiegelt die wesentlichen strategischen Entscheidungen zugunsten einer generischen und flexiblen Systemarchitektur für diese Aufgabe wider. Spezifische Beschreibungsschemata für Daten und Prozesse sowie deren gegenseitige Abhängigkeiten fördern das Systemverständnis und bilden ein beliebig konfigurierbares Kontrollinstrument zur Spezifikation und Reproduzierbarkeit auch komplexer Prozessketten. Die Implementierung der hier vorgestellten Komponenten verfolgt eine auf minimale Abhängigkeiten abzielende Strategie. Eine von bestehenden Lösungen abstrahierende Grundstruktur soll den Austausch von Komponenten erleichtern und als grundlegende Architekturmerkmale sowohl Stabilität als auch Flexibilität gleichermaßen bieten.

Kontakt zum Autor:

Dr. Peter Lorkowski
Niedersachsen Wasser Kooperations- und Dienstleistungsgesellschaft mbH
Georgstraße 4
26919 Brake
Fon: 0160/90566923
E-Mail: lorkowski@niedersachsen-wasser.de

Literatur

[1] Kamusoko, Courage: Explainable Machine Learning for Geospatial Data Analysis, Boca Raton FL, 2025

[2] Brinkhoff, Thomas: Geodatenbanksysteme in Theorie und Praxis, 4., überarbeitete und erweiterte Auflage, Berlin, 2022

[3] Lorkowski, Peter: Monitoring Continuous Phenomena: Background, Methods and Solutions, Boca Raton FL, 2021

[4] Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V. (DWA): Arbeitsblatt DWA-A 102-2/BWK-A 3-2, Grundsätze zur Bewirtschaftung und Behandlung von Regenwetterabflüssen zur Einleitung in Oberflächengewässer – Teil 2: Emissionsbezogene Bewertungen und Regelungen, Hennef, 2022

[5] Evans, Eric: Domain Driven Design: Tackling Complexity in the Heart of Software, Boston, 2004

[6] Taylor, R. N.; Medvidovic, Nenad; Dashofy, Eric: Software Architecture: Foundations, Theory and Practice, Danvers MA, 2010

FOSSGIS bei OpenCode.de

Torsten Friebe / David Arndt

Regionalverband Ruhr

OpenCode ist aus einer Initiative des Bundesministerium des Innern und für Heimat hervorgegangen. Im letzten Jahr (2024) ist der Betrieb der Plattform an die ZenDiS GmbH übergegangen, die seitdem den Betrieb und die Weiterentwicklung der Plattform verantwortet. Im Zuge der Weiterentwicklung gab es am 15. Januar 2025 einen kompletten Relaunch.

Ziel der Plattform ist es Open-Source-Lösungen in der Öffentlichen Verwaltung zu stärken und die digitale Souveränität voranzutreiben. Die Plattform stellt einen Lizenzkatalog und Informationen zu Open-Source-Kriterien und -Definitionen bereit.

Die Vision einer zentralen Plattform, die gleichermaßen von Bund, Ländern und Kommunen genutzt werden kann, um gemeinsam an Open-Source-Lösungen zu arbeiten ist hiermit umgesetzt worden.

OpenCode steht für alle Verwaltungsangestellten und Open-Source-Dienstleister zur Entwicklung, Kommunikation und zum Austausch von Softwarecode zur Verfügung.

Die Plattform weißt das Potenzial auf sichere Softwarelieferketten zu schaffen und auszubauen. Gemeinsam mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI) wird wird dieses gerade umgesetzt. Dabei sollen automatisierte Qualitäts- und Sicherheitschecks auf der Plattform umgesetzt werden.

Nutzern steht eine Wissensdatenbank rund um das Thema Lizenzen, Beschaffung und Nachnutzung bereit. Dabei ist es auch möglich über eine Diskussionsplattform miteinander ins Gespräch zu kommen.

Kontakt zum Autor:

Torsten Friebe
lat/lon GmbH
Im Eillig 1
53343 Wachtberg
0228/24333784
friebe@lat-lon.de

Dipl. Geogr. David Arndt
Regionalverband Ruhr
Kronprinzenstraße 6
45128 Essen
0201/2069-412
arndt@rvr.ruhr

Literatur

[1] Pressemitteilung ZenDiS GmbH zum Relaunch von OpenCode.de: https://zendis.de/ressourcen/opencode-relaunch-vom-softwareverzeichnis-zur-plattform-fuer-die-digitale-souveraenitaet

[2] Webauftritt ZenDiS GmbH: https://www.zendis.de

[3] Webauftritt openCode: https://www.opencode.de

Barrierefreies Routing mit MOTIS

Felix Gündling

MOTIS v2 routet nicht nur die erste und letzte Meile auf OpenStreetMap Daten, sondern auch Umstiegswege. Um dies effizient auch für große Gebiete zu ermöglichen und Stockwerke korrekt zu behandeln, wurde ein skalierbarer, profilbasierter Router (OSR) entwickelt. Hierbei wird in der Verbindungssuche die zeitliche Verfügbarkeit von Aufzügen und Einstiegshilfen am Bahnhof berücksichtigt. Wege und andere Informationen werden nach Level unterteilt dargestellt.

Die Berechnung von Umstiegen beim Routing im öffentlichen Verkehr benötigt insbesondere für mobilitätseingeschränkte Menschen eine höhere Präzision in den Daten als das, was Fahrplandaten in vielen Fällen heute bereitstellen. Um diese Lücke zu schließen, setzt MOTIS bei der Berechnung von Umstiegen auf OpenStreetMap Daten. Diese werden sowohl mit den Fahrplandaten als auch mit den Aufzugsdaten verknüpft. Da bisherige Routing Engines für eine weltweite Abdeckung einen hohen Speicherbedarf haben, wurde eine neues profilbasiertes Routing entwickelt, das durch ein effizientes Datenmodell mit sehr wenig RAM auskommt (aktuell ~16 GB Routing Daten für weltweite OSM Daten). Zusätzliche Profile können ohne zusätzlichen Speicherverbrauch umgesetzt werden. Beim Routing werden für mehrstöckige Bauwerken auch die Stockwerksinformation berücksichtigt. Die Wiederverwendung von OSM Knoten in mehreren Wegen auf verschiebenen Ebenen wird hierbei unterstützt. Das Routing kann zur Laufzeit Knoten ausschließen, was z.B. genutzt wird, um nicht verfügbare Aufzüge auszuschließen. Nach einer Verknüpfung der Fahrplandaten auf die OpenStreetMap Daten (hier ist insbesondere das Stockwerk relevant), werden alle Umsteigebeziehungen für alle Kombinationen von funktionierenden bzw. nicht funktionierenden Aufzügen berechnet und in den ÖV-Router eingespeist. Dieser kann nun zeitabhängig Fahrstuhlausfälle sowie die Verfügbarkeit von Einstiegshilfen berücksichtigen.

Verarbeitung offener Satellitendaten mit freier Software für die visuelle Anwendung

Christoph Hormann

Was steckt dahinter, wenn man in einem interaktiven Kartendienst im Internet die Satellitenbild-Ebene betrachtet? Dieser Vortrag gibt einen Einblick auf Grundlage von 20 Jahren praktischer Erfahrung in der Produktion derartiger Darstellungen.

Satellitenbild-Darstellungen sind heute in öffentlich zugänglichen Kartendiensten allgegenwärtig und beliebt. Und viele dieser Bilder basieren zumindest teilweise auf offenen Satellitendaten. Anders als bei kartographischen Darstellungen werden jedoch die praktischen Verfahren zur Produktion solcher Visualisierungen öffentlich sehr wenig diskutiert.

Dieser Vortrag erläutert zunächst die Rolle und Geschichte von Luft- und Satellitenbild-Ebenen in Kartendiensten als eigene Visualisierungsform zwischen originären Fotos und kartographischen Darstellungen und gibt einen Überblick über die grundsätzlichen Herausforderungen bei der Satellitenbild-Verarbeitung für diese Anwendung.

Darauf aufbauend wird der Stand der Technik im Bereich großräumiger Satellitenbild-Zusammenstellungen erläutert was Methoden, Datenquellen und Qualität der Ergebnisse betrifft.

Daneben werden einige der speziellen Probleme bei der Verwendung offener Datenquellen und der Nutzung freier Software in diesem Bereich anhand konkreter Beispiele erläutert.

10 Jahre openSenseMap: Neue Features für die Zukunft der Umweltdaten

Frederick Bruch

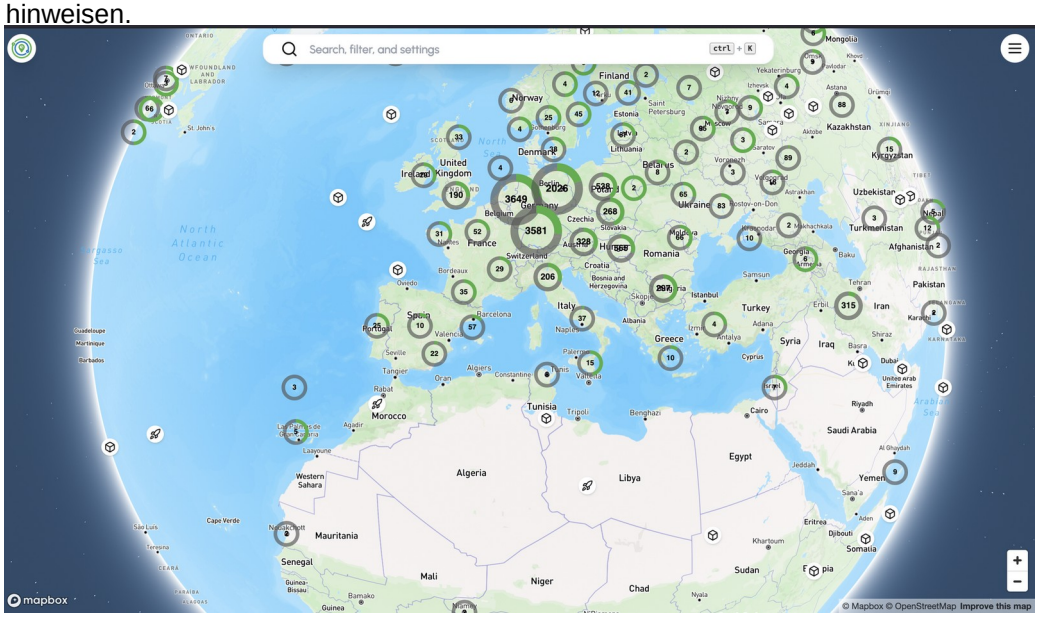
Die openSenseMap ist eine interaktive Karte für Umweltdaten, die von Schulen, Städten und Bürger*innen weltweit genutzt wird. Seit 2014 wurden über 14.000 Sensoren registriert und mehr als 10 Milliarden Messwerte in Echtzeit bereitgestellt. Die neue Version verbessert die Datenvisualisierung und bietet viele weitere neue Features. In unserem Lightning-Talk stellen wir diese Neuerungen vor und feiern 10 Jahre openSenseMap, die einen wichtigen Beitrag zu Forschung und Bildung leistet.

Die openSenseMap ist eine für jeden frei nutzbare interaktive Karte mit Umweltdaten. Diese Daten werden von Schulen, Forschungseinrichtungen, Städten und engagierten Bürger:innen weltweit erhoben und genutzt. Sensoren können sowohl an stationären als auch mobilen Geräten Messungen zu verschiedenen Umweltphänomenen erheben und stellen diese in Echtzeit über die eigene RESTful API als offene Daten zur Verfügung. Zudem gibt es ein angeschlossenes Archiv in dem alle Daten seit dem Start der openSenseMap in 2014 gespeichert werden.

Die Nutzung ist vielfältig: engagierte Bürgerinnen messen die Luftqualität und den Lärmpegel in ihrer Straße, Schülerinnen experimentieren im Unterricht und werden spielerisch für Naturwissenschaften, Technik und die Erforschung ihrer Umwelt begeistert, Wissenschaftlerinnen nutzen die umfangreichen frei verfügbaren Daten für komplexe Forschungsprojekte, Städte betreiben Smart City-Anwendungen und informieren ihre Bürgerinnen live über Temperatur, Luftfeuchtigkeit, Luftdruck, Feinstaub und geben nützliche Tipps, wie die Standorte freier Parkplätze.

Als Reaktion auf die stetig wachsende Anzahl registrierter Stationen (über 14.000) und die über 10 Milliarden übermittelten Messwerte wurde eine neue Version der openSenseMap Plattform entwickelt. Dabei musste die Technologie teilweise neu konzipiert werden, um das steigende Datenvolumen effizient verarbeiten zu können. Die neue Version bringt zahlreiche Features mit sich, darunter die Anbindung an eine spezielle Sensordatenbank, die eine effizientere und strukturierte Verwaltung der gesammelten Daten ermöglicht. Zudem ist die Plattform jetzt mit einer Open Badges Plattform verknüpft, die es den Nutzern erlaubt, für ihre Beiträge digitale Auszeichnungen zu erhalten. Darüber hinaus bietet die neue Version erweiterte Möglichkeiten zur Durchführung detaillierter Datenanalysen, wodurch Bürger*innen, Forschende und Organisationen tiefere Einblicke in Umweltphänomene gewinnen können.

In diesem Lightning-Talk möchten wir genau diese neuen Features präsentieren und auf unser Nutzertreffen hinweisen.



Datenklassifikation: Mehrwert oder Manipulation?

Jochen Schiewe[1]

[1] HafenCity Universität Hamburg, Labor für Geoinformatik und Geovisualisierung (g2lab)

DOI: 10.5281/zenodo.14616647

Zusammenfassung: Dieser Beitrag behandelt zwei grundlegende Probleme bei der Datenklassifikation, d. h. der Gruppierung quantitativer Werte, z. B. für die Erzeugung von Choroplethenkarten. Das erste Problem besteht darin, dass konventionelle Verfahren wie äquidistante oder Quantil-Einteilungen keine Bewahrung räumliche Muster in mono-temporalen sowie von Veränderungseigenschaften in multi-temporalen Datensätzen vorsehen. Es werden daher entsprechende Verfahren vorgestellt, die eine entsprechende Optimierung vornehmen (*Feature-Preserving Data Classification*). Das zweite Problem besteht darin, dass es weder Standardverfahren, noch „richtige“ oder „falsche“ Methoden für die Datenklassifikation gibt bzw. geben kann. Es wird ein zweckorientierter Einsatz verbunden mit einem kritischen Umgang und einer Bereitstellung von Metadaten vorgestellt.

Schlüsselwörter: Datenklassifikation, Choroplethenkarte, räumliche Muster, Veränderungen

1 Einführung

Bei der Verarbeitung von quantitativen Daten wird häufig eine Datenklassifikation vorgenommen. Eine solche Klassifikation dient z. B. der Reduktion von Datenmengen, Rechenzeiten bei Verarbeitungen und ganz besonders der Erzeugung einer kontrastreicheren Visualisierung (z. B. in thematischen Karten, bei denen Klassen „bis zu 100 Einwohner pro km^2“, usw. gebildet werden). Im Folgenden wird die Anwendung der Datenklassifikation für solche Choroplethenkarten aus Illustrationszwecken im Fokus stehen.

Ein Problem bei der Auswahl und Anwendung von Methoden der Datenklassifikation besteht darin, dass nach Anwendung der Verfahren

- in mono-temporalen Karten wichtige Informationen (insbesondere räumliche Muster wie Extremwerte, Cluster oder Hot Spots) sowie
- in multi-temporalen Karten (wiedergegeben als Kartenanimationen oder sog. *small multiples*) gewisse Veränderungseigenschaften (z. B. große Werteunterschiede oder Trends)

nicht mehr erkennbar sein können. Die unterschiedlichen GIS-Softwareprodukte bieten Standardverfahren und -einstellungen an (z. B. äquidistant, Quantile), die solche Eigenschaften per se nicht bewahren und häufig unreflektiert eingesetzt werden.

Darüber hinaus wird im übergeordneten Kontext das Thema der Datenklassifikation – z. B. auch im Feld der Kritischen Kartographie - kontrovers diskutiert, bis hin zum Vorwurf der Manipulationsmöglichkeit. Dies liegt zum einen daran, dass eine Klassifikation immer mit einem Informationsverlust verbunden ist. Zum anderen erzeugen verschiedene Methoden für ein und denselben Datensatz sehr unterschiedliche Ergebnisse, z. B. bei der Erzeugung von Choroplethenkarten, und daraus resultierende Interpretationen (Abschnitt 2).

Basierend auf diesen Problemstellungen bestehen die Ziele dieses Beitrages darin, alternative, erweiterte Methoden der Datenklassifikation vorzustellen, die eine Bewahrung räumlicher Muster (Abschnitt 3) bzw. zeitlicher Veränderungen (Abschnitt 4) bewahren können, und

einen kritischen Umgang bei der Anwendung von Methoden der Datenklassifikation zu verfolgen (Abschnitt 5).

2 Datenklassifikation: Status Quo

2.1 Varianten

Eine Datenklassifikation bei der Gestaltung von Choroplethenkarten ist optional. Ein Verzicht auf die Klassifikation bedeutet jedoch eine größere visuelle Heterogenität, geringere visuelle Kontraste und eine schwierigere Extraktion von Werten aus den nicht klassifizierten Farbwerten.

Entscheidet man sich für eine Datenklassifikation, gibt es eine Reihe unterschiedlicher, auch in der Standard GIS-Software implementierter Verfahren, die unterschiedliche Regeln und Zwecke verfolgen. Ein gängiger und intuitiv „erwartbarer" Ansatz ist die *äquidistante Gruppierung.* Diese gleichabständigen Klassengrenzen sind jedoch für schiefe Datenverteilungen weniger geeignet, da eine sehr homogene Kartenansicht erzeugt wird und viele Unterschiede nicht angezeigt werden können (Abb. 1, links).

Alternativ wird eine Gleichverteilung der Werte in Klassen und einem resultierenden, verbesserten ästhetischen Eindruck postuliert. (Gruppierung nach *Quantilen;* Abb. 1, mitte). Sollen einzelne Extremwerte farblich hervorgehoben werden, ist diese Einteilung allerdings ungeeignet, da die obersten und untersten Klassen i. d. R. mehr als einen Wert beinhalten.

Ein anderer Ansatz ist die Berücksichtigung von statistischen Parametern. Hier gibt es z. B. die Methode der *natural breaks* (auch: *Jenks optimal*), die Grenzen an lokalen Minima im Histogramm setzt (Abb. 1, rechts), oder die Histogramm-Angleichung (z. B. mit engen Klassenbreiten in der Mitte einer Normalverteilung).

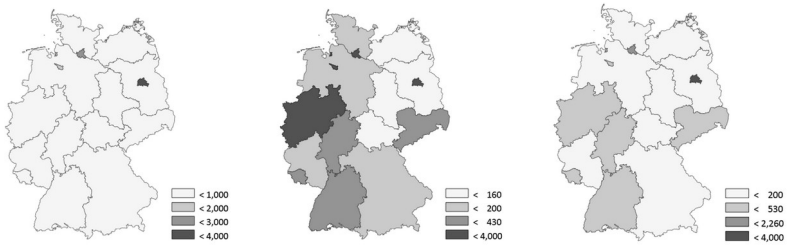


Abb. 1: Anwendung verschiedener Methoden zur Datenklassifikation für einen Datensatz (Bevölkerungsdichte in Deutschland, 1995), v.l.: äquidistant, Quantile, *natural breaks* (nach Schiewe 2022)

Aus den unterschiedlichen Zwecken wird bereits deutlich, dass es keine „Standard"-Methode geben kann. Abb. 1 verdeutlicht zudem, dass unterschiedliche Methoden sehr unterschiedliche Ergebnisse – und damit unterschiedliche Interpretationen – erzeugt werden.

Alle diese Methoden sind rein datenbasiert - räumliche Beziehungen (wie z. B. Nachbarschaften) und insbesondere die Erhaltung räumlicher Muster werden nicht berücksichtigt und können bei der Klassifizierung verloren gehen. Andererseits ist gerade die Erkennung von räumlichen Mustern eine zentrale Aufgabe bei Choroplethenkarten, die durch die Klassifikation erschwert oder gar unmöglich gemacht werden kann.

2.2 Grundlegende Arbeiten

Das Thema Datenklassifikation für kartographische Zwecke wird z. B. in den Überblicksartikeln von Cromley & Cromley (1996) oder Coulsen (1987) umfangreich behandelt. Der Schwerpunkt liegt dabei

auf datengetriebenen Methoden. Einige empirische Studien beschäftigen sich auch mit dem Vergleich solcher Methoden zur Beantwortung typischer Kartennutzungsaufgaben (z. B. Goldsberry & Battersby 2009; Brewer & Pickle 2002). Darüber hinaus wurden interaktive Tools entwickelt, um die „optimale“ Wahl für eine bestimmte Anwendung zu finden, z. B. durch die Verwendung von verknüpften Ansichten zwischen Datenhistogramm und Choroplethenkarte (Andrienko et al., 2001).

Über das bereits erwähnte Problem der Vernachlässigung des räumlichen Kontextes ist relativ wenig veröffentlicht worden; Armstrong et al. (2003) geben dazu einen Überblick. Versuche, Wertunterschiede von räumlichen Nachbarn zu beschreiben und zu erhalten, wurden z. B. von Smith (1986) oder Monmonier (1972) veröffentlicht. Häufig wird eine Vereinfachung der dargestellten Muster angestrebt, damit der Kartennutzer die groben Trends in der Darstellung schneller erfassen kann, ohne durch einen detaillierten und „fleckigen“ Eindruck gestört zu werden (Andrienko et al. 2002; MacEachren 1994). Chang & Schiewe (2019) haben ein Konzept zur Erhaltung ausgewählter räumlicher Muster vorgestellt – dieser Ansatz wird in diesem Beitrag wieder aufgegriffen (Abschnitt 3).

Ferner gibt es keine systematische Behandlung von multi-temporalen Datenklassifikationen in der Literatur. Schiewe (2023) hat ein Konzept erstellt und für die Erhaltung einfacher Veränderungseigenschaften eingesetzt. Cybulski (2022) führte eine empirische Studie zu den Auswirkungen zeitlicher Trends von Mustern in Choroplethenkarten durch und kam ebenfalls zu dem Schluss, dass in Zukunft ein erhöhter Bedarf an effektiver und effizienter kartographischer Visualisierung multitemporaler Daten besteht.

3 Bewahrung räumlicher Muster

Die Tatsache, dass die o. g. Standardverfahren der Datenklassifikation rein datenbasiert arbeiten und keine Nachbarschaften berücksichtigen, führt dazu, dass gewünschte räumliche Auffälligkeiten oder Muster nicht garantiert bewahrt werden können. Abb. 2 macht dies am Beispiel eines lokalen Maximums deutlich, das je nach Klassifikation als solches „graphisch ablesbar“ ist oder nicht.

Chang & Schiewe (2019) bzw. Schiewe (2021) haben eine Methode entwickelt, die ausgehend von der a priori Spezifikation eines möglichen Musters (z. B. lokales Extremum, Hot/Cold Spot, Cluster) eine optimale Bewahrung dieses Musters für einen gegebenen Datensatz ermittelt. Hierzu wurde auch ein entsprechendes QGIS-Plugin entwickelt (https://gitlab.com/g2lab/aChor). Die grundlegende Idee besteht darin, dass zwischen Originalwerten von möglichst allen Regionen, die einen signifikanten Unterschied aufweisen (z. B. lokales Maximum vs. kleineren Werte oder Cluster vs. Nicht-Cluster) auch ein Klassenunterschied festgelegt wird (und damit auch ein Farbunterschied in der kartographischen Darstellung).

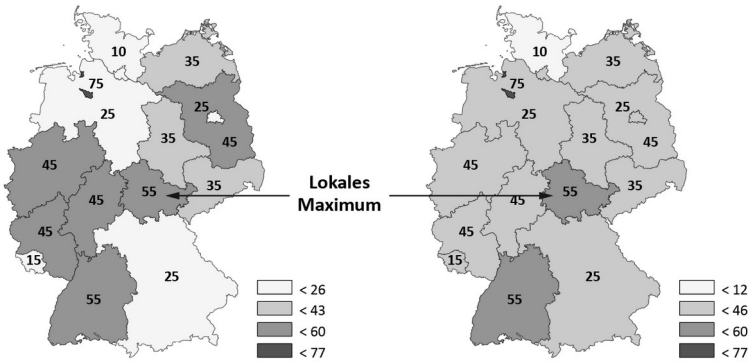


Abb. 2: Lokales Maximum (Werte größer als bei allen direkten Nachbarn) – nach Datenklassifikation nicht deutlich (da in selber Klasse wie einige Nachbarn; links) sowie bewahrt (rechts); Abbildung verändert nach Schiewe 2022

Hierzu werden alle betreffenden Wertepaare bestimmt und in einem *line sweep*-Diagramm aufgetragen. *Line sweep*-Verfahren haben sich bereits zu anderen Zwecken als ein effizientes Hilfsmittel be-

währt, z. B. für eine räumliche Partionierung im Kontext einer Overlay-Operation (Kriegel et al 1992). Hier wird nun die orthogonale *line sweep* von links nach rechts über diese Intervalle bewegt. An allen Startwerten der Intervalle werden Haltepunkte definiert. Diese stellen eine mögliche, gewünschte Klassengrenze dar. Wenn c Klassen erforderlich sind, werden c-1 Linien erstellt. Für jede Linie an jedem Haltepunkt wird die Anzahl der Schnittpunkte gezählt. Dann werden alle Kombinationen aller Linien (für eine gegebene Anzahl von Klassen) in Betracht gezogen, um die beste Lösung zu finden (d. h. so viele Bedingungen wie möglich zu erfüllen).

Abb. 3 zeigt das Prinzip für die Aufgabe „Erhalte lokales Minimum“: Zuerst wird der nächstgrößere Wert in der Nachbarschaft des lokalen Minimums identifiziert und das entsprechende Intervall (x= [51;59]) im *line sweep*-Diagramm aufgetragen. Dieses Vorgehen wird für alle anderen lokalen Minima im Datensatz wiederholt. Dann werden – für die vorgegebene Anzahl von drei Schnitten – die Platzierung der Linien gesucht, die möglichst alle Intervalle schneidet.

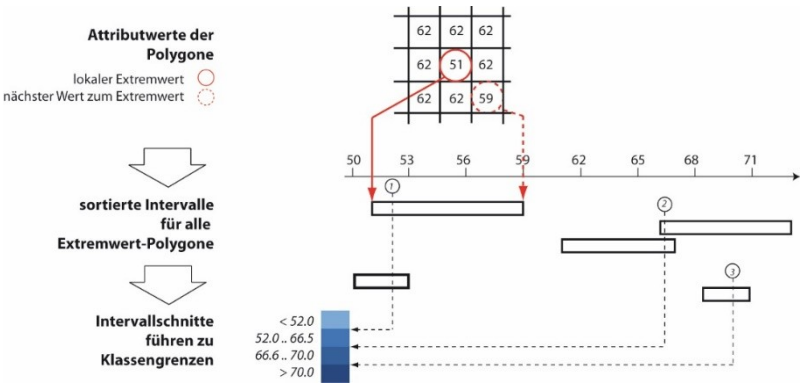


Abb. 3: Prinzip des Verfahrens zur Bewahrung lokaler Minima durch *line sweep*-Verfahren (*featue preserving data classification*)

Abhängig von der Anzahl und Verteilung der Intervalle sowie der Anzahl gewünschter Klassen kann es u. U. nicht möglich sein, alle gewünschten Intervalle zu schneiden. Das Verfahren kann also nur eine bestmögliche (d. h., optimale Lösung) finden. Im Vergleich zu konventio-nellen Verfahren (äquidistant, Quantile) konnte für diesen Ansatz aber eine stets bessere Bewahrung der räumlichen Muster nachgewiesen werden (Chang & Schiewe 2019).

4 Bewahrung Veränderungseigenschaften

Geht man von den oben verwendeten mono-temporalen Darstellungen auf multi-temporale Choroplethenkarten über, ergibt sich die Aufgabenstellung, bestimmte Veränderungen in der Zeitserie auch nach der Datenklassifikation zu bewahren, d. h. sichtbar zu machen.

Analog zum o. g. Vorgehen muss auch hier a priori definiert werden, welche Veränderungseigenschaften erhalten werden sollen. Dies ist allerdings komplexer als im mono-temporalen Fall - folgende Parameter gilt es zu betrachten:

- Area of Interest: Angabe der räumlichen Einheiten (z. B. Bundesländer) - von einer einzelnen Region bis hin zur Gesamtheit aller Regionen im Datensatz.
- Time of Interest: Angabe des Zeitintervalls (Start- und Endepoche) und der zeitlichen Auflösung (Lag).
- Change Properties: Diese können mit Hilfe verwandter (geo-)statistischer Maße beschrieben werden (in einigen Fällen durch Hinzufügen notwendiger Schwellenwerte):

  - Differenz der Attributwerte für ausgewählte räumliche Einheiten innerhalb des ausgewählten Zeitintervalls: globales Maximum/Minimum, lokales Maximum/Minimum, größte/kleinste Differenzen (z. B. ausgedrückt durch Perzentile) usw.
  - Trend: (monotoner) Anstieg/Abfall für ausgewählte räumliche Einheiten innerhalb eines ausgewählten Zeitintervalls, usw.
  - Multi-temporale, räumliche Muster: Zeitliche Veränderung der Kennwerte von Hot/Cold Spots, Clustern usw.
- Visualisierungskontext: Dieser berücksichtigt insbesondere die beabsichtigte Kartennutzung, den Kartennutzer, die Dauer der Präsentation oder das Kartenlayout. All diese Faktoren haben einen Einfluss auf die Anzahl von Klassen, aber auch auf mögliche Gestaltungsmerkmale für
  - die visuelle Betonung von Veränderungseigenschaften, oder
  - die Vermeidung von „False Positives“ (auch: „fake changes“), d. h., dass „unwichtige“ Veränderungen nicht in der gleichen Weise dargestellt werden wie „wichtige“ Veränderungen.

Basierend auf dieser a priori Festlegung werden nun wieder alle Paare oder Tupel von Originalwerten gesucht, die die Bedingung erfüllen (und erhalten bleiben sollen). Dieses Vorgehen wird in folgendem Beispiel demonstriert: Gesucht ist eine Datenklassifikation mit fünf Klassen, die die 90 % größten, positiven Attributwertänderungen für alle räumlichen Einheiten für alle gegebenen Zeitepochen mit einer zeitlichen Auflösung von einer Epoche beibehält, wobei diese Änderungen mit einer Klassendifferenz von 2 (zu Zwecken der visuellen Hervorhebung) dargestellt werden. Darüber hinaus sollen alle anderen Änderungen mit einer Klassendifferenz von weniger als 2 dargestellt werden, um keine „fake changes“ zu suggerieren.

Abb. 4 zeigt das resultierende *line sweep*-Diagramm und die erzeugten Klassengrenzen. In diesem Fall konnten alle Bedingungen erfüllt werden. Abb. 5 vergleicht eine konventionelle multi-temporale Choroplethenkarte basierend auf einer äquidistanten Einteilung (oben) mit der auf der Grundlage der *feature preserving data classification*.

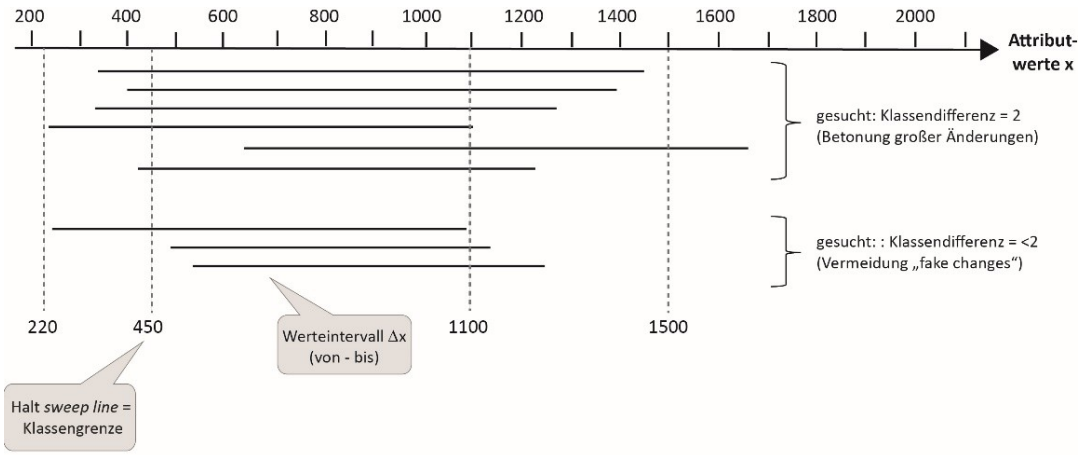


Abb. 4: *line sweep*-Verfahren für multi-temporalen Datensatz (Monatswerte Covid-19 für Januar bis Mai 2022; Daten: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Inzidenz-Tabellen.html)

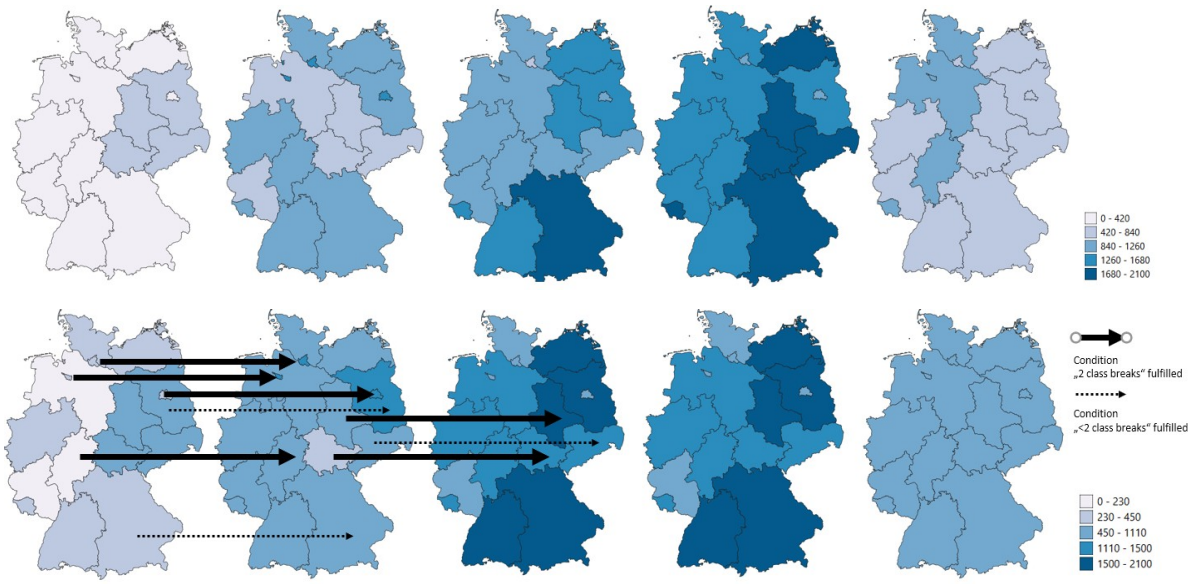


Abb. 5: Multi-temporale Choroplethenkarten für o. g. Datensatz Covid-19 – äquidistante Einteilung (oben) und *feature preserving data classification* (unten; mit Hervorhebung der Wertepaare, die die Veränderungsbedingungen erfüllt haben)

5 Mehrwert oder Manipulation?

Es wurde bereits ausgeführt, dass es keine „Standard“-Methode für die Datenklassifikation gibt bzw. geben kann, und andererseits sehr unterschiedliche Ergebnisse – sowie abgeleitete Interpretationen – erzeugt werden. Die in den Abschnitten 3 und 4 vorgestellten Verfahren zu Bewahrung räumlicher Muster sowie von Veränderungseigenschaften haben die Auswahlmöglichkeiten für die Durchführung der Datenklassifikation sogar noch erweitert.

Es bleibt festzuhalten, dass es keine „richtige“ oder „falsche“ Methode zur Datenklassifikation gibt (vorausgesetzt, die Methode selbst wird korrekt ausgeführt). Jedes einzelne Verfahren hat seine „Daseinsberechtigung“ – sei es, um übliche Erwartungen zu erfüllen (äquidistant), ein gutes Abbild der statistischen Eigenschaften zu geben (*natural breaks* oder Histogramm-Angleichung) oder aber ausgewählte räumliche Muster oder Veränderungseigenschaften zu bewahren (siehe Abschnitte 3 und 4). Aus dieser Vielfalt kann man aber durchaus auch einen Mehrwert ableiten: Es ist eine zielgerichtete Anpassung an einen Zweck zur Darstellung möglich. Dies setzt die a priori Definition eines solchen Zweckes voraus – was bei der Nutzung von Varianten oder gar Standardeinstellungen in der GIS-Software (z. B. *natural breaks* in *QGIS* bzw. *ArcGIS*) nicht geschieht. Und natürlich kann der Zweck auch – bewusst oder unbewusst – in einer Manipulation bestehen – und ist letztlich nicht auszuschließen.

Daraus ergibt sich ein Plädoyer für eine bewusste und kritische Anwendung der Datenklassifikation – sowohl von Seiten der Erstellenden, aus auch von Seiten der Nutzenden. Dies beinhaltet insbesondere die Bereitstellung von Metadaten, die den Zweck verbal artikulieren und basierend auf verfügbaren Daten das Klassifikationsergebnis nachvollziehbar machen.

6 Zusammenfassung und Ausblick

Dieser Beitrag hat zwei grundsätzliche Problem im Zusammenhang mit der Klassifikation von Daten (z. B. als vorbereitende Maßnahme zur Darstellung in Choroplethenkarten) aufgegriffen. Zum einen berücksichtigen konventionelle Verfahren (wie äquidistante oder Quantil-Gruppierungen) nicht die Be-

wahrung bestimmter räumlicher Muster (in mono-temporalen Karten) oder bestimmter Veränderungseigenschafen (in multi-temporalen Karten). Hierfür wurden Methoden vorgestellt, die basierend auf einer a priori Definition der zu bewahrenden Eigenschaften alle Wertepaare selektieren, die die Bedingungen erfüllen. Diese Paare werden in einem *line sweep*-Diagramm aufgetragen, in dem die Suche nach optimalen Klassengrenzen stattfindet.

Die geschilderte Vorgehensweise kann noch optimiert werden. Dies betrifft den hohen Rechenaufwand, der durch die Berücksichtigung aller Schnittkombinationen für mehrere *line sweep*-Linien entsteht – hier sind bereits angedachte Näherungslösungen weiter zu entwickeln. Ferner beruht die a priori Auswahl der zu bewahrenden Eigenschaften bisher auf einer subjektiven Vorauswahl. Hier ist es denkbar, dass eine vorgeschaltete, automatisierte Sequenz von geostatistischen Methoden alle relevanten Eigenschaften abprüfen kann.

Das zweite behandelte Problem betrifft die Tatsache, dass es kein Standardverfahren der Datenklassifikation geben kann, sondern eine Vielzahl von Varianten, die verschiedene Zwecke erfüllen können. Dies kann gleichermaßen als Mehrwert (flexibler Einsatz), als auch als Möglichkeit der Manipulation angesehen werden. Hieraus ergibt sich ein Plädoyer für eine bewusste und kritische Anwendung der Datenklassifikation. Dies sollte auch durch die Bereitstellung von Metadaten geschehen, die den Zweck verbal artikulieren und basierend auf verfügbaren Daten das Klassifikationsergebnis nachvollziehbar machen. Eine entsprechende Entwicklung von (OGC-) Standards in diesem Kontext ist wünschenswert.

Kontakt zum Autor:

Prof. Dr. Jochen Schiewe
HafenCity Universität Hamburg, Henning-Voscherau-Platz 1, 20457 Hamburg
Telefon: +49 40 428275442 / E-Mail: jochen.schiewe@hcu-hamburg.de

Literatur

Andrienko, G., Andrienko, N. & Savinov, A.: Choropleth maps: Classification revisited. Proceedings International Cartographic Conference, 9 p., 2001

Armstrong, M.P., Xiao, N. & Bennett, D.A.: Using Genetic Algorithms to Create Multicriteria Class Intervals for Choropleth Maps. Annals of the Association of American Geographers, 91(3): 595-623, 2003

Brewer, C.A. & Pickle, L.: Evaluation of Methods for Classifying Epidemiological Data on Choropleth Maps in Series. Annals of the Association of American Geographers, 92(4): 662-681, 2002

Chang, J. & Schiewe, J.: An open source tool for preserving local extreme values and hot/coldspots in choropleth maps. KN – Journal of Cartography and Geographical Information. 68(6): 307-309, 2019

Cromley, E.K. & Cromley, R.G.: An analysis of alternative classification scheme for medical atlas mapping. European Journal of Cancer, 32A(9) 1551-1559, 1996

Coulsen, M.R.C.: In the matter of class intervals for choropleth maps: With particular reference to the work of George Jenks. Cartographica, 24(2): 16-39, 1987

Cybulski, P.: An Empirical Study on the Effects of Temporal Trends in Spatial Patterns on Animated Choropleth Maps. ISPRS International Journal of Geo-Information 11 (5) 273 DOI: 10.3390/ijgi11050273, 2022

Goldsberry, K. & Battersby, S.: Issues of Change Detection in Animated Choropleth Maps. Cartographica, 44(3): 201-215, 2009

Kriegel, H.-P., Brinkhoff, T. & Schneider, R. (1992): An Efficient Map Overlay Algorithm Based on Spatial Access Methods and Computational Geometry. In: Gambosi, G., Scholl, M.,

Six, H.-W. (eds.), “Geographic Database Management Systems”, Workshop Proceedings, Capri, May 1991, Springer-Verlag: 194-211.

MacEachren, A.M.: Some truth with maps: A primer on symbolization and design. Association of American Geographers, Washington, DC, 1994

Monmonier, M.: Contiguity-biased class-interval selection and location allocation models. Geographical Review, 62: 203-228, 1972

Schiewe, J.: Data classification methods for preserving spatial patterns, Proc. Int. Cartogr. Assoc., 4, 95, https://doi.org/10.5194/ica-proc-4-95-2021, 2021.

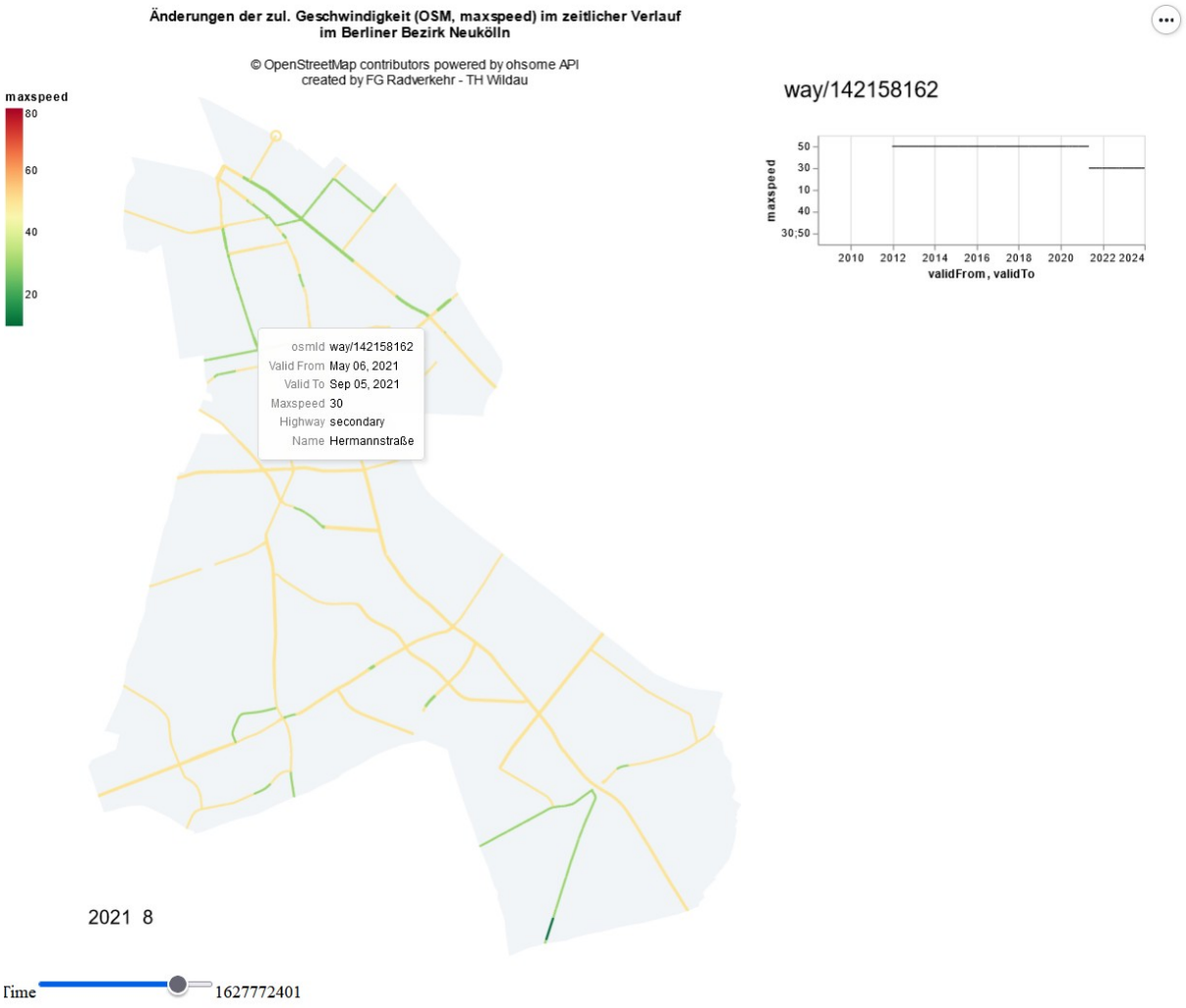
Schiewe, J.: Kartographie. Visualisierung georäumlicher Daten. Springer Spektrum. ISBN 978-3-662-65440-8, 2022

Schiewe, J.: Preserving change information in multi-temporal choropleth maps through an extended data classification method. The Cartographic Journal, DOI: 10.1080/00087041.2023.2267944, 2023

Smith, R.M.: Comparing traditional methods for selecting class intervals on choropleth maps. The Professional Geographer, 38(1): 62-67, 1986

Verkehrliche Maßnahmen aus historischen OSM-Daten identifizieren

Simon Metzler



Eine zentrale Datenbank zu verkehrlichen Maßnahmen fehlt in Deutschland. Dieses Projekt entwickelt einen Prozess, der mithilfe von OSM-Daten und ohsome versucht verkehrliche Maßnahmen räumlich und zeitlich zu identifizieren. Anhand von Prototypen und Szenarien für Benchmarking soll die Methode getestet werden. Die Herausforderungen bestehen insbesondere in der Differenzierung echter Maßnahmen von anderen Änderungen und der Komplexität bei Geometrieanpassungen.

In Deutschland existiert bisher keine einheitliche, öffentlich zugängliche Datenbank, die verkehrliche Maßnahmen und Änderungen räumlich und zeitlich darstellt. Eine solche Datenbasis könnte jedoch wertvolle Erkenntnisse zur Bewertung verkehrlicher Maßnahmen liefern, indem sie deren Auswirkungen – beispielsweise auf Verkehrsnachfrage oder Unfallhäufigkeit – transparenter macht.

Eine mögliche Lösung wäre die Entwicklung eines Prozesses, der mithilfe der OpenStreetMap (OSM)-Datenbank und des ohsome-Frameworks [1] spezifische Maßnahmen wie bspw. Geschwindigkeitsbeschränkungen identifiziert und deren räumliche sowie zeitliche Veränderungen erfasst.

Vorgehen im Projekt:

1. Grundlage schaffen:

* Im ersten Schritt wurde ein grundlegender Verarbeitungsprozess in Python entwickelt, um Daten aus ohsome zu extrahieren und zu verarbeiten.

2. Erstellung von Prototypen:

* Prototyp 1: Die Entwicklung eines ersten Prototyps wurde in einem kleinen Testgebiet durchgeführt, bei dem Änderungen der zulässigen Höchstgeschwindigkeit (maxspeed) zum 1. Januar eines Jahres ausgewertet wurden.

* https://radverkehr.github.io/digitalisierungsmodul/SoSe2024/viz/historic_osm-maxspeed_hermann-str.html

* Prototyp 2: Ein weiterer Prototyp stellt Geschwindigkeitsänderungen (maxspeed) auf Hauptverkehrsstraßen im Berliner Bezirk Neukölln dar. Die Visualisierung enthält einen Zeitstrahl zur Darstellung der Änderungen über einen längeren Zeitraum sowie eine Verlinkung der Way-Historie.

* https://radverkehr.github.io/osm_history_analysis/viz/maps_nk_no-basemap_PrimSecTert_wayChart_24-11-04.html

3. Ein Benchmarking soll zeigen, wie gut die Methode funktioniert und ob bzw. welche Probleme auftreten.

* Beispielsweise wird untersucht, wie schnell bzw. ob neue Fahrradstraßen in OSM getaggt wurden. Dazu werden als Vergleich die amtlichen Daten zu Fahrradstraßen der Berliner Senatsverwaltung [2] herangezogen und mit dem Zeitpunkt der Änderung in OSM (bicycle_road=yes) verglichen.

* Zusätzlich soll untersucht werden, ob eine Reduktion der zulässigen Geschwindigkeit oder die Einrichtung separater Radinfrastrukturen zu einem Rückgang der Fahrradunfälle in den jeweiligen Straßenabschnitten führt. Grundlage hierfür sind Daten aus dem Unfallatlas [3].

Herausforderungen bei der Umsetzung:

* Unterscheidung von Änderungen und Maßnahmen: Wann handelt es sich bei einer Änderung in OSM um eine tatsächliche verkehrliche Maßnahme, und wann nur um eine Korrektur, Verfeinerung oder ein Missverständnis?

* Umgang mit Geometrieänderungen: Änderungen können zu veränderten Geometrien führen (z.B. durch Aufteilung oder Ergänzung von Straßenabschnitten), was Auswirkungen auf die entsprechenden Way-IDs haben kann und die eindeutige Nachverfolgbarkeit erschwert.

[1] https://docs.ohsome.org/ohsome-api/v1/

[2] https://gdi.berlin.de/services/wfs/fahrradstrassen

[3] https://unfallatlas.statistikportal.de/

Erfahrungsbericht: MerginMaps für die Inventarisierung von Vogelnestern

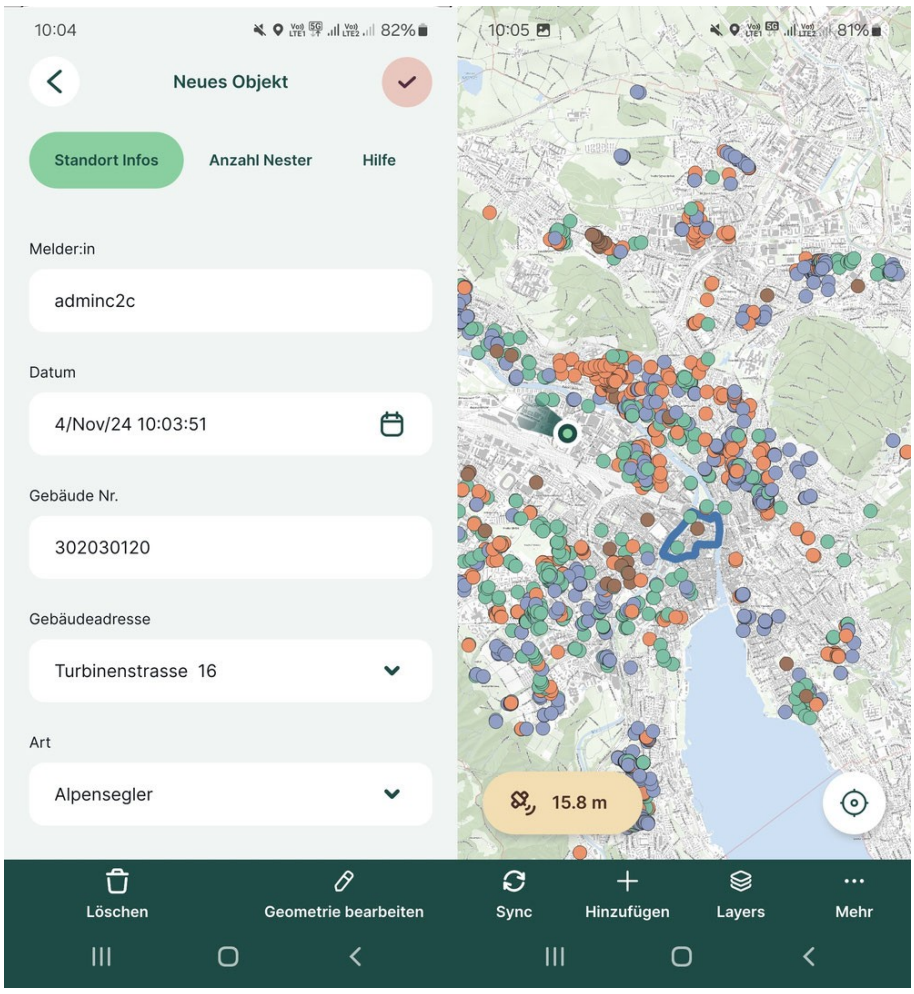
Geneviève Hannes

In Zürich steht eine Gebäudebrüterinventarisierung an und diese soll in einem knappen Zeitrahmen mit einer mobilen App erfolgen. MerginMaps ermöglicht es uns, dank der Integration mit QGIS eine komplexe Feldaufnahmelogik schnell umsetzen. Der Bestand wird somit direkt vor Ort mit der Mobilen App erfasst, während die Daten zentral in unserer eigenen Cloud-Infrastruktur gespeichert werden. Eine innovative Lösung, die auf OpenSource Komponenten basiert.

In unserem Vortrag stellen wir vor, wie und warum wir Mergin Maps für die Inventarisierung von Gebäudebrütern einsetzen. Unser Ziel war es, eine effiziente Lösung zu finden, die das Erfassen von Geodaten vor Ort ermöglicht, und diese sollte zeitnah in Betrieb genommen werden können. Die Feldaufnahmen werden von rund 70 Freiwilligen durchgeführt und somit soll die App möglichst leicht und intuitiv zu bedienen sein.

Wichtig war für uns, dass wir die den Mergin Maps Server in unserer eigenen Cloud-Plattform hosten können. In einem Erfahrungsbericht teilen wir die Herausforderungen und Erfolge, die uns bei der Einrichtung und dem Betrieb dieser Self-Hosting-Lösung begegnet sind.

Ausserdem zeigen wir, wie eine komplexe Feldaufnahmelogik einfach in QGIS umgesetzt und dann dank einem Plugin direkt für MerginMaps verwendet werden kann



Unser technischer Stack besteht aus Open-Source-Komponenten: QGIS als GIS-Werkzeug, PostGIS/PostgreSQL und die Mergin Maps App für die mobile Datenerfassung. Diese Kombination ermöglicht es, die Gebäudebrüter-Daten zuverlässig und sicher zu erfassen, zu verwalten und jederzeit zu synchronisieren.

OpenStreetMap FeatureInfo

Jakob Miksch

Der Vortrag stellt eine Implementierung vor, die es erlaubt Informationen zu OSM Objekten um einen Ort herum über eine einfache Schnittstelle abzufragen. Die Idee dahinter ist, das Rad nicht neu zu erfinden, sondern bewährte Software-Komponenten zu kombinieren und auf etablierte Standards zu setzen.

Die Daten werden mit osm2pgsql in eine PostgreSQL/PostGIS Datenbank geladen und kontinuierlich auf dem neusten Stand gehalten. Die Abfrage wird mit einer parametrisierten SQL-Abfrage durchgeführt und das Ergebnis mittels pg_featureserv über eine REST-Schnittstelle verfügbar gemacht. Weitere Ideen sind die Schnittstelle konform zu OGC API Processes zu machen. Eine mögliche Anwendung wäre die Nutzung als Backend für Objektabfrage in der offiziellen OpenStreetMap Webseite.

Pünktlich zur Uni? Bus-Tracking mit der EFA JSON API

Christoph Friedrich

Im "Bahnland Bayern" sind Echtzeit-Mobilitätsdaten über eine Schnittstelle namens "EFA JSON API" verfügbar. Indem man diese mitschneidet, kann eine Pünktlichkeitsstatistik z.B. der Würzburger Stadtbusse erstellt werden, die aufgrund der nachwievor fehlenden Straßenbahnanbindung der Uni gerade zum morgendlichen Vorlesungsbeginn nicht prickelnd aussieht.

Es wird eine kurze Übersicht über die API gegeben, dann der Ausleseprozess erklärt und was die Schwierigkeiten dabei waren, und anschließend die Ergebnisse der Analyse präsentiert. Insgesamt angelehnt an den Vortrag "BahnMining" beim CCC vor fünf Jahren.

2.5D Indoor-Karten auf Basis von OpenStreetMap-Daten

Richard Karl Fuchs, Danny Schober, Jacques-Maurice Walther und Jan Schmalfuß-Schwarz

Technische Universität Dresden

Motivation

Ein zentraler digitaler Info-Punkt kann in einem Gebäude aufgestellt werden, um bei der Navigation durch das Gebäude und beim Auffinden von Räumen zu unterstützen. Dafür kann auf eine Vielzahl digitaler Karten- und Navigationsanwendungen zurückgegriffen werden, die für blinde oder sehbeeinträchtigte Anwender:innen jedoch häufig nicht nutzbar sind [1]. Durch geeignete Darstellungsformen sowie die Verbindung mehrerer Ausgabemodalitäten, etwa der Kombination eines visuelles Displays mit einer haptischen Stiftplatte, könnte allerdings ein Info-Punkt mit einem hohen Maß an Barrierefreiheit entstehen, der die Navigation für unterschiedliche Nutzer:innen-Gruppen ermöglicht, einschließlich Personen mit Farbfehlsichtigkeiten, Blindheit oder Gehbehinderungen.

Die vorliegende Arbeit basiert auf einer studentischen Projektarbeit mit dem Ziel, einen barrierefreien Info-Punkt zu realisieren, um sehende und sehbeeinträchtigte Personen bei der Navigation durch das Gebäude zu unterstützen. Um die Wegfindung auf dem visuellen Display zu erleichtern, wird dafür zunächst eine 2.5D-Ansicht der Gebäudeinformationen umgesetzt. In dieser können mehrere Etagen übereinander angezeigt werden und die Höheninformation wird durch eine perspektivische Darstellung angedeutet. Dadurch kann häufiges Hin-und-Herschalten zwischen einzelnen Etagen vermieden und die Navigation über mehrere Ebenen hinweg übersichtlicher gestaltet werden. Für die 2.5D-Ansicht ergibt sich eine Vielzahl an Entwurfsentscheidungen, etwa zum Abstand der Etagen, der Transparenz von Räumen und Etagen oder der Darstellung von verbindenden Elementen, wie Treppen und Aufzügen. Diese werden in der vorliegenden Arbeit beschrieben und ggf. bestehende alternative Umsetzungen werden vorgestellt. Die Benutzeroberfläche wird für die Unterstützung der 2.5D-Ansicht und das Umschalten von dieser Darstellung zur 2D-Ansicht angepasst. Weiterhin sind bei der Oberflächengestaltung des visuellen Displays grundlegende Aspekte der Barrierefreiheit zu berücksichtigen, um die Interaktion für Nutzende mit Farbfehlsichtigkeit zu ermöglichen. Um sicherzustellen, dass auch Nutzer:innen im Rollstuhl alle Schaltflächen des Touch-Displays erreichen können, wird ein Rollstuhlmodus realisiert.

Aus dem Ziel, einen Info-Punkt zu implementieren, der auch für blinde oder sehbeeinträchtigte Personen nutzbar ist, ergibt sich die Notwendigkeit, die Kartendarstellung auf dem Monitor um eine haptische Repräsentation, zum Beispiel in Form einer Stiftplatte mit ausfahrbaren Pins, zu ergänzen. Dadurch können die Gebäude- und Raumumrisse, die Position von Wänden und Türen, aber auch Hinweise und Raumnummern in Braille-Darstellung ertastet werden. Damit der Info-Punkt gleichzeitig durch sehende und seehbeinträchtigte Nutzer:innen bedient werden kann, ist jeweils hervorzuheben, welcher Karten-Ausschnitt auf der Stiftplatte dargestellt wird, und diese wiederum ist zu aktualisieren, wenn sich der Ausschnitt auf dem visuellen Display verändert. Dieser Ansatz wurde im Rahmen der studentischen Arbeit diskutiert, jedoch nicht in der vorliegenden Arbeit aufgegriffen, da sich diese der Darstelung einer Karte in 2.5D widmet.

Als Grundlage dient die bestehende Anwendung *Mapable*[5] des Forschungsprojekt „Accessible Maps“[6] der Technischen Universität Dresden. Die bestehende Code-Basis wurde angepasst, um die Weite-

5 Demo: https://accessiblemaps.de:3001/, Quellcode: https://github.com/AccessibleMaps/Mapable (je Besucht am 23.09.2024)

6 Accessible Maps – Barrierefreie Karten zur Verbesserung der Mobilität im Beruf. URL: https://tu-dresden.de/ing/informatik/ai/mci/forschung/forschungsgebiete/accessible-maps-barrierefreie-karten-zur-verbesserung-der-mobilitaet-im-beruf (Besucht am 23.09.2024)

rentwicklung und Integration der beschriebenen Funktionen zu ermöglichen. Als Datenformat wird dabei das *Simple Indoor Tagging* Schema vorausgesetzt.

Forschungshintergrund

Die Darstellungsweise der Karte und ihre Übereinstimmung mit der tatsächlichen Umgebung ist essentiell, damit die Betrachter:innen ein solides Verständnis der räumlichen Lage erhalten und damit die Navigation auch in unbekannten Umgebungen ermöglicht wird [3]. Zheng und Hsu beschreiben, dass der Betrachtungswinkel vieler 3D-Karten zu Verdeckungen führe und es Nutzer:innen mitunter schwer falle, die Entfernungen einzuschätzen. Eine 2D-Darstellung, die das räumliche Layout nur in einer Draufsicht abbildet, muss durch die Nutzenden hingegen zuerst in ein dreidimensionales mentales Modell übertragen werden, um es zur Navigation verwenden zu können [3]. Eine Alternative stellen 2.5D-Karten dar, die die Höheninformation aus einer festen Perspektive abbilden. Zheng und Hsu beschreiben diese Darstellungsform wie folgt: „By defining the map user as the origin of the map's coordinate system, the 2.5D method employs perspective drawing techniques to present direction and depth on the map“ [3]. In Bezug auf den geeigneten Betrachtungswinkel stellen die Forschenden fest, dass eine 2.5D-Karte mit im Winkel von 60° ein schnelleres Identifizieren der geeigneten Routen ermöglichte als die Darstellungen mit 30° und 90°: „[...] 60° maps allowed superior performance in simulated wayfinding (avoiding confusion and hesitation, doubling back, and going off-route), a higher consistency between the taken route and the cognitive map, and a lower orientation error rate“ [3]. In einer weiteren Untersuchung im Rahmen ihrer Studie zeigte, dass eine 2.5D-Karte, bei der weniger relevante oder verdeckende Gebäude transparent dargestellt werden und weiterhin einige Texturen der Häuserfronten in die Kartendarstellung übernommen werden, Vorteile gegenüber nicht-transparenten Darstellungen und solchen ohne Textur-Details bieten. Die Umsetzung mit Transparenz und Gebäude-Datails ermöglichte eine schnellere Wegfindung und reduzierte die wahrgenommene Verwirrung der Proband:innen [3]. Neben einer Empfehlung für diese Darstellung fassen die Autor:innen zusammen, dass wesentliche Orientierungspunkte an Kreuzungen hervorgehoben werden sollten, um die Navigation zu erleichtern [3].

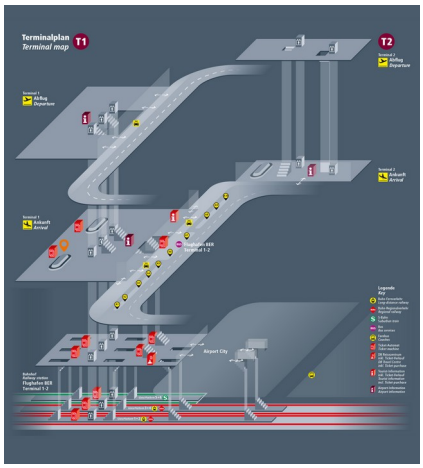


Abb. 1: 2,5D-Karte der Terminals am Flughafen Berlin Brandenburg, https://flughafenexpress.deutsche-bahn.com/fex-de/Ihr-Weg-zum-Terminal-5240798 (Besucht am 24.9.2024)

Zheng und Hsu untersuchten die geeignete Darstellung von Stadtplänen, wie Riedinger et al. feststellten, sind 3D-Pläne darüber hinaus in einer Vielzahl von Anwendungsbereichen verbreitet, etwa der Architektur, Planung und Industrie [2]. In Detail-reduzierter Darstellung, die mehrere Etagen und deren Verbindungen abbildet, werden 2.5D-Karten auch in Einkaufszentren und Flughäfen verwendet, siehe dazu Abb. 1 des Flughafens Berlin Brandenburg. Diese Darstellungsform, die die Verbindung der einzelnen Etagen hervorhebt, wichtige Informationen durch Symbole und Labels darstellt und einige Details, unter anderem Zwischenwände, ausspart, stellt aufgrund der schnellen Erfassbarkeit die Grundlage für die 2.5D-Darstellung der Gebäudepläne im Rahmen dieser Arbeit dar.

Darstellung der Stockwerke in 2.5D

Im Gegensatz zu einer 2D-Ansicht werden bei der 2.5D-Ansicht die Stockwerke im 3D-Raum übereinander geschichtet dargestellt, bleiben jedoch flach. Eine native Unterstützung einer 3D-Ansicht und der benötigten Interaktionsmöglichkeiten wie Kippen und Schwenken ist für das verwendete Karten-Framework dabei essentiell.

Das in *Mapable* eingesetzte *Leaflet* erfüllt diese Bedingungen nicht, weswegen diese Arbeit den Wechsel zum Open-Source Projekt *Maptalks*[7] macht. *Maptalks* unterstützt nativ das Platzieren von flachen Objekten auf verschiedenen Höhen, welche basierend auf der Distanz zur Kamera korrekt dargestellt werden.

In einer 2.5D-Ansicht kann stets nur ein Ausschnitt aller Stockwerke gezeigt werden, da ansonsten die Lesbarkeit und Benutzbarkeit stark eingeschränkt wäre. Deswegen wird sich in dieser Anwendung auf drei zu einer Zeit sichtbare Stockwerke beschränkt: Das ausgewählte Stockwerk und gegebenenfalls ein Stockwerk darüber und darunter. In Gebäuden mit vielen Stockwerken könnte dies zu einer Verminderung der Intuition führen, jedoch haben diese meist einen ähnlichen Aufbau zwischen Stockwerken, insbesondere im Bezug auf die Position von Verbindungen zwischen Etagen. Der Abstand zwischen zwei Stockwerken ist dabei statisch auf 12m festgesetzt, damit die Überdeckung durch die nächsthöheren Stockwerke möglichst gering ausfällt, genauso ist der Winkel der Kamera zur Oberfläche fest gesetzt. Diese statischen Werten sollten allerdings je nach Form und Grundfläche eines Gebäudes reevaluiert und gegebenenfalls angepasst werden, um optimale Sichtbarkeit zu erzielen.

Beim Wechsel zwischen Stockwerken und beim Übergang von der klassischen 2D-Draufsicht in die 2.5D-Ansicht muss gewährleistet werden, dass die Nutzenden ihre Orientierung zu jedem Zeitpunkt behalten. Die Anwendung greift dafür auf Animationen zurück, welche die räumlichen Zusammenhänge und Veränderungen hervorheben und damit die Bedienung erleichtern sollen. Beim Wechsel von der 2D- in die 2.5D-Sicht wird sowohl die Kameraperspektive angepasst, als auch die ausgewählte Ebene in das zentrale Sichtfeld gehoben, sodass mindestens eine weitere anliegende Ebene darunter als auch darüber angedeutet werden kann. Beim Wechsel zwischen den Ebenen wird die neu ausgewählte Ebene opak und in das Sichtfeld bewegt. Anliegende, bisher nicht sichtbare Ebenen werden eingeblendet, während zuvor sichtbare Ebenen, die nun nicht mehr gezeigt werden sollen, ausgeblendet werden.

Werden mehrere Ebenen gleichzeitig und teilweise sich gegenseitig überlagernd dargestellt, so treten visuelle Überdeckungen auf und die Darstellung ist zu überladen, um einfach damit arbeiten zu können.

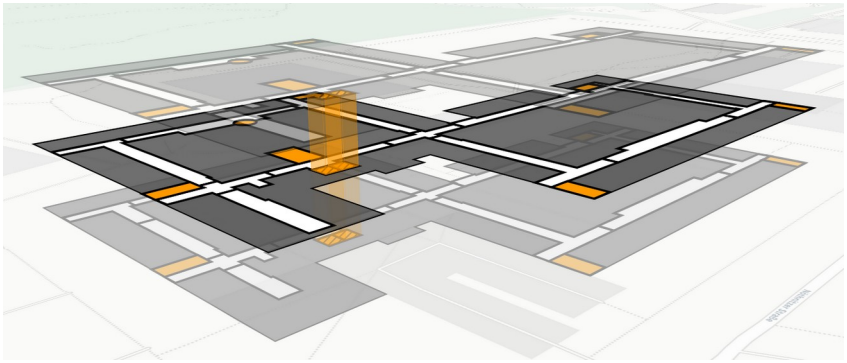


Abb. 2: Ebenendarstellung mit reduzierten Informationen

Auch dient die Darstellung der Ebenen in 2.5D primär der Orientierung, während das Vermitteln von speziellen Informationen zu ausgewählten Räumen in Kontrast zur 2D-Sicht nebensächlich ist. Deswegen wurde versucht, die Ebenendarstellung in der 2.5D-Sicht so einfach wie möglich zu gestalten und es wurde sich allein auf notwendige Informationen beschränkt, wobei der Fokus auf den vertikalen Verbindungen zwischen Stockwerken liegt. Somit gibt es keine Unterscheidung zwischen einzelnen Räumen, sondern lediglich zwischen den Fluren, den Räumen als Gesamtes und einem potentiell ausgewählten Raum. Die Darstellung ist in Abbildung 2 zu sehen.

7 URL: https://github.com/maptalks/maptalks-canvas (Besucht am 01.02.2025)

Anzeige von Treppenhäusern und freistehenden Treppen

Die 2.5D-Darstellung eines Gebäudes in der Kartenanwendung bietet den Nutzenden einen klaren Mehrwert, indem sie wichtige Strukturen wie Treppenhäuser, Aufzüge und freistehende Treppen zwischen den Ebenen in 3D anzeigt. Diese Elemente ermöglichen eine intuitive Orientierung und erleichtern die Navigation durch das Gebäude, da die vertikale Bewegung zwischen den Stockwerken anschaulich dargestellt wird. Besonders in komplexen Gebäuden, wie Einkaufszentren oder Bürogebäuden, hilft dies den Nutzenden, schnell den optimalen Weg zu finden, z. B. um schnell zum nächsten Aufzug oder Treppenhaus zu gelangen. Dies verbessert nicht nur die Nutzendenerfahrung, sondern spart auch wertvolle Zeit bei der Orientierung.

Das verwendete Framework *Maptalks* ermöglicht es, Polygonen und anderen Formen eine Höhe zu geben, allerdings ist es nicht möglich, einzelne Punkte auf unterschiedliche Höhen zu setzen. Ebenfalls können keine vertikalen Linien angezeigt werden, welche nicht bis auf die Grundlinie gehen. Eine Möglichkeit, diese Limierungen zu umgehen, bietet die Integrierung von *THREE.js*[8] in *Maptalks*.

Mit *THREE.js* können beliebige geometrische Objekte in 3D dargestellt werden. Die Integrierung erlaubt es ebenfalls aus *Maptalks*-Polygonen durch Extrusion Prismen zu erstellen. Um einzelne Koordinaten zu verändern, müssen die Knoten jedoch modifiziert werden, der Code dafür wurde aus einem Beispiel übernommen[9].
Treppenhäuser und Aufzüge (In der Anwendung „simple staircases“ genannt) lassen sich in der Regel durch die Extrusion der Grundfläche als Prismen darstellen, was in Abbildung 3 zu sehen ist. Da die Transparenz es erschwert, verschiedene Objekte zu unterscheiden, sind an den Kanten zusätzliche Zylinder angefügt, um die vertikalen Abgrenzungen deutlich zu machen.

Freistehende Treppen (In der Anwendung „complex staircases“ genannt) sind dagegen schwieriger zu implementieren. In *OpenStreetMap* müssen dabei neben den Umrissen der Treppen, welche für die Darstellung in 2D verwendet werden, auch die Mittellinien der Treppen hinterlegt sein. Diese Linie liegt genau in der Mitte der Treppe und speichert die Breite der Treppenstufen als OSM-Attribut. Ebenfalls sind an den Punkten der Linie gegebenenfalls weitere Informationen wie Stockwerk dieses Punktes gegeben, somit kann die Höhe der Treppe akkurat wiedergegeben werden. Wenn die Ebenen-Information nicht vorhanden ist, wird zwischen den nächsten gegebenen Punkten interpoliert, wodurch gleichmäßig steigende Treppen erreicht werden können.

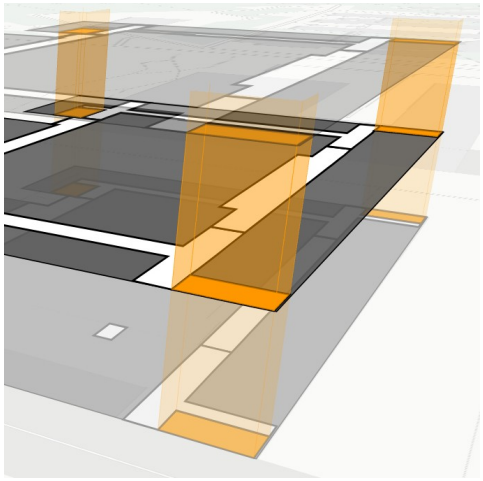


Abb. 3: Treppenhäuser in der Informatik-Fakultät der TU-Dresden

Da nur die Mittellinie der Treppe und nicht die linke und rechte Kante der Treppenstufen gegeben sind, müssen diese zunächst extrapoliert werden. Einfach ausgedrückt muss die Linie also Segment für Segment senkrecht verschoben werden, wobei die Segmente mit einander verbunden bleiben müssen. Für jede Treppendarstellung muss die Mittellinie mehrfach verschoben werden. Einmal je nach rechts und links für die Breite der Stufen, und je einmal für die Breite des Geländers, welches standardmäßig eine Breite von 5 cm hat. Mithilfe dieser Linien können die verschiedenen Prismen erzeugt werden, sowohl die vertikal-flachen für die Stufen als auch die horizontal-flachen für die Geländer.

Die entstandenen Treppen sind in Abbildungen 4 und 5 zu sehen.

8 URL: https://threejs.org/
9 Siehe https://maptalks.org/maptalks.three/demo/custom-bridge.html (Besucht am 21.10.2024)

Fazit und Ausblick

In der vorliegenden Arbeit wurden zunächst die Motivation und Teilziele vorgestellt sowie eine Auswahl an Publikationen referenziert, die die Vorteile von 2.5D-Karten hervorheben. Die Umsetzung der Kartendarstellung wurde angepasst, um eine 2.5D-Darstellung von mehreren Stockwerken zu ermöglichen. Die Darstellung der Aufzüge und Treppen in 3D könnte die Navigation für eine Vielzahl unterschiedlicher Nutzer:innen-Gruppen erleichtern. Um eine gute Übersichtlichkeit zu gewährleisten, wurde auf eine vereinfachte Ebenendarstellung geachtet, Wechsel zwischen Ebenen oder Perspektiven werden außerdem durch Animationen verdeutlicht.

Die vorliegende Arbeit kann als Grundlage für die praktische Umsetzung des barrierefreien Infopunkts dienen, der die Navigation für eine Vielzahl von Nutzer:innen ermöglichen kann und insbesondere Unterstützung für Personen mit Gehbehinderungen, Farbfehlsichtigkeiten, Sehschwächen oder Blindheit bietet. Da im Rahmen dieser Arbeit keine Nutzer:innen-Tests durchgeführt wurden, erscheint es jedoch zunächst relevant, die Nutzbarkeit und die Oberflächengestaltung im Rahmen anknüpfender Forschungsarbeiten zu evaluieren. Dafür eignet sich eine qualitative Proband:innen-Studie, die die Gebrauchstauglichkeit untersucht, zum Beispiel durch Verwendung des System-Usability-Scale.

Weiterführende Forschungsansätze zur Weiterentwicklung der bestehenden Anwendung könnten unter anderem das Anzeigen einer Route zwischen der aktuellen Position und dem Zielpunkt umfassen, um die Wegfindung zu erleichtern. Im Rahmen der vorliegenden Arbeit wurde die 2.5D-Darstellung außerdem nur für den Andreas-Pfitzmann-Bau und das Hörsaalzentrum der Technischen Universität Dresden umgesetzt und getestet. Um auch die Geoinformationen anderer Gebäude in geeigneter Form darstellen zu können, erscheint es zunächst empfehlenswert, die Gebäudeansicht so zu rotieren, dass die Außenkanten parallel zum Bildschirmrand ausgerichtet sind. Dies kann sowohl in der 2D- als auch in der 2.5D-Ansicht eine platz effiziente und verständnisfördernde Darstellung unterstützen. Bei Gebäuden mit einer außergewöhnlich hohen Anzahl an Etagen kann es weiterhin empfehlenswert sein, die 2,5D-Darstellung dahingehend anzupassen, dass die einzelnen Ebenen mit ausreichendem vertikalen Abstand angezeigt werden, um ein Verdecken von Details zu verhindern. Weiterhin ist anzumerken, dass in einigen Gebäuden Räume auch dann zu einer Etage gezählt werden, wenn nicht alle Räume plan mit dem Stockwerk ausgerichtet sind und dadurch sozusagen halbe Stockwerke bilden. Ein Beispiel dafür ist der Georg-Schumann-Bau der TU Dresden (ehemals Landgericht am Münchner Platz), in welchem sich Ebenen auf halber Höhe der darunterliegenden Etage befinden, wodurch sich eine Stapelung alle 0,5 Stockwerke ergibt. Dies ist in der aktuellen Anwendung jedoch noch nicht darstellbar und ohne größere Umplanung der Struktur nicht leicht umzusetzen.

Kontakt zum Autor:

(Herr) Richard Karl Fuchs
TU Dresden
Richard_Karl.Fuchs@mailbox.tu-dresden.de

Literatur

[1] J. Ducasse, M. Macé, and C. Jouffrais, ‘FROM OPEN GEOGRAPHICAL DATA TO TANGIBLE MAPS: IMPROVING THE ACCESSIBILITY OF MAPS FOR VISUALLY IMPAIRED PEOPLE’, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XL-3-W3, pp. 517–523, 2015, doi: 10.5194/isprsarchives-XL-3-W3-517-2015.

[2] C. Riedinger, M. Jordan, and H. Tabia, ‘3D models over the centuries: From old floor plans to 3D representation’, in 2014 International Conference on 3D Imaging (IC3D), 2014, pp. 1–8. doi: 10.1109/IC3D.2014.7032583.

[3] M.-C. Zheng and Y.-W. Hsu, ‘How 2.5D Maps Design Improve the Wayfinding Performance and Spatial Ability of Map Users’, Informatics, vol. 8, no. 4, 2021, doi: 10.3390/informatics8040088.

Sheepsmeadow: Agentenbasierte Modellierung einfach erlernen

Michael Brüggemann

'Sheepsmeadow' ist ein in Java entwickeltes offenes Simulationsprogramm, das den Nutzer:innen das Themengebiet "Agentenbasierte Modellierung" (ABM) spielerrisch näher bringen soll. Agenten sind Wölfe und Schafe, die miteinander auf einer Wiese interagieren.

'Sheepsmeadow' gibt zudem die Möglichkeit, eigene Entscheidungsregeln (Actions) direkt im Quellcode zu ergänzen und so die Simulation zu verändern. Es soll dabei helfen, die Entwicklung von ABMs mithilfe des MASON Frameworks zu erlernen.

Agentenbasierte Modelle (ABM) bieten die Möglichkeit, komplexes Zusammenspiel einzelner Individuen mit ihrem Umfeld darzustellen. Ein Agent ist hier ein Individuum, welches eigene Ziele (eine Agenda) verfolgt. Agenten können in solchen Modellen ganz unterschiedliche Dinge sein (Menschen, Tiere, Teilchen, Autos, Pflanzen). Die Agenten haben hierbei individuelle Charakteristiken, wodurch die Gesamtheit der Agenten heterogen dargestellt werden kann.

Solche Systeme zu analysieren ist vor allem im Bereich der Geoinformatik/Landschaftsökologie interessant, da so das räumliche Zusammenspiel der Agenten (auf einem Raster) beobachtet werden kann. Agenten haben dabei eine Position als Pixelkoordinate und können nur mit Agenten in ihrer direkten Nachbarschaft interagieren.

'Sheepsmeadow' stellt ein solches Agentenbasiertes Modell dar. In der Simulation interagieren Wölfe und Schafe (die Agenten) miteinander und mit ihrem Umfeld auf einem Raster (der Wiese) über sog. "Actions". Eine "Action" stellt eine Regel dar, die das Verhalten des Agenten darstellt. Diese Regeln sind an das Eintreten besonderer Bedingungen geknüpft.

Die Simulation als ganzes wird über seine "Modellparameter" gesteuert. Diese bestimmen u.a. die Größe des Rasters auf dem sich die Agenten bewegen, die Anzahl der Wölfe und Schafe oder deren Fortpflanzungswahrscheinlichkeit.

'Sheepsmeadow' erlaubt einfaches und intuitives Anpassen der Parameter des Modells, wodurch die Nutzer:innen direkt erleben können, wie sich diese Veränderungen auf das System der Agenten auf der Wiese auswirkt. Hier kann z.B. analysiert werden, ab welchem Verhältnis von Wölfen zu Schafen, die Schafspopulation nicht fortbestehen kann. So können Personen mit wenig bis keinen Programmierkentnissen leicht unterschiedliche Konfigurationen des Modells erstellen und die Veränderungen im Verhalten der Agenten beobacheten.

Ein weiterer Lernzweck ist es, das Nutzer:innen ihre Kenntnisse in der Programmiersprache "Java" zu erweitern. Durch den offenen Quellcode haben Nutzer:innen zudem die Chance, selbst neue "Actions" zu ergänzen. Hierfür stellt 'Sheepsmeadow' ein einfaches Interface in Java zur Verfügung. Dieser Teil von 'Sheepsmeadow' richtet sich an Personen, die bereits in einer (anderen) Programmiersprache programmieren können, welche aber erlernen möchten wie man ein Agentenbasiertes Modell in einer Software mit UI implementiert. Ebenso bietet 'Sheepsmeadow' die Möglichkeit Objektorientiere Entwurfskonzepte zu erlernen/anzuwenden.

Eine alternative Plattform, um Agentenbasierte Modellierung zu erlernen wäre "NetLogo". NetLogo bietet bereits eine Wolfs-Schafs-Simulation (WolfSheepPredation). Diese kann ebenso um neue Entscheidungsregeln erweitert werden (wenn die Programmierkenntnisse dafür vorhanden sind). 'Sheepsmeadow' zielt hier darauf ab, dass sich Nutzer:innen mit dem Modellierungsframework MASON auseinadersetzen und erlernen, wie man mithilfe dieses Frameworks eigene Modelle erstellen kann/ ein bestehendes Modell erweitern kann.

Sheepsmeadow: Agentenbasierte Modellierung einfach erlernen

Durch die Einstellbarkeit der Modelparameter und der Möglichkeit, eigene Actions im Quellcode zu integrieren, eignet sich 'Sheepsmeadow' perfekt als Lernsoftware, um unterschiedliche Facetten der Agentenbasierten Modellierung zu erlernen (Nutzen des Modells, Parametrisieren des Modells, Implementieren neuer Logiken, Analyse der Simulationsergebnisse).

Als Maintainer des Repositories trete ich dafür ein, dass das Projekt unter offenen Standards (weiter-)entwickelt wird. Das Repository ist hier zu finden:

https://github.com/MichaelBrueggemann/sheepsmeadow.

XPlanung mit Open Source Software

Torsten Friebe[1], Julian Zilz[1]

[1] lat/lon GmbH, Wachtberg

Abstract

Im April 2022 wurde der Quellcode der Software xPlanBox der Firma lat/lon (https://www.lat-lon.de/) im Rahmen eines Pilotprojekts auf der openCode-Plattform des Zentrum für Digitale Souveränität der Öffentlichen Verwaltung (ZenDiS) veröffentlicht. Die Software liegt nun als Open-Source-Lösung vor und kommt im Rahmen des Onlinezugangsgesetz (OZG) und des "Einer-für-Alle"-Prinzips (EfA) zum Einsatz. Dieser Beitrag stellt kurz die openCodePlattform sowie deren Funktionen vor. Auf der FOSSGIS-Konferenz 2024 wurde das Projekt und die Einsatzmöglichkeiten der Software bereits präsentiert. In diesem Jahr sollen die Neuerungen mit der Version 8.0 präsentiert werden und welche Möglichkeiten sich für den Betrieb auf einer Kubernetes-Plattform ergeben.

Schlüsselwörter: openCode, Open Source, XPlanung, XPlanGML, digitale Souveränität, OZG, EfA, deegree, xPlanBox, GIS, WMS, WFS, INSPIRE

Was ist die openCode-Plattform des ZenDiS?

Die openCode-Plattform (https://opencode.de/de) des ZenDiS ist eine Plattform, auf der Open-Source-Software-Projekte entwickelt und bereitgestellt werden, die im Zusammenhang mit allen Ebenen der öffentlichen Verwaltung stehen. openCode wurde durch das Bundesministerium des Innern und für Heimat (BMI) sowie den Ländern Baden-Württemberg und Nordrhein-Westfalen initiiert und wird seit Januar 2024 durch ZenDiS betrieben.

Die Plattform bietet eine Vielzahl von Vorteilen für Entwickler und Nutzer, die Open-Source-Software für die deutsche Verwaltung schreiben. Sie beinhaltet neben einer GitLab-Instanz für die Quellcodeverwaltung, einen Wissensbereich und ein Diskussionsforum für Nutzer sowie ein Softwareverzeichnis mit den Projekten und Lösungen der öffentlichen Verwaltung, die zur Nachnutzung bereitstehen (vgl. Abb. 1).

Um die rechtssichere Bereitstellung und Nachnutzung des Quellcodes zu ermöglichen, müssen alle Inhalte auf openCode klare lizenzrechtliche Anforderungen erfüllen; siehe https://opencode.de/de/wissen/softwareverzeichnis-einsatz-in-der-verwaltung/richtlinie-zum-softwareverzeichnis und https://opencode.de/faq#Nutzung-und-Rechtliches [1,2,3].

Was sind die Vorteile der openCode-Plattform?

Insgesamt bietet die openCode-Plattform des ZenDiS eine ideale Umgebung für die Entwicklung und Bereitstellung von Open-Source-Software für die öffentliche Verwaltung, indem es die Zusammenarbeit und Kommunikation zwischen Entwicklern und Nutzern, die größtenteils aus der öffentlichen Verwaltung sind, erleichtert und die Transparenz und Sicherheit der Software erhöht. Mit einer Anzahl an 5.300 Nutzenden in über 2.200 Projekten (Stand 15.01.2025), was einer Zunahme von 100 Prozent in den letzten 12 Monaten entspricht, weißt die Plattform zudem eine steigende User-Community auf [2].

Einer der wichtigsten Vorteile der Plattform ist, dass Richtlinien definiert sind, in welchem Stadium sich eine Software befindet. Durch die im Projekt zu hinterlegende *publiccode.yml* (https://yml.publiccode.tools/) Datei, einem Metadaten Standard für Software der öffentlichen Verwaltungen, werden wich-

tige Informationen über das Stadium der Software, deren Einsatzbereich und weitere Informationen zur Qualität und dem Reifegrad der Software bereitgestellt [4]. Somit stehen potentiellen Anwendern diese Informationen über das Software-Verzeichnis zur Verfügung und können über die Plattform direkt mit anderen Nutzern in Kontakt treten.

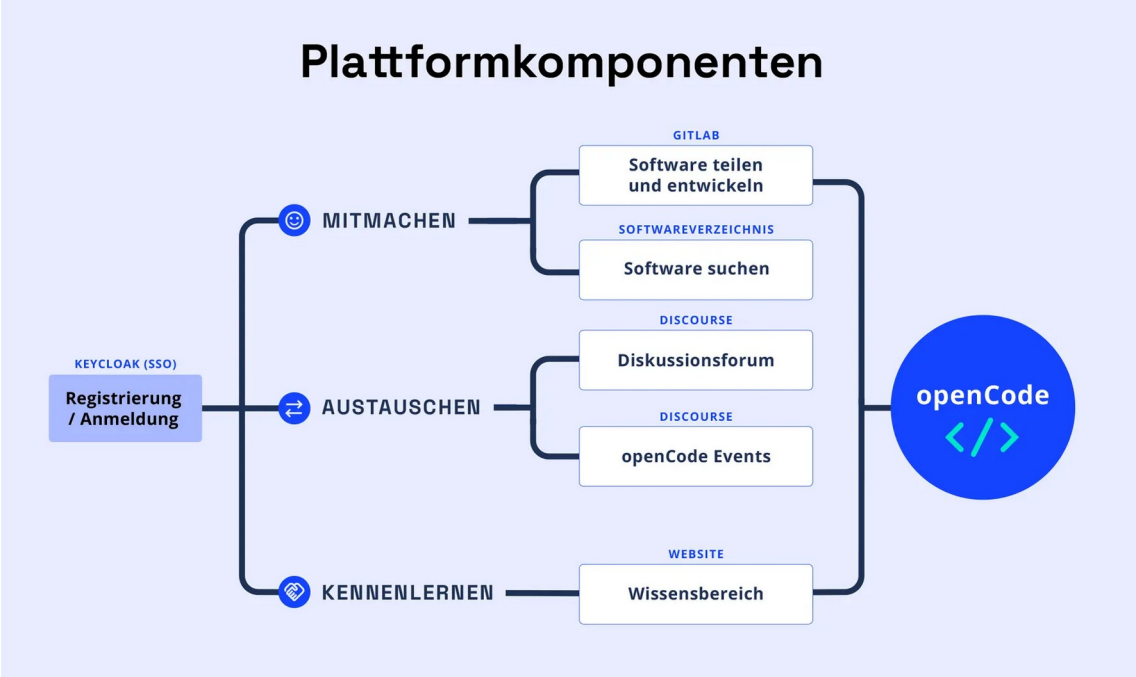


Abbildung 1: Darstellung Plattformkomponenten (Quelle: https://opencode.de/_next/image?url=https%3A%2F%2Fcdn.opencode.de%2Fwp-content%2Fuploads%2F2025%2F01%2Ffbd-zendis-infografik-plattformkomponenten.png&w=1920&q=75)

Mit der Veröffentlichung der Software xPlanBox auf der openCode-Plattform unter https://gitlab.opencode.de/diplanung profitiert das Projekt neben den lizenzrechtlichen Rahmenbedingungen auch von der Möglichkeit über das Software-Verzeichnis von Nutzern, Anwendern und Entwicklern, die speziell zu dem Themengebiet suchen, gefunden zu werden. Diese fach- und zielgruppenspezifische Umgebung waren ausschlaggebend für die Entscheidung, den Quellcode auf openCode zu veröffentlichen und nicht auf GitHub, Bitbucket oder anderen in der Entwickler-Community genutzten Plattformen. Da die Komponenten der xPlanBox auch im Rahmen des OZG-Umsetzungsprojekte "Bürgerbeteiligung und Information" aus dem Themenfeld "Bauen & Wohnen" für die Bereitstellung von räumlichen Planwerken im Internet verwendet und perspektivisch über den FIT-Store der FITKO (https://www.fitko.de/produktmanagement/fit-store) bereitgestellt werden soll, war die Wahl der openCode-Plattform geboten.

Warum haben wir uns für die AGPL entschieden?

Die Affero General Public License (AGPL), Version 3.0 ist eine spezielle Form der GPL (General Public License), die sich besonders für Software eignet, die in der Cloud ausgeführt wird. Einer der wichtigsten Unterschiede zwischen AGPL und GPL ist, dass AGPL-lizensierte Software auch dann veröffentlicht werden muss, wenn sie als Software as a Service (SaaS) über ein Netzwerk bereitgestellt wird. Dies ist für die Software xPlanBox passend, da es sich um eine Server-Anwendung handelt, die perspektivisch Teil der deutschen Verwaltungscloud (DVC) (https://deutsche-verwaltungscloud.de) wird.

Wenn eine Software in einer Cloud-Umgebung ausgeführt wird, ist es nicht erforderlich, dass Benutzer die Software auf ihrem Arbeitsplatzrechner installieren, noch die Quellcode-Dateien für die Erstellung

einer ablauffähigen Ausführungseinheit, dem sogenannten Binary, herunterladen müssen. Mit der AGPL wird sichergestellt, dass der Quellcode auch dann unter gleichen Bedingungen bereitgestellt werden muss, wenn ein Cloud-Service diesen nutzt. Dies ermöglicht es Benutzern, die Software zu überprüfen, zu ändern und anzupassen.

Die AGPL bietet somit auch einen Schutz für die Community und die Benutzer von Open-Source-Software. Es stellt sicher, dass die Software nicht einfach genommen und als geschlossene, proprietäre Software verwendet werden kann, ohne dass die Änderungen an den Quellcode zurückgegeben werden, auch dann, wenn die Software nicht zum Download angeboten wird. So bleibt die Software für die weitere Verwendung und Weiterentwicklung durch die Community zugänglich und entspricht somit auch dem EfA-Prinzip.

Was ist XPlanung?

Im Oktober 2017 hat der IT-Planungsrat die verbindliche Einführung der Standards XPlanung und XBau beschlossen. Daraus folgt, dass IT-Verfahren die neuen Standards zur Raum- und Bauleitplanung unterstützen sollen.

Aus dem Standard XPlanung ergeben sich verschiedene Vorteile. Einerseits stellt der Standard eine einheitliche und transparente Raumordnungsplanung dar, die es ermöglicht, Pläne und Entscheidungen auf den verschiedenen Planungsebenen abzustimmen und zu koordinieren. Des Weiteren werden XPlanungsdaten für die Öffentlichkeit zugänglicher und können in weiteren Anwendungskontexten (z.B. zur Erstellung von Risikokarten) genutzt werden.

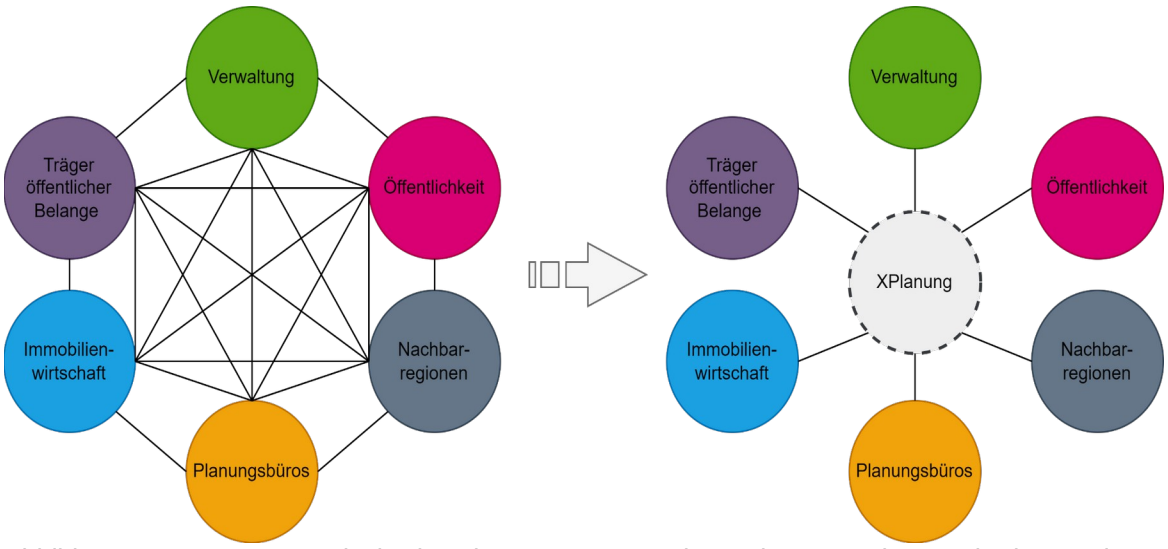


Abbildung 2: Datenaustausch der im Planungswesen tätigen Akteuren, ohne und mit gemeinsamem XPlanung-Standard, nach XLeitstelle (Abbildung nach [5])

Der Standard XPlanung und das dazugehörige Datenaustauschformat XPlanGML sind bereits seit 2003 in Entwicklung und stammen aus den E-Government-Initiativen Deutschland-Online und Media@Komm-Transfer. Im Oktober 2014 wurde dann auf Grundlage eines Antrags der Freien und Hansestadt Hamburg sowie eines Beschlusses des IT-Planungsrats der Bedarf nach „Austauschstandards im Bau- und Planungsbereich“ in die Standardisierungsagenda aufgenommen. Folglich kam es im Oktober 2017 zu der oben genannten verbindlichen Einführung durch den IT-Planungsrat. Für die kontinuierliche Pflege und Weiterentwicklung der Standards XPlanung wurde im Januar 2018 die XLeitstelle in Betrieb genommen. Als zentrale Geschäfts- und Koordinierungsstelle stellt die XLeitstelle ebenfalls unterstützende Werkzeuge zur Konformitätsprüfung des Dateiformats XPlanGML sowie eine Informationsplattform bereit.

Die XLeitstelle koordiniert heute die Weiterentwicklung des Standards XPlanung und von XPlanGML. Das XPlanGML ist das zugrundeliegende Dateiformat für die Speicherung von XPlanungsdaten. Das UML-Datenmodell von XPlanGML definiert ein objektorientiertes, semantisches Datenaustauschformat für Bauleitpläne, Regional- und Landschaftspläne.

Das Datenmodell entspricht einer Formalisierung gesetzlicher Grundlagen:

- BauGB, BauNVO, PlanzV für Bebauungs- und Flächennutzungspläne
- ROG (+ Ländergesetze) für Regionalpläne
- BNatSchG (+ Ländergesetze) für Landschaftspläne

Neben der XLeitstelle spielt die AG-Modellierung bei der Weiterentwicklung im Zusammenhang mit dem Standard XPlanung eine wichtige Rolle. Die AG-Modellierung ist ein Expertengremium und setzt sich aus Fachleuten der Planungsämter, Softwarehersteller sowie weiteren Beteiligten der Planungspraxis zusammen. Ziel der AG-Modellierung sind die Bearbeitung von Änderungsanträgen sowie die Erweiterung und Verabschiedung des Standards XPlanung [6,7].

Zur Unterstützung der Umsetzung des XPlanung-Standards hat die XLeitstelle einen Leitfaden (2. Auflage: https://xleitstelle.de/sites/default/files/2023-07/Leitfaden_Xplanung_2_Auflage.pdf) sowie eine Handreichung (3. Auflage: https://xleitstelle.de/sites/default/files/2023-01/Handreichung_3_Auflage_2023-01-04.pdf) zur Erstellung von XPlanGML-Dateien veröffentlicht.

Was ist die xPlanBox?

Die xPlanBox dient der Verarbeitung und Verwaltung von Daten im Rahmen des Standard XPlanung. Sie ermöglicht es Behörden und anderen Nutzern, XPlanungsdaten einfach und effizient zu verarbeiten und zu verwalten. Die xPlanBox kann in der Praxis eingesetzt werden, um die Verarbeitung und Verwaltung von Daten im Rahmen der nationalen Raumordnungsplanung zu vereinfachen.

Technisch basiert die xPlanBox auf verschiedenen OSS-Projekten. Für die Geodaten-Dienste wurde das OSGeo-Projekt deegree verwendet, als Datenbank kommt PostgreSQL mit der PostGIS-Erweiterung zum Einsatz, und für die Verarbeitung von Rasterdaten wird der MapServer verwendet. Die Transformation der XPlanungsdaten in das INSPIRE Planned Land Use (PLU) Schema, wird durch das Werkzeug HALE ermöglicht.

Was sind die Einsatzmöglichkeiten der xPlanBox?

Mit der xPlanBox können Planer und Entscheidungsträger verschiedene Funktionen ausführen, um die Planung und Verwaltung von Raumordnungs-Prozessen zu unterstützen. Einige der wichtigsten Funktionen, die mit der xPlanBox ermöglicht werden, sind:

- **Verwaltung** von XPlanungsdaten: Die xPlanBox ermöglicht es, XPlanungsdaten zu verwalten.
- **Prüfung** von XPlanGML-Dateien: Mit der xPlanBox können Planer XPlanGML-Dateien validieren.
- **Verwaltung** von Änderungen: Die xPlanBox ermöglicht es, Änderungen an XPlanungsdaten zu verwalten, um sicherzustellen, dass die Daten immer aktuell und korrekt sind.
- **Automatisierte Prozesse**: Die xPlanBox ermöglicht es, automatisierte Prozesse durchzuführen, wie z.B. die Überprüfung der Plausibilität von Planungsdaten, die Erstellung von Statistiken und die Erstellung von Prüfberichten.

- **Interoperabilität**: Die xPlanBox ermöglicht die Interoperabilität von verschiedenen Systemen und Plattformen, indem es die Möglichkeit bietet, Daten über das einheitliche Format XPlanGML und standardisierte Schnittstellen des OGC, wie WFS und WMS, auszutauschen.
- **Veröffentlichung von räumlichen Planungen**: Mit der xPlanBox lassen sich räumliche Planungen wie Bebauungspläne, Flächennutzungspläne und Regionalpläne veröffentlichen und so für die Öffentlichkeit zugänglich machen.
- **Importieren und Exportieren** von räumlichen Planungen: Mit der xPlanBox können räumliche Planungen aus anderen Systemen importiert und in andere Systeme exportiert werden.
- **Visualisierung** von räumlichen Planungen: Die xPlanBox ermöglicht die Visualisierung von räumlichen Planungen über die verfügbaren WMS-Dienste.
- **Zusammenarbeit und Kommunikation**: Mit der xPlanBox können die Beteiligten an einer Planung, wie z.B. Behörden, Planer, Bürger, miteinander kommunizieren und zusammenarbeiten.
- **Compliance**: xPlanBox ermöglicht die Einhaltung von gesetzlichen Vorgaben und Standards, wie z.B. der INSPIRE-Richtlinie, da neben den Fachdaten diensten mit den XPlanungsdaten auch Dienste für den Abruf der Daten im INSPIRE PLU-Schema bereitgestellt werden.

Wer sind die Anwender der xPlanBox?

Die xPlanBox wird hauptsächlich von Behörden und öffentlichen Einrichtungen in Deutschland genutzt, die für die Erstellung, Verwaltung und Veröffentlichung von räumlichen Planungen verantwortlich sind. Dazu gehören beispielsweise:

Kommunale Behörden: Städte und Gemeinden, die für die Erstellung und Verwaltung von Bebauungsplänen und Flächennutzungsplänen verantwortlich sind.

Landesbehörden: Landesbehörden, die für die Erstellung und Verwaltung von Bauleitplänen und anderen räumlichen Planungen zuständig sind.

Bundeseinrichtungen: Bundesbehörden, die für die Erstellung und Verwaltung von räumlichen Planungen auf nationaler Ebene verantwortlich sind.

Planungsbüros: Planungsbüros und Beratungsunternehmen, die für die Erstellung von räumlichen Planungen für Kommunen und andere öffentliche Einrichtungen beauftragt werden.

Bürger: Die xPlanBox ermöglicht es auch Bürgern, sich an der Planung ihrer Stadt und Gemeinde zu beteiligen und die Planungen einzusehen.

XPlanung und die xPlanBox können auch von anderen Akteuren im Bereich der Raumordnungsplanung genutzt werden, wie zum Beispiel von Wissenschaftlern, Umweltverbänden und von Interessengruppen, die an der Entwicklung von Plänen und Projekten beteiligt sind.

Da die Software Open-Source ist, ermöglicht dies auch kleinen und mittelständischen Unternehmen, die an der Raumordnungsplanung beteiligt sind, die Software zu nutzen, ohne hohe Kosten für kommerzielle und proprietäre Softwarelösungen aufwenden zu müssen.

Welche Komponenten beinhaltet die xPlanBox?

XPlanValidator

XPlanGML-Dateien können auf Korrektheit geprüft werden, indem man sie gegen eine XML-Schema-Datei (XSD) validiert. Eine XSD ist eine Art "Vorlage" für die Struktur und Inhalte einer XML-Datei. Sie legt fest, welche Elemente und Attribute in der XML-Datei enthalten sein müssen und welche Regeln

für die Verwendung dieser Elemente und Attribute gelten. Für XPlanGML-Dateien gibt es eine offizielle XSD, die von der XLeitstelle bereitgestellt wird. Diese XSD kann verwendet werden, um XPlanGML-Dateien auf Korrektheit zu prüfen.

Es gibt auch verschiedene Tools und Anwendungen, die die Validierung von XML-Dateien gegen eine XSD automatisch durchführen können, z.B. XmlSpy, Oxygen XML Editor oder xmllint. Auch die meisten Entwicklungsumgebungen (IDEs) bieten integrierte Möglichkeiten zur Validierung von XML-Dateien gegen eine XSD.

Es ist wichtig zu beachten, dass die Validierung gegen die XSD lediglich die syntaktische Korrektheit der XPlanGML-Datei prüft und nicht die semantische oder inhaltliche Korrektheit der Daten in der Datei. Um die inhaltliche Korrektheit sicherzustellen, müssen die XPlanGML-Dateien auch semantisch und geometrisch und bei Bedarf auch manuell geprüft werden.

Der XPlanValidator vereint die drei technischen Prüfungen auf syntaktische, semantische und geometrische Korrektheit. Die Komponente kann über eine Web-Oberfläche oder ein REST-API aufgerufen werden. Bei der XLeitstelle steht eine öffentliche Instanz der Anwendung (https://xleitstelle.de/validator) zur Prüfung von Daten bereit. Auf Länderebene wurden zusätzlich zu den Standardregeln Erweiterungen als sogenanntes Profil entwickelt, die in den XPlanValidator eingebunden werden können. So liegen bereits für die Länder Berlin, Brandenburg und Bayern Erweiterungen für den XPlanValidator auf der openCode-Plattform bereit.

XPlanManager

Der XPlanManager dient der Verwaltung von Planwerken innerhalb der xPlanBox. Mit dem XPlanManager können XPlanGML-Dateien und XPlanArchive importiert, editiert oder gelöscht werden. Die Komponente kann über eine Web-Oberfläche oder ein REST-API aufgerufen werden.

XPlanDienste

Die XPlanDienste stellen folgende Endpunkte zur Verfügung:

XPlanWMS, XPlanWerkWMS und XPlanArtWMS

XPlanMapServer

XPlanMapProxy

Der XPlanWMS, XPlanArtWMS und XPlanWerkWMS sind auf dem Standard Web Map Service (Version 1.1.1 und 1.3.0) des Open Geospatial Consortium (OGC) basierende Kartendienste. Während der XPlanWMS die Inhalte planübergreifend darstellt, visualisiert der XPlanWerkWMS genau ein einzelnes Planwerk. Der XPlanArtWMS wiederum stellt je einen Endpunkt für jede Planart wie z.B. BPlan oder FPlan zur Verfügung. Über den XPlanMapServer können im XPlanGML referenzierten Rasterdaten abgerufen werden. Der XPlanMapProxy kann als Proxy kaskadierend für den XPlanWMS und XPlanMapServer mit Cache-Funktion verwendet werden, um den Zugriff auf Raster- und Vektorebenen des XPlanWMS und XPlanMapServer zu beschleunigen. Der Zugriff auf XPlanMapProxy erfolgt über die Standardschnittstellen WMS und WMTS.

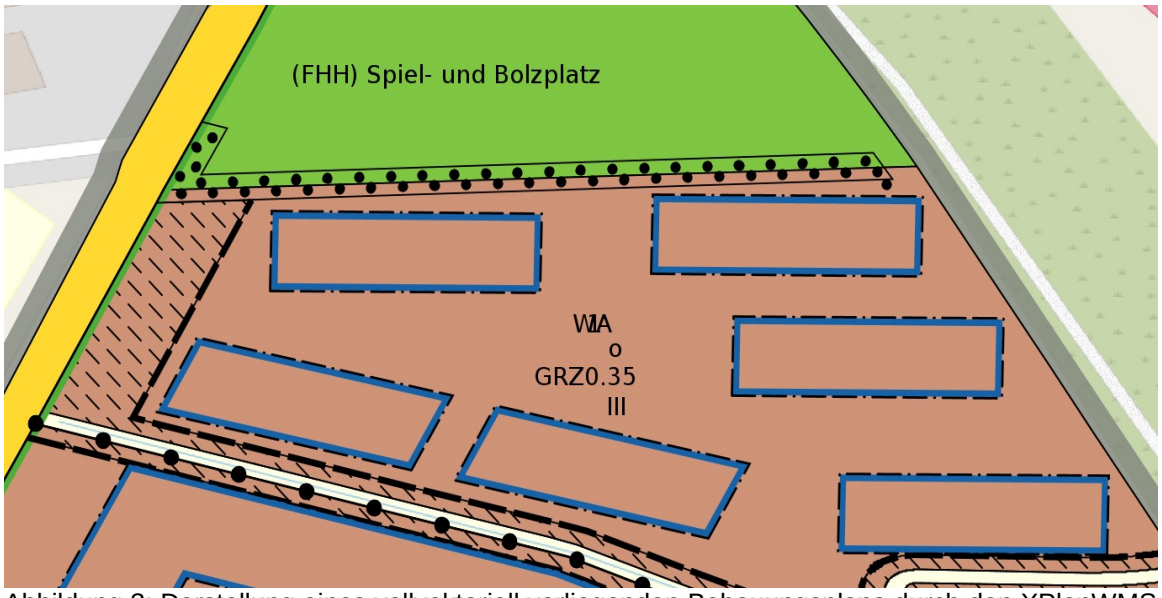


Abbildung 3: Darstellung eines vollvektoriell vorliegenden Bebauungsplans durch den XPlanWMS (Maßstab 1:280)

XPlanWFS

Der XPlanWFS ist ein auf dem Standard Web Feature Service (Version 1.1.0 und 2.0.0) des Open Geospatial Consortium (OGC) basierender Dienst zur Abfrage von Vektordaten und stellt Endpunkte für jede XPlanGML-Version bereit. Zu jeder unterstützten XPlanGML-Version stellt die xPlanBox einen XPlanWFS-Endpoint bereit. Jeder Dienst nutzt das Datenbankschema in der XPlanDB, das zur entsprechenden XPlanGML-Version gehört.

XPlanSynWFS

Der XPlanSynWFS ist ein auf der Open Source Software deegree basierender Dienst zur Abfrage von Vektordaten und ist konform zum Standard Web Feature Service (Version 1.1.0 und 2.0.0) des Open Geospatial Consortium (OGC). Der Dienst dient der Abbildung des synthetisierten XPlanGML-Anwendungsschemas (XPlanSynGML). Dieses stellt eine vereinfachte und zusammenfassende Form der von der xPlanBox unterstützten XPlanGML-Versionen dar.

XPlanDB

Die zentrale Datenhaltung der xPlanBox basiert auf einer PostgreSQL-Datenbank. Die XPlanDB enthält für jede unterstützte XPlanGML-Version ein eigenes Datenbankschema. Die Datenbankschema je XPlanGML-Version unterscheiden sich nur im Namen des Schemas und nutzen das BLOB-Mapping von deegree für die Ablage von XPlanGML-Dokumenten.

Welche Werkzeuge und Tools kommen bei der Erstellung von XPlanGML noch zum Einsatz?

Wenn man XPlanung und XPlanGML erstellen und verwalten möchte, kommen neben der xPlanBox auch weitere Software-Tools zum Einsatz. Einige Beispiele sind:

- **GIS-Software** (Geographic Information System): Diese Art von Software ermöglicht es, die geografischen Daten wie Karten, Luftbilder und andere geografische Informationen zu einem Plan zu erfassen, zu verarbeiten, zu analysieren und darzustellen. So kann mit dem QGIS-

Plugin XPlan-Umring (https://plugins.qgis.org/plugins/xplan-umring/) ein Plan im sogenannten Raster-Umring-Szenario erstellt werden. Mit dem QGIS-Plugin XPlan-Reader (https://plugins.qgis.org/plugins/xplan-reader/) kann eine XPlanGML-Datei in QGIS geladen und dargestellt werden.

- **CAD-Software** (Computer-Aided Design): Diese Art von Software ermöglicht es, technische Zeichnungen und Pläne zu erstellen und zu bearbeiten. In diesem Bereich gibt es aktuell keine Open-Source-Lösung, die für die Erstellung von XPlanungsdaten geeignet ist.
- **Datenbank-Management-Software**: Diese Art von Software ermöglicht es, die Daten zu speichern, zu verwalten und abzufragen. Für die xPlanBox kommt die PostgreSQL Datenbank mit der Erweiterung PostGIS zum Einsatz.
- **ETL-Software**: Diese Art von Software ermöglicht es, Daten von einer Datenquelle z.B einer CAD-Datei mittels eines Extract, Transform, Load (ETL) Prozesses in das Zieldaten-Format XPlanGML zu überführen. Im Bereich XPlanung kommt das Open-Source Werkzeug HALE (https://github.com/halestudio/hale) zum Einsatz.

Es ist wichtig zu beachten, dass diese Tools nicht unbedingt verwendet werden müssen, und es gibt auch andere Open-Source Tools und Lösungen, die für die Erstellung und Verwaltung von XPlanungsdaten und XPlanGML verwendet werden können.

xPlanBox Version 8.0 – jetzt Cloud-ready

Die neueste Major Release-Version der xPlanBox, 8.0, eröffnet neue Möglichkeiten für den Betrieb der Software. Die xPlanBox ist nun Cloud-ready und kann auf Infrastrukturen wie Kubernetes oder Minikube installiert werden. Das Deployment via Docker Container wird ebenfalls weiterhin unterstützt. Alle der bereits im vorherigen Abschnitt aufgelisteten Komponenten der xPlanBox wurden auf Java 17 und Spring Boot-Anwendungen umgestellt, wodurch Sicherheit und Performance verbessert wurden.

Die XPlanManagerAPI, XPlanValidatorAPI und XPlanDokumenteAPI unterstützen nun eine zentrale Authentifizierung und Autorisierung über Keycloak. Weiterführend wurden die Schnittstellen so optimiert, dass sie eine effizientere und skalierbare Verarbeitung ermöglichen. Die interne Kommunikation zwischen den Komponenten erfolgt nun über den Message Broker RabbitMQ, wodurch eine zuverlässige und nachverfolgbare Nachrichtenübertragung sichergestellt wird. Durch die Unterstützung asynchroner API-Aufrufe können Anfragen schneller und ressourcenschonender bearbeitet werden.

Für die Ablage der Daten ist ab der Version 8.0 ein S3-Objektspeicher erforderlich, in dem neben dem Rasterdaten auch die anderen Begleitdokumente abgelegt werden. Der Zugriff auf die Rasterdaten aus dem S3-Objektspeicher ist dann über den XPlanMapServer als auch den XPlanMapProxy möglich.

Der XPlanValidator unterstützt in der Version 8.0 nun GeoJSON als Ausgabeformat für Geometrien mit Validierungsfehlern. Auch wurde der Validierungsbericht im JSON-Format nochmal grundlegend überarbeitet und unterstützt jetzt auch die Ausgabe von Validierungsergebnissen von mehreren Planwerken innerhalb einer XPlanGML-Datei.

Ausblick

Mit den Möglichkeiten die Software in der DVC zu betreiben, fallen auch für kleinere Kommunen die technischen Hürden für den Betrieb. Darüberhinaus unterstützt die xPlanBox Nutzer beispielsweise auch dabei die Planungsdaten in die Digitalen Zwillinge der Kommunen einzubinden und so Planungsalternativen sowie verschiedene Planungsstände in den Beteiligungsportalen anzeigen zu können.

Kontakt zu den Autoren:

Torsten Friebe, Julian Zilz
lat/lon GmbH
Im Ellig 1
53343 Wachtberg
Tel.: ++49(0)228 24 333 784
info@lat-lon.de

Literatur und Quellen

[1] Vom Softwareverzeichnis zur Plattform für die Digitale Souveränität: Leonhard Kugler im Gespräch über openCode: https://www.digitale-verwaltung.de/SharedDocs/interviews/Webs/DV/DE/2025/leonhard-kugler-interview.html

[2] openCode-Relaunch: Vom Softwareverzeichnis zur Plattform für die Digitale Souveränität: https://www.zendis.de/ressourcen/opencode-relaunch-vom-softwareverzeichnis-zur-plattform-fuer-die-digitale-souveraenitaet

[3] Die Plattform der Öffentlichen Verwaltung für den Austausch von Open-Source-Software: https://opencode.de/de/was-ist-opencode

[4] Richtlinie zum Softwareverzeichnis: https://opencode.de/de/wissen/softwareverzeichnis-einsatz-in-der-verwaltung/richtlinie-zum-softwareverzeichnis

[5] Über XPlanung: https://xleitstelle.de/xplanung/ueber_xplanung

[6] Krause, K.-U., Duan, X., Horenczuk, J. u. H. Leuner (2020): XPlanung und XBau. In: Informationen zur Raumentwicklung 47, Nr. 3. S. 6–21.

[7] Geoinformationen in der Raumplanung: https://www.bbsr.bund.de/BBSR/DE/veroeffentlichungen/izr/2020/3/izr-3-2020.html

Online-Karten für die Verkehrswende mit OpenData und FOSS

Johanna Klitzschmüller

Mit OpenData und freier/open-source Software lässt sich eine Online-Kartenanwendung bauen, mit der sich die Straßenraumaufteilung der Stadt Kiel darstellen und untersuchen lässt. Damit kann diese Online-Karte ein weiteres Werkzeug im Kampf für eine Verkehrswende sein.

Aus der kritischen Kartographie wissen wir, dass Karten Machtinstrumente sind. Sie lassen sich im Rahmen von bspw. Gegenkarten / Counter-Maps damit auch nutzen, Gegenmacht aufzubauen bzw. alternative Bilder unserer Umwelt zu schaffen. Anhand eines Beispiels aus Kiel zeigt der Vortrag, wie sich mit OpenData und freier/open-source Software eine Online-(3D-)Karte erstellen lässt, mit dem sich die Aufteilung des Straßenraums im Stadtgebiet darstellen lässt. Mit einer solchen Kartendarstellung können Kämpfe für eine Verkehrswende unterstützt werden.

Um die Kartenanwendung zu erstellen, benötigt es die genannten Daten zur Straßenraumaufteilung. Nach einer Anfrage via FragDenStaat stehen diese für die Stadt Kiel als OpenData zur Verfügung. Als Software wird eine Kartenanwendung für den Browser wie MapLibre benötigt und um die Daten zur Verfügung zu stellen, wird ein WMS-Server (bspw. QGIS-Server) und/oder TileServer benötigt.

Mit EnMAP-Box und OSM durch Namibia: ein hyperspektraler Praxisbericht

Benjamin Jakimow

Verbuschung von Savannen bedroht deren Artenvielfalt und Nutzung als Weideland. Unser Vortrag zeigt, wie wir Verbuschung in Namibia bei Feldarbeiten mithilfe von QGIS, der EnMAP-Box und anderer freier und open-source Software dokumentiert haben, um sie anschließend mithilfe von EnMAP- und anderen Satellitendaten zu beschreiben.

In vielen Savannengebieten weltweit findet eine zunehmende Verbuschung auf Kosten natürlicher Gräser statt. Sie bedroht einheimische Tiere und Pflanzen, und reduziert die Nutzbarkeit und den Wert kostbaren Weidelandes. Um diese Verbuschungsprozesse besser zu verstehen und sie kartieren zu können, wurden im Oktober 2024 Feldarbeiten entlang der Savannengebiete Namibias in einem Korridor von Windhoek bis zur angolanischen Grenze durchgeführt. Dabei wurden Vegetations- und Bodencharakteristika erfasst, sowie Spektralmessungen und Drohnenflüge vorgenommen mit dem Ziel, relevante Vegetationstypen besser in Zeitserien von multi- und hyperspektralen Satellitendaten erkennen zu können.

Bei unseren Arbeiten wurde verschiedene FOSS Software, insbesondere QGIS und die EnMAP-Box verwendet. Unser Vortrag zeigt, wie wir diese zur Planung und Durchführung der Feldarbeiten eingesetzt haben, sowie bei der anschließenden Datenauswertung nutzen konnten. Wir stellen den aktuellen Stand unserer Forschung vor und geben einen Ausblick, was dies für die weitere Entwicklung der EnMAP-Box an der HU-Berlin bedeutet.

Digitaler Zwilling Niedersachsen auf Basis der Unreal Engine und LGLN Open Data

Julian Müller, Vincent-Aleister Raveling



Dieses Projekt aus dem Bereich "digitaler Zwilling" erkundet die Verbindbarkeit von offenen Geodaten des LGLNs mit der Unreal Engine 5 um moderne Echtzeitanalysen zu entwickeln. Als Szenegrundlage dient das Geländemodell mit digitalen Orthophotos zur Texturierung und LOD 1 und 2 Gebäudemodellen. Auch 3D-Meshes und Punktwolken können geladen werden. Alle Daten werden aus der Cloud bezogen, was die Dateigröße der Anwendung an sich gering hält.

Gezeigt wird prozedurale Generierung von Wohneinheiten in urbanen Gebieten basierend auf dem B-Plan am Beispiel Osnabrück, einziehen von Schallschutzmauern und einer Echtzeitanalyse und rudimentäre GIS-Funktionen, wie z.B. Strecken- und Flächenmessungen innerhalb der Szene. Es ist vorgesehen diese GIS-Funktionen als open-source Plugins zu entwickeln, die mindestens über den Epic Games Marketplace (jetzt "FAB") bezogen werden können. Ziel dieser Entwicklungen ist die architektonische Grundlage für einen Geobasiszwilling zu schaffen, der mit Fachdaten auch von anderen Ämtern angereichert werden kann und soll Analysemöglichkeiten zur Einschätzung von geplanten Vorhaben bieten. Ultimativ sollen die Ergebnisse der Analysen in Vorschlägen für Handlungsmaßnahmen münden.

Artificial Ground Truth Data Generation for Map Matching with Open Source Software

Adrian Wöltche

Zenodo-DOI: https://zenodo.org/records/14774143

Abstract: Map matching is a widely used technology for assigning tracks recorded by Global Navigation Satellite Systems (GNSS) to existing road networks. Due to the measurement uncertainty of GNSS positions, the biggest challenge is to map them accurately. To develop and verify suitable algorithms for map matching, ground truth, i.e., the traveled routes in the road network, is required as a reference. However, GNSS recorded tracks naturally lack the ground truth routes. Providing this data is time-consuming and costly in these cases, as it requires manual correction of the routes based on human memorization. This is not practical on a large scale, e.g., with floating car data (FCD). This is why there exist only a few isolated ground truth data sets that were created in this way for map matching. To close this gap, we introduce and evaluate in this work a new open source tool-chain for artificially generating large amounts of simulated ground truth routes for map matching. Based on these routes, we generate simulated FCD and we apply comparably authentic and parameterizable artificial GNSS noise with varying noise characteristics. The generated data allows to thoroughly evaluate and improve the performance of existing map matching algorithms and facilitates in future research the development of novel algorithms based on sufficiently large and diverse labeled data. In this work, we evaluate different scenarios of varying noise characteristics of our artificially generated ground truth data to compare the robustness, individual strengths, and weaknesses of existing open source map matching implementations. Our new approach of artificially generating ground truth data for map matching addresses the existing lack of sufficient available reference data for ongoing map matching research.

Keywords: artificial ground truth generation, traffic simulation, map matching, open source

DOI: 10.5281/zenodo.14774143

1 Introduction

The input to a map matching algorithm is a sequence of points, i.e., a track, acquired during the movement of a tracking device, e.g., a GNSS tracker, and an existing road network. Map matching algorithms then assign the points of the track to the edges of the existing road network graph and deduce a path through it containing all the assigned edge positions [4]. This results in a route, also called a “match”, that aligns exactly with the road network. In contrast, the input track does not lie on the road network due to GNSS measurement uncertainty.

Map matching is a method to resolve the discrepancies between recorded GNSS tracks and a given road network in a most plausible way, according to defined evaluation functions. These algorithms operate under the assumption that the given road network is complete and accurate [35, 36] and that only the given recorded tracks have significant discrepancies to the roads. Although road networks are not completely accurate by design, e.g., curves are usually modeled by short consecutive segments instead of splines, this is not an issue in map matching, as the goal is to find the most representative route within the road network rather than the true real-world positions of a track.

A common use case for map matching is to match large scale FCD tracks, see Figure 1 for an example with artificial data. The routes that are shown in the left panel, which align exactly with the road network, do generally not exist in practice for real-world FCD. Only the noisy tracks, i.e., the FCD tracks recorded by GNSS trackers, that are shown in the middle panel exist in practice. For an exam-

ple with real-world FCD that aligns with this Figure 1, also see the respective Figure 1 in [32]. Map matching the FCD tracks from the middle panel results in the routes that are shown in the right panel, which again align exactly with the road network and are close to the original routes shown in the left panel, which would ideally be the desired result. Thus, map matching results in an aggregation of all given FCD tracks onto existing road network segments.

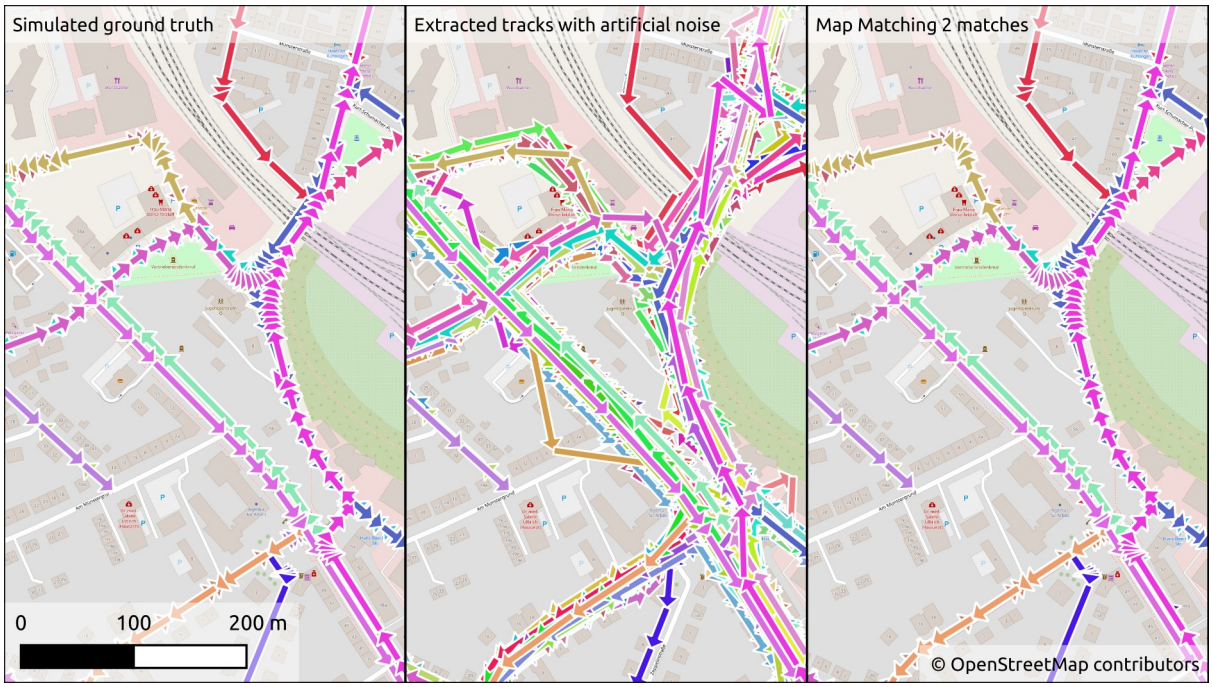


Figure 4: Map Matching of artificial ground truth data. In the left panel is the simulated ground truth, in the middle panel are the artificially noised tracks from our approach, and in the right panel are the map matching results from our own open source software “Map Matching 2”.

Such aggregated matches enable the analysis of various traffic and movement pattern topics, such as road usage statistics or traffic flow analysis. These can provide valuable input for city and traffic planning use cases, such as road maintenance intervals, planning bypass constructions, or planning public transportation networks. In addition, not only large-scale FCD, but also individual, personal GNSS tracks, such as GPX records, can be individually matched as required for information purposes.

One of the main challenges in map matching is to find the routes that most accurately represent the given input tracks in the given road network, i.e., find the routes that would actually be driven in the road network. There are two major issues that complicate this process. First, due to data quality challenges [4], the discrepancy between the tracks and the road network can vary drastically. Second, there is usually no ground truth route in a given road network for a given input track available to train and evaluate algorithms against.

Ground truth in the field of map matching “is a sequence of connected road segments of the traveled route on the map” [35] and “represents the correct path taken by the vehicle through the road network” [24]. Thus, ground truth in the context of map matching is not the actual “true geographical position” from which a GNSS point was measured but rather the “correct match” within the defined network.

The current lack of comprehensive ground truth data severely limits the ability to evaluate map matching algorithms, models, and metrics. Today, only a few ground truth data sets ex-

ist, as creating such data requires manually mapping GNSS records to predefined road networks, which is expensive. Other “reference” map matching data sets that were created by manually correcting map matching results are by definition not “ground truth” as the correct route in the road network remains unknown to the authors. These data sets are arguably subjective and sometimes also contain small oversight errors. Furthermore, evaluating map matching algorithms only on existing data sets bears the risk of introducing a general optimization bias. As such, the general availability of diverse and scalable ground truth data sets for map matching is an important aspect for ongoing map matching research.

Therefore, in this work we present an open source process for generating artificial ground truth map matching data sets by simulating routes on a given road network. These routes are then virtually driven by vehicle agents in a traffic micro-simulation software, with periodic emission of the exact virtual positions into sequence of points. On these tracks we then apply an artificial, parameterizable noise generation algorithm to simulate GNSS measurement uncertainty. The result of this process are artificial ground truth routes in a given road network and corresponding artificial GNSS tracks with synthetic noise applied.

This artificial generation approach differs from real-world cases where GNSS recordings are manually mapped to a road network after collection. In such cases, additional discrepancies arise because the recorded tracks and the defined road network usually come from slightly different true original geopositions. In contrast, in our artificial process, the ground truth route is created before the track positions are emitted, and both originate from the same geopositions. Thus, there exist no discrepancies between the ground truth route and the track positions before artificial noise is added to the track, which eliminates a source of error for the training and evaluation of map matching algorithms that real-world data usually contains.

Simulating artificial GNSS noise comes with some challenges. Real-world GNSS noise has various sources of disturbance, such as weather phenomena or satellite opacity, which cannot be accurately represented in a virtual process to date. However, current map matching models are not designed to accurately resolve real-world GNSS disturbances, as they rely on simplified statistical methods that evaluate, for example, distances or angles between points. Thus, we hypothesize that generating artificial noise with similar statistical distributions as real-world GNSS noise is sufficient for producing meaningful ground truth data for map matching research.

As long as there is no technology for measuring real-world geopositions with absolute precision under any circumstances, or a technology that is able to completely resolve the measurement uncertainty of GNSS measurements, there will always be some discrepancies between recorded GNSS geopositions and already defined road networks, which generally also rely on recorded GNSS geopositions. This ongoing discrepancy is addressed with map matching and as such, robust and accurate map matching solutions remain needed.

1.1 Literature and State of the Art

Our work is based on various existing research in the fields of GNSS (i.e., GPS error distribution), traffic simulation (i.e., the open source tool SUMO), and map matching (i.e., data sets, models and algorithms, benchmarking, and our own open source implementation).

Previous approaches to provide map matching ground truth routes hand-matched [16, 24] self-recorded GNSS tracks. Currently, there exist only two tracks with one ground truth route each, one from [16] and one from [24]. There has also been work on hand-correcting [15, 23, 32] the results of an initial map matching of tracks with the ground truth routes being unknown to the authors. However, this only lead to an additional 100 [15], 10 [23], and 264 [32]

tracks and some of the “reference” routes, i.e., not “ground truth” routes, contain small errors due to human oversight.

In our previous work [32] we described, evaluated, and compared the three real-world GNSS map matching data sets from [15, 16, 24] and our own fourth real-world data set with existing open source map matching tools against our own map matching implementation “Map Matching 2” (https://github.com/iisys-hof/map-matching-2) [32] in a detailed benchmark. Among the other open source map matching solutions were, e.g., Barefoot (https://github.com/bmw-carit/barefoot) [19], Fast map matching (https://github.com/cyang-kth/fmm) (FMM) [33], Map Matching based on GraphHopper (https://github.com/graphhopper/graphhopper#mapmatching), Open Source Routing Machine (https://github.com/Project-OSRM/osrm-backend) (OSRM) [20], and Valhalla (https://github.com/valhalla/valhalla). Subsection 1.3 and Table 1 in our previous work [32] provide more details on the technologies and models of these tools, which we also benchmark in this work.

All third-party tools rely on the Hidden Markov Model (HMM) as a stochastic method for evaluating alternate paths between possible matching solutions for finding an optimal solution concerning the defined models and metrics, except for FMM, which also implements ST-Match [18], a simpler stochastic model. Our own “Map Matching 2” tool primarily relies on a more robust Markov Decision Process (MDP) [32], but we also implemented the traditional HMM for direct comparison purposes. There are several review works about the history and current state of research about map matching [4, 6, 10, 11, 25]. As of today, stochastic methods are the most accurate models. However, the general accuracy still needs to improve, which, in our opinion, mainly originates from the challenging noise characteristics of GNSS noise in general.

The results from the benchmark in our previous work [32] on the real-world GNSS map matching data sets showed that our solution outperforms the other solutions in terms of accuracy and mostly also in terms of computational resources, especially when put in relation with the achieved accuracy. Due to some recent software updates of our own “Map Matching 2” solution, our open source code repository now contains examples with newer accuracy and computational resource benchmark results in the “data” folder for all mentioned real-world GNSS data sets. The achieved accuracy currently is 98.73 % [15], 99.80 % [16], 98.59 % [23] (with two large data errors manually corrected, see repository explanation), 99.89 % [24], and 99.58 % [32].

Concerning GNSS, there exist various implementations, e.g., Global Positioning System (GPS) (https://www.gps.gov/), Galileo (https://www.gsa.europa.eu/), GLONASS (https://www.glonass-iac.ru/en/), or Beidou (http://en.beidou.gov.cn/). However, in this work we focus on GPS, because currently existing map matching data sets, as well as OpenStreetMap (OSM), all use the GPS spatial reference system WGS-84. Furthermore, SUMO also emits GPS coordinates as virtual FCD, and real-world FCD also generally contains GPS coordinates. Therefore, GPS is currently the most common and compatible choice.

The most relevant topic of GPS for this work is the GPS error distribution. Due to atmospheric disturbances [13] and signal reflections [22], GPS is said to have an accuracy range of at least several meters. In [30] the error range with smartphones in an open space was measured to be within 4.9 m. In [12] the ≈ 5 m range of smartphone GPS accuracy in an open space is confirmed, and it is shown that the error also depends on the GPS site, i.e., location on the globe. Furthermore, it is shown that near buildings the error is ≈ 3 − 4 times higher, up to ≈ 15 m − 20 m, which is especially relevant for FCD, i.e., driving in cities while navigating with smartphones. When satellite opacity [9] happens, i.e., when

a GPS receiver loses contact with enough GPS satellites at a time for locating its position, for example, because of buildings, hills, bridges, or tunnels, large outliers can occur. These outliers are outside the regular noise distribution and can reach several hundreds of meters in practice. For example, the map matching data set from Newson and Krumm [24] contains such an outlier with > 100 m.

GPS error is spatially and temporally autocorellated, as was extensively researched by analyzing stationary GNSS records [8, 9, 26, 27, 29]. In [8] the authors show that the GNSS distance error mostly follows a Rayleigh distribution. A Rayleigh distribution is a probability distribution for non-negative Gaussian distributed values. The GNSS longitude and latitude measurements mostly follow a Gaussian distribution, which is also shown in [9], which makes the GNSS distance error as the "geometric sum of latitude and longitude errors" [9] similar to a Rayleigh distribution [9]. In [8] the authors also state that using an error distribution that does not account autocorrelation leads to a different kind of error, as "the data would fluctuate randomly without the autocorrelation phenomena" [8]. This would happen, e.g., with using a Gaussian distribution for simulating GNSS error.

Autocorrelation means that succeeding measurements with close timestamps generally have similar shifts on the axes within the accuracy means. GNSS measurements are time-dependent on previous values, but over time they become less dependent. The reasons for this are that within close timestamps, the satellite positions, atmospheric disturbances, and earth rotation are similar and they diverge the more time goes by. In [9] the authors show a GNSS error autocorrelation function, a GNSS error over time, and moving average smoothing of the error over time, to give an impression of the autocorrelation of GNSS measurements. The authors also show an example of how the distance error significantly raises when satellite opacity happens. Such outliers lie far outside the original error distribution and autocorrelation. In [26] the spatial and temporal autocorrelation is also confirmed. Moreover it is shown that GPS distance errors overestimate actual distances, i.e., a GPS track recorded at a static position is always longer than the static recording device has actually moved.

In [27] the authors simulate GPS error with an Autoregressive (AR) model and they validate their synthetic GPS data with real-world collected data to show that autocorrelated models are able to simulate the noise characteristics of real-world GPS data. In [29] the Ornstein-Uhlenbeck (OU) process with the Euler-Maruyama method is applied for simulating GPS noise. In the same work, also the simple Gaussian distribution, the AR model, the Moving Average (MA) model, and the Autoregressive Moving Average Model (ARMA) are evaluated. The authors explain that while the AR, MA, and ARMA models are autocorrelated, only the OU model is additionally mean-reverting, which means that the noise generated over time not only depends on its previous values but simultaneously returns to a defined value, for example, zero, with the attraction being greater the further away the current error is from the defined value. The authors in [29] also show in their Figure 6 that OU generated noise is similarly autocorrelated as real GPS noise, in their case most closely comparable up to two hours. Since the OU process is also based on a Gaussian process for the random value generation, the OU process gives a similar GNSS error distribution as real GPS error over time.

On the topic of traffic simulation, there are various implementations that have been proven to model traffic with sufficient accuracy [3, 21, 28]. Simulation of Urban MObility (SUMO) [14] is an open source micro traffic simulation tool that supports OSM road networks and FCD output for generating routes that are virtually driven while simultaneously emitting GPS coordinates as GPS tracks. Traffic can be generated in various ways, for example, randomly, with origin-destination (OD) matrices, based on road counting data, turn probabilities, or based on activities, e.g., based on population distribution with the SAGA tool-chain [5].

There are a few previous works that generate map matching ground truth data based on a traffic simulation. However, the complete processes are not openly available. In [7] traffic is simulated based on SUMO, and the onboard “traceExporter.py” tool is used for applying artificial noise based on a simple Gaussian distribution. As such, the artificial error is not autocorrelated. The authors use a sampling interval of 30 s with a Gaussian distribution standard deviation σ of 5 m. Using a low sampling interval helps in mitigating the effects of autocorrelation, as samples further apart are less autocorrelated, especially with moving vehicles in cities. In [17] traffic is also simulated based on SUMO, and the open source tool GNSS-INS-SIM (https://github.com/Aceinna/gnss-ins-sim) is used for generating artificial noise, again, based on a simple Gaussian distribution. They use sampling intervals of 1 s, 5 s, and 10 s with respective Gaussian distribution standard deviation σ of 5 m, 10 m, and 20 m. Since the used noise distributions are not autocorrelated in these works, only low noise can be used with high sampling intervals, or else the track shape degrades significantly when high uncorrelated normal noise is applied on a high sampling interval.

1.2 Approach

To address the ongoing lack of available map matching ground truth data, we present in this work an open process based on an open source software for generating diverse artificial map matching ground truth data sets on OSM road network extracts. We use SUMO for simulating the artificial ground truth routes and for extracting the FCD positions, initially without error. We then apply a new open source implementation for adding artificial autocorrelated GNSS noise based on the Ornstein-Uhlenbeck process to the FCD. This tool also implements the simple Gaussian distribution and the ARMA model for comparison purposes. We also implement parameterizable moving average smoothing on the generated noise to account for excessive jittering of the underlying Gaussian random values. Additionally, our tool allows for applying simple random outliers with parameterizable distribution closeness along the tracks for simulating potential satellite opacity. We also deploy additional settings and tools for analyzing the generated noise distributions and the track characteristics. We evaluate the generated synthetic GNSS noise distributions against real-world GPS recordings to show that our process is able to achieve authentic GNSS noise results within certain limitations. Finally, we compare multiple diverse artificial generated map matching ground truth data sets with different open source map matching tools and our own implementation to reveal different strengths and weaknesses of all tools depending on the input track data noise characteristics.

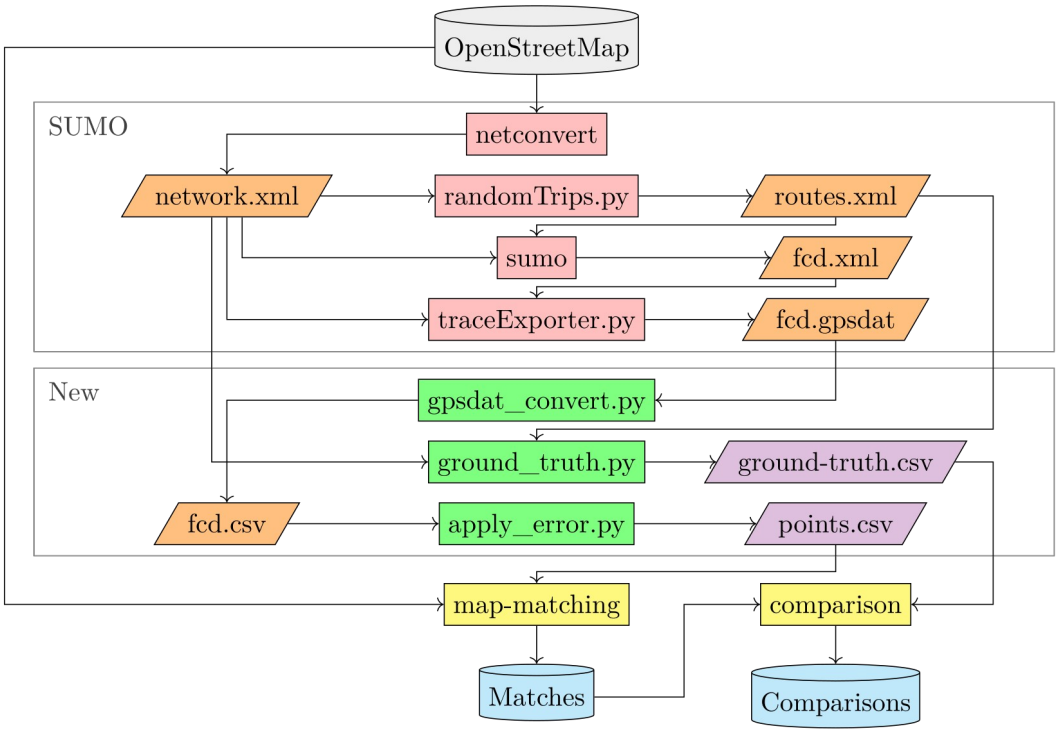


Figure 5: Technology roadmap for our artificial ground truth generation open source tool-chain. The rectangle shapes show processes and tools, the trapezium shapes show artifacts, the cylinders show the input and output data. The processes and tools in the “SUMO” frame are provided by SUMO, the tools in the “New” frame below are newly provided from us. The “ground-truth.csv” and “points.csv” files are the essential results of our pipeline that are further matched and evaluated. The “map-matching” and “comparison” processes are existing open source algorithms. The rest are intermediate files.

2 Method

In Figure 2 we present the step-by-step process of how we generate artificial map matching ground truth data with our new approach. First, the OSM road network is converted into the SUMO format with “netconvert”. We use some additional parameters for reducing the internal lane and junction area sizes so that the emitted FCD stays close to the original geopositions from OSM. Next, we generate movements with “randomTrips.py” from the SUMO package. Instead of random movements, more sophisticated trip generation algorithms may be used, but random movements can be generated for any region or city without additional knowledge. The resulting movements are then simulated with the “sumo” command, i.e., virtual vehicles drive the generated routes, and FCD is emitted every second. The raw FCD XML output then is converted into the intermediate GPSDAT format with the SUMO “traceExporter.py” utility.

We continue our workflow with converting the GPSDAT format into a FCD CSV file with Well- Known-Text (WKT) point representation, as this facilitates further processing in our pipeline. We also export the ground truth routes from the SUMO network and the generated routes XML file, i.e., the exact OSM edges are exported into WKT lines. Finally, we apply our artificial noise generation on the input FCD tracks, which results in the noisy points CSV file. This file is then matched with map matching, and the results are compared with the extractted ground truth routes. In this way, we are able to measure the accuracy of a map matching algorithm depending on the defined artificial noise characteristics on the FCD tracks.

2.2 Artificial GNSS noise

One of the key aspects of this work is the artificial noise generation process. With our “apply_error.py” tool, simple Gaussian, ARMA, or OU distributed noise can be generated. In addition, simple outliers can be added with a closeness parameter that describes how likely outliers lie directly next to each other, for simulating satellite opacity. Outliers are sampled from a Gaussian distribution, as they do not follow the autocorrelation of regular GNSS noise.

The Gaussian (normal) distribution is uncorrelated and, although frequently used in previous research, does not correctly represent GNSS error over time:

$$X \sim N(\mu, \sigma^2)$$

with μ := zero (mean); σ := error (standard deviation);

The autoregressive moving-average (ARMA) process is autocorrelated but contains no term for mean-reversion, which means that it offsets further from the zero-mean the longer it runs. The white noise term is sampled from a simple Gaussian distribution:

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-1} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-1}$$

with φ_i := autoregressive terms; θ_i := moving-average terms; ε_t := white noise; p := autoregressive order; q := moving-average order;

Therefore, the Ornstein-Uhlenbeck process with Euler-Maruyama method contains a mean-reversion term, so that it stays within close bounds of the zero-mean even over a long time, which comes closer to real GNSS noise autocorrelation. The Wiener process term is also sampled from a simple Gaussian distribution:

$$X_t = X_{t-1} + \theta(\mu - X_{t-1})\Delta t + \sigma W_{t-1}\sqrt{\Delta t}$$

with θ := rate of mean reversion; μ := long-term mean; σ := volatility; W := Wiener process (standard Brownian motion); Δt := time step size;

In addition to the noise processes, we apply a moving average smoothing on the sampled noise. For high-sampled noise, the autocorrelated processes still jitter more than regular GNSS noise, which comes from the underlying Gaussian distribution randomness. With low sampling intervals, the smoothing can gradually be disabled completely, as the autocorrelation is lower, the longer the time distance between two samples is.

2.2 Noise comparison

In Figure 3 we show, for an example track, an exaggerated noise distribution of simple Gaussian noise, compared with exaggerated ARMA and OU noise. The noise generation seed is the same in all variants. It can be seen that the Gaussian noise does not yield practical results for high-sampled tracks due to the missing autocorrelation. The ARMA process yields much more authentic results, but it diverges from the zero-mean as time passes, which leads to larger differences between the input and the output track, especially at the end of the tack. The OU process is autocorrelated and also mean-reverting, and thus is the only noise process in comparison, which is suitable for long-time high-sampling noise generation. We can see that compared to ARMA, it slowly reverts towards the zero-mean as time passes. We can also see how the moving average smoothing further reduces the jittering of the noise that comes from the underlying Gaussian process.

Normal (Gaussian) ARMA Ornstein-Uhlenbeck

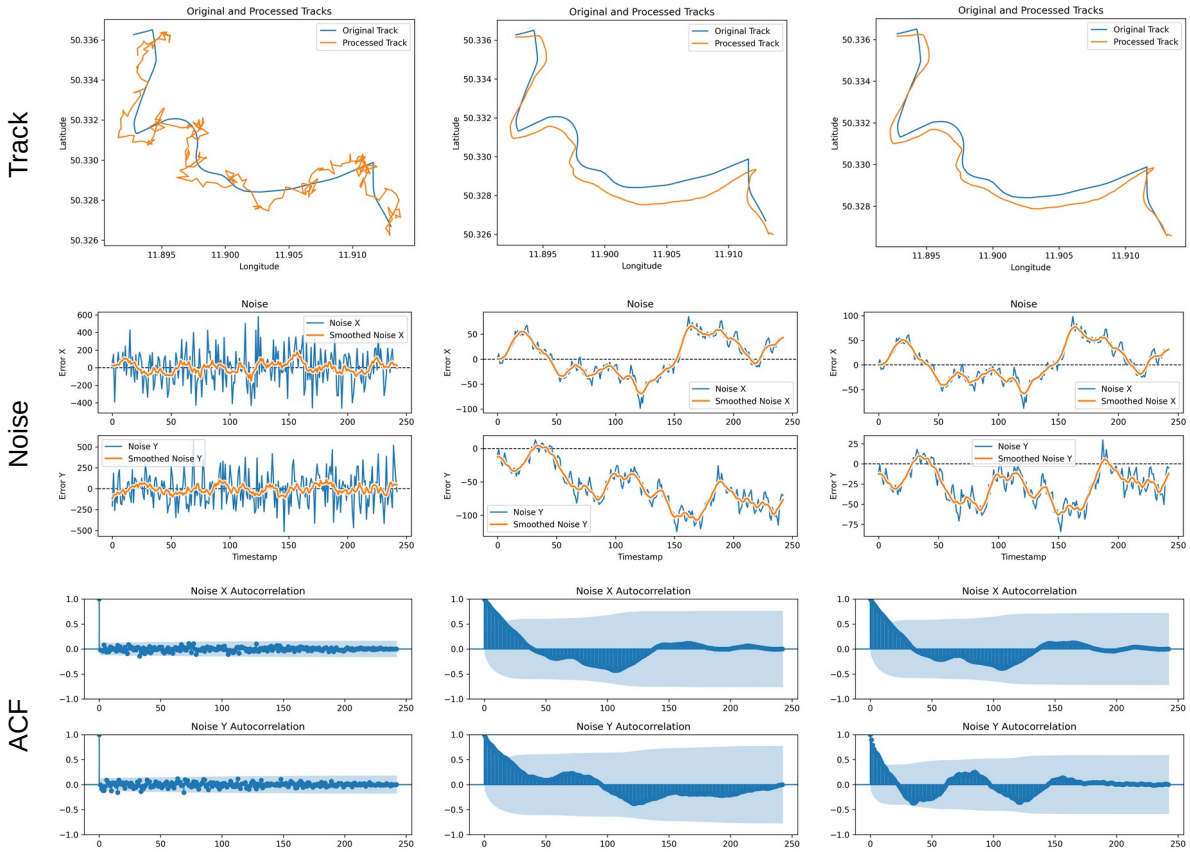


Figure 6: Noise comparison of normal (Gaussian), autoregressive moving-average (ARMA), and Ornstein-Uhlenbeck (OU) noise on an example track from the SUMO simulation. Also is shown the noise after moving average smoothing with a window size of 10. In addition, the autocorrelation functions (ACF) are shown.

2.3 Hypertuning and Noise characteristics validation

We made two experiments for validating the artificially generated GNSS noise against real-world GNSS data. In the first experiment, we compare the GNSS noise distributions from a training of artificially generated data against real GNSS data, i.e., FCD, by learning the error, outlier distribution, and outlier error parameters from our artificial noise generation process. These parameters were selected because the real-world GNSS data contains an unknown error distribution and unknown outliers. We first matched the original FCD with our “Map Matching 2” application and exported the distances of all track GNSS points to their final road positions, called “target distances”. Then we used the Python library Optuna [1] and trained the artificial error generation parameters until the distances from the artificial data, called “hypertuning distances”, were similar to the target distances. We also evaluated and tuned the outlier occurrence and outlier distance distribution. For the time delta, we used 5 s to reflect the median time delta of 5 s from the real-world target FCD data. In Figure 4, we can see that both error distance distributions are similar to a Rayleigh distribution and also similar to each other after hypertuning. This shows that our artificial noise generation process can learn the parameters from real-world GNSS data and it can generate artificial noise that has comparable distance error distributions as real-world GNSS data. In addition, we show an exemplary excerpt from the real-world target FCD and the artificially generated FCD after hypertuning. The track and the outlier distance error distributions around the roads are of a similar magnitude, although the artificial data of course consists of different routes and different concrete outliers than the real-world FCD. The real-world FCD data has

≈ 3.34 % of outliers and the artificial FCD after hypertuning has ≈ 3.38 % of generated outliers, both analyzed with Tukey's fences with $k = 3$ on the distances.

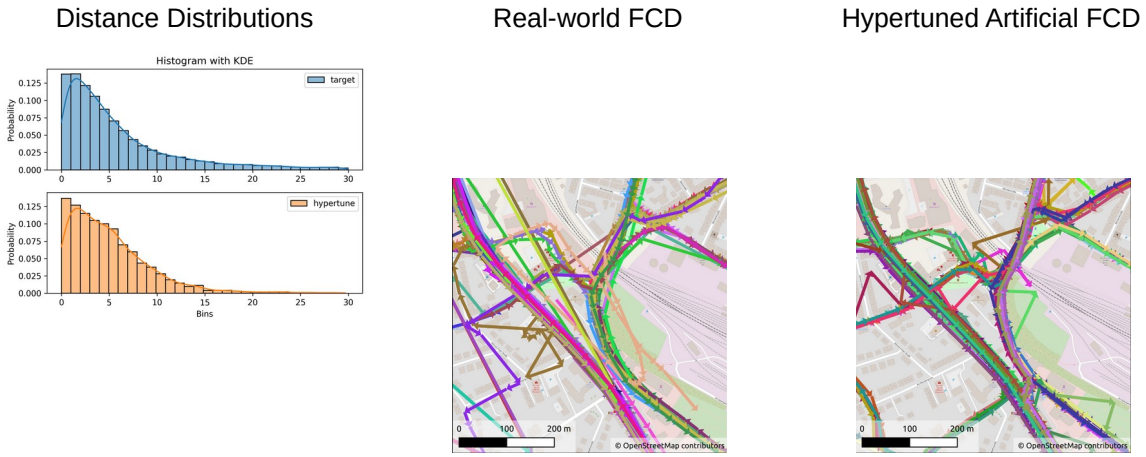


Figure 7: Hypertuning results of the GNSS error parameter estimation on real-world FCD. The left panel shows the target and hypertuned distance error distributions. The center panel shows the distance distribution and outliers of the real-world target FCD on a map. The right panel shows the distance distribution and outliers of the artificially noised FCD with the hypertuned parameters on a map.

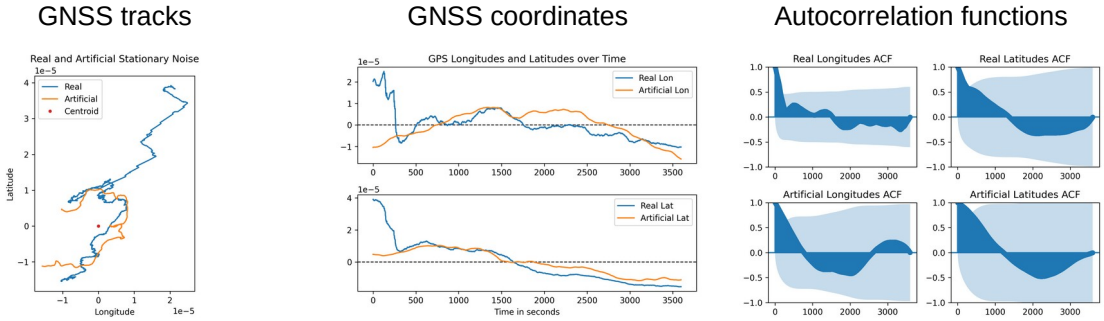


Figure 8: Hypertuning results of the GNSS error parameter estimation on a real-world stationary GNSS record of one hour. The left panel shows the real-world and the artificial GNSS track with their centroid. Both start in the north, go to the south, and lie in a similar region close to each other. The middle panel shows the coordinates in comparison. We can see that our artificial coordinates overall fit to the real curves, with some unrepresented discrepancies in smaller parts. Moreover, the real values have a very fine jittering that our moving average smoothed curves have not. The autocorrelation functions (ACF) have similar dimensions, although the longitude could not be represented as good as the latitude in this example.

In a second experiment, we recorded GNSS points with a smartphone in a static position at a window in a building for roughly one hour. We anonymized the data by translating it to the (0, 0) coordinate and then we hypertuned the artificial process on a simulated track that was created from the centroid of the recorded GNSS track as “assumed real position”. In this experiment, the error, the OU reversion speed, the moving average smoothing, and the random seed parameters were hypertuned. The parameters were chosen because the real GNSS points this time not only have an unknown error distribution but, due to being static, also a very long run-time, which needs for a finer autocorrelation and smoothing factor tuning. Moreover, we chose between 100 distinct random seed parameters to find one that represents the actual curve of the recorded track better than other seeds. In practice, the seed has no impact on the distributions on average over many random values, but for comparing a specific track shape, it helps to select a favorable seed that generates the noise in a similar direction.

Figure 5 shows that we can generate an artificial track that has a somewhat resembling track shape within similar bounds as the original real-world GNSS record. Although we could not represent the start of the curve as good as the remainder, the dimensions and the autocorrelations, especially for the latitudes, show that our process can generate fairly authentic GNSS noise even for longer durations.

We can see that our artificial noise generation has some limitations, for example, real GNSS noise contains a very fine jittering that we cannot accurately represent yet. Without moving average smoothing, the artificial jittering is too high, as shown in Figure 4. With moving average smoothing, the jittering is too small in direct comparison. With additional work, it should be possible to re-apply a fine noise again after smoothing. However, we believe that this missing fine jittering is not an impactful issue, since the jittering is very fine in comparison to the actual distance error.

Although the artificially generated GNSS noise is only a model of real-world GNSS noise, we hypothesize that this is sufficient for an evaluation, because map matching metrics are very simplified in contrast to real-world GNSS error dynamics and they also can contain terms, such as direction changes of the computed road network routes, see Figure 11 in [32], that are not even present in the GNSS tracks. As such, the artificial GNSS error does not need to be perfectly accurate, as map matching algorithms are not designed to accurately resolve GNSS errors.

3 Experiments

For analyzing the strengths and weaknesses of existing open source map matching tools, we execute a benchmark on one sample of 1000 simulated SUMO tracks on an OSM road network extract from the city of Hof, Bavaria, Germany. For generating the artificial noise, we used the hypertuned parameters as well as custom-chosen settings for testing a variety of more authentic and more extreme noise characteristics with and without simple outliers.

One limitation of our benchmark is that different map matching tools process the road network data differently. For example, some tools might remove roads that others keep, and some tools handle turn restrictions, while others ignore them. This naturally impacts the map matching results even when the original road network input is the same for all tools. However, to reflect practical real-world scenarios, in which this limitation similarly exists between the different tools, we explicitly allow and keep this situation in our experiments.

Table 1: Parameter definitions for our benchmark experiments. We define as B the baseline experiments without error applied, as H5 the hypertuned scenario with 5 s sampling interval, as E the typical error scenarios with L = low, M = mid, and H = high error, as O the outlier scenarios, and as S the sparse scenarios with 10 s to 30 s between two points. We define the error standard deviation σ, the smoothing window size w, the outlier probability p, the outlier standard deviation error σ, the time step δ, and the Ornstein-Uhlenbeck (OU) process Δt. All other parameters were the default parameters from the “apply_error.py” tool.

	H5	B	EL	EM	EH	OL	OM	OH	B10	B20	B30	S10	S20	S30
error *σ*	0.8	0	1	2	3	1	2	3	0	0	0	1	2	3
smooth *w*	2	1	10	20	30	10	20	30	1	1	1	1	1	1
outlier *p*	0.022					0.01	0.02	0.03						
outlier *σ*	43.81					30	40	50						
time *δ*	5	1	1	1	1	1	1	1	10	20	30	10	20	30
OU *Δt*	5	1	1	1	1	1	1	1	10	20	30	10	20	30

We define multiple experiment scenarios according to Table 1. The baseline scenarios are run without errors, i.e., the baseline is matching the unmodified simulated FCD without any noise added to com-

pare how well a tool matches the FCD without noise already. We hypothesize that a tool cannot surpass the baseline with the same sampling interval after noise has been added to the FCD.

Table 2: Hypertuned scenario detailed results from our benchmark. Accuracy is the “weighted mean correct fraction”, Time is in elapsed seconds, CPU resources are the sum of user and sys CPU seconds, MEM is the maximum resident set size in MiB. The Time, CPU, and MEM resources are taken from the best out of ten runs to reduce the impact of software running simultaneously.

Software	Version	Language	Accuracy	Time	CPU	MEM
Map Matching 2	1.0.5	C++	0.9953	1.41	84.76	1429
Barefoot	2021-08-18	Java	0.9737	34.30	4348.23	22,347
OSRM	5.27.1	C++	0.8955	6.58	7.08	46
GraphHopper	10	Java	0.8204	8.29	681.78	1860
Valhalla	3.5.1	C++	0.7693	8.34	64.53	7426
FMM (HMM)	2024-04-22	C++	0.7416	18.95	31.27	1538
FMM (STMatch)	2024-04-22	C++	0.5256	0.69	24.21	68

$$a = \frac{c}{max(g, m)}$$

with c := length of correct match parts; g := length of ground truth; m := length of match

The accuracy metric, i.e., correct fraction, is computed after the accuracy equation above, as was already suggested in [2, 31, 34] and in our previous work [32]. There, we computed the “mean correct fraction” as the average of all individual accuracies, without taking the match and ground truth lengths into account. In this work we compute the “weighted mean correct fraction” by computing the accuracy over the sum of all correct matched parts over the length of all matches, respective ground truth routes. This means that the accuracy metric in this work is weighted after track length, so that small tracks with low accuracy don’t impact the overall accuracy metric as much as it would be the case in the unweighted variant. As such, the accuracy results are not directly comparable with the results from our previous work.

In Table 2 we have the results from the hypertuned noise tracks with additional details about CPU and memory resource consumption. The CPU seconds are the time in seconds the application would need when only using one thread on one CPU core. If both values are close to each other, it means that the application did not scale well with multithreading. A higher difference fraction means that the individual CPU cores were used more in parallel. This table also shows the used software versions and programming languages. We can see that our “Map Matching 2” implementation has the highest accuracy, a comparably good resource consumption, and it is the second fastest solution. Barefoot is closely following, but uses much more CPU and memory resources. OSRM uses the least resources, but also has a much lower accuracy. GraphHopper again needs more CPU resources and is also less accurate. Valhalla uses the second most memory and is only slightly more accurate than FMM, which comes in last. In particular, the simpler STMatch model does not give good accuracy, although it does not use many resources either.

Table 3: Experiment results from our benchmark. The values are the accuracy metric “weighted mean correct fraction” between 0.0 and 1.0 with 0.0 having 0 % accuracy and 1.0 having 100 % accuracy.

	Map Matching 2	Barefoot	OSRM	GraphHopper	Valhalla	FMM (HMM)	FMM (STMatch)
H5	0.9953	0.9737	0.8955	0.8204	0.7693	0.7416	0.5256
B	0.9968	0.9960	0.8753	0.9860	0.9321	0.9955	0.5678
EL	0.9960	0.9950	0.7404	0.6407	0.9296	0.5644	0.2517
EM	0.9931	0.9779	0.2843	0.2795	0.9159	0.3704	0.1591
EH	0.9823	0.9275	0.1632	0.2024	0.8491	0.2842	0.1230
OL	0.9957	0.9706	0.6099	0.4996	0.5964	0.4747	0.2086
OM	0.9843	0.9018	0.1677	0.2239	0.2601	0.2955	0.1163
OH	0.9525	0.8006	0.0805	0.1618	0.1495	0.2122	0.0812
B10	0.9951	0.9964	0.9529	0.9857	0.9856	0.9908	0.9520
B20	0.9904	0.9933	0.9424	0.9832	0.9870	0.9802	0.9803
B30	0.9830	0.9883	0.9493	0.9804	0.9817	0.9655	0.9647
S10	0.9936	0.9937	0.9265	0.8798	0.9826	0.8251	0.6540
S20	0.9839	0.9508	0.6881	0.7736	0.9808	0.6861	0.5691
S30	0.9603	0.8773	0.4795	0.7300	0.9220	0.6220	0.5355

In Table 3 we can see the accuracy metrics of all evaluated track sets in comparison with all tools. We can see that our own “Map Matching 2” implementation achieves high accuracy in all situations, even with large outliers (OH) where most other tools drastically degrade in accuracy. Barefoot often stays close and also handles noise comparably well, and it is the only other tool that can tolerate outliers up to some extent. In some of the sparse scenarios, it even has slightly higher accuracy than our solution, but it is less accurate than our solution in dense scenarios with 1 s sampling interval and when the error is higher. OSRM and GraphHopper both benefit from sparse-sampled tracks and degrade quickly with higher error and especially with outliers applied. Valhalla is not so sensitive to higher error and has good accuracy in high and especially in low-sampling situations (S10, S20), but it degrades a lot with outliers. FMM only achieves some accuracy with sparse and low error, especially with its simpler STMatch model. Already in the baseline scenario with 1 s sampling interval, the STMatch model of FMM has an exceptionally low accuracy, although no noise has been applied yet. Overall, our own implementation is the most robust in comparison and can be recommended for any tracks with any noise characteristics. We also have discovered several individual strengths and weaknesses of the other open source map matching tools, depending on the noise and sampling interval characteristics of the GNSS tracks.

4 Discussion

In this work, we have presented a new approach for generating artificial ground truth map matching data sets with open source software based on traffic simulation with artificial autocorrelated noise generation. We have shown that our new approach can generate authentic synthetic noise within certain limitations and also intentional extreme noise for worst case analysis. We have also shown how this

process can be used to reveal hidden strengths and weaknesses of map matching algorithms in practice.

Our artificial noise still has some limitations. Currently, we only generate simple outliers by moving individual points, selected by a closeness probability parameter, to a random position from a Gaussian distribution that is independent of the OU process. Additional analysis of real-world outliers is required to further improve the authenticity of the artificial outliers and to make them more diverse.

Furthermore, the application of traffic micro-simulation may limit the size of the road networks that can be used efficiently. However, the amount of GNSS tracks that can be generated by our approach, and that we have artificially generated in this work, already exceeds all existing map matching data sets combined. In this way, our new approach of artificially generating map matching ground truth data facilitates further improvements in map matching research and development.

Data and codes availability

Our open source tools and benchmark are available under AGPL 3.0 (https://www.gnu.org/licenses/agpl-3.0.en.html) at FigShare (https://doi.org/10.6084/m9.figshare.27952545).

Contact to the Author:

M.Sc. Adrian Wöltche
Hochschule für Angewandte Wissenschaften Hof
Alfons-Goppel-Platz 1, 95028 Hof
+49 9281 409 6277
adrian.woeltche@iisys.de

Bibliography

[1] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A Next-generation Hyperparameter Optimization Framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, pages 2623–2631, New York, NY, USA, 7 2019. ACM. doi:10.1145/3292500.3330701.

[2] Georgia Ayfantopoulou, Marios Nikolaos Militsis, Josep Maria Salanova Grau, and Socrates Basbas. Improving Map Matching of Floating Car Data with Artificial Intelligence Techniques. Information, 13(11):508, 10 2022. doi:10.3390/info13110508.

[3] Nelson Baza Solares, Ruben Velasquez Martínez, Cristian Torres Bohórquez, Yerly Martínez Estupiñán, and Cristian Poliziani. Traffic Simulation with Open-Source and Commercial Traffic Microsimulators: A Case Study. Communications - Scientific letters of the University of Zilina, 24(2):E49–E62, 4 2022. doi:10.26552/com.c.2022.2.e49-e62.

[4] Pingfu Chao, Yehong Xu, Wen Hua, and Xiaofang Zhou. A Survey on Map-Matching Algorithms. In Databases Theory and Applications, Lecture Notes in Computer Science, pages 121–133. Springer International Publishing, Cham, 1 2020. doi:10.1007/978-3-030-39469-1_10.

[5] Lara Codeca, Jakob Erdmann, Vinny Cahill, and Jerome Haerri. SAGA: An Activity-based Multimodal Mobility Scenario Generator for SUMO. SUMO Conference Proceedings, 1:39–58, 6 2022. doi:10.52825/scp.v1i.99.

[6] Corina Dima and Cosmin Popa. A review of map-matching algorithms. Annals of the "Dunarea de Jos" University of Galati. Fascicle II, Mathematics, Physics, Theoretical Mechanics, 2(46):121–125, 12 2023. doi:10.35219/ann-ugal-math-phys-mec.2023.2.09.

[7] Jakob Erdmann and Rüdiger Ebendt. Benchmarking Trajectory Matchers with SUMO. In European Simulation and Modelling Conference, 28th European Simulation and Modelling Conference, pages 226–230, Porto, Portugal, 10 2014. Eurosis-ETI. URL: https://elib.dlr.de/93676/.

[8] Vincenzo Pasquale Giofre, Vittorio Astarita, Giuseppe Guido, and Alessandro Vitale. Localization issues in the use of ITS. In 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS, pages 427–432. IEEE, 6 2017. doi:10.1109/mtits.2017.8005710.

[9] Elans Grabs, Ansis Klūga, and Alberts Zeļenkovs. Accuracy Estimation of GPS Receiver Parameters with Re-Reference System in Static Mode. Telecommunications and Electronics, 8:31–36, 3 2008. URL: https://ortus.rtu.lv/science/en/publications/4827.

[10] Ajay Kumar Gupta and Udai Shanker. A Comprehensive Review of Map-Matching Techniques: Empirical Analysis, Taxonomy, and Emerging Research Trends. International Journal of Web Services Research (IJWSR), 19(1):1–32, 9 2022. doi:10.4018/ijwsr.306243.

[11] Zhengfeng Huang, Shaojie Qiao, Nan Han, Chang-an Yuan, Xuejiang Song, and Yueqiang Xiao. Survey on vehicle map matching techniques. CAAI Transactions on Intelligence Technology, 6 2020. doi:10.1049/cit2.12030.

[12] K. Izet Ünsalan and M. Ünsalan. Classroom study of GNSS position accuracy using smartphones. Scientific Bulletin of Naval Academy, XXIII(1):83–89, 7 2020. doi:10.21279/1454-864x-20-i1-011.

[13] Serđo Kos, David Brčić, and Ivo Musulin. Smartphone application GPS performance during various space weather conditions: A preliminary study. In Proceedings of the 21st International Symposium on Electronics in Transport, ISEP 2013, pages 15–18, Ljubljana, Slovenija, 3 2013. Elektrotehniška zveza Slovenije. URL: https://www.bib.irb.hr/623174.

[14] Daniel Krajzewicz, Georg Hertkorn, Christian Rössel, and Peter Wagner. SUMO (Simulation of Urban MObility) - An open-source traffic simulation. In A. Al Akaidi, editor, Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM20002), 4th Middle East Symposium on Simulation and Modelling, pages 183–187, 8 2002. URL: https://elib.dlr.de/6661/.

[15] Matej Kubicka, Arben Cela, Philippe Moulin, Hugues Mounier, and S.I. Niculescu. Dataset for testing and training of map-matching algorithms. In 2015 IEEE Intelligent Vehicles Symposium (IV). IEEE, 6 2015. doi:10.1109/ivs.2015.7225829.

[16] Hengfeng Li, Lars Kulik, and Kotagiri Ramamohanarao. Spatio-temporal trajectory simplification for inferring travel paths. In Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL '14, pages 63–72, New York, NY, USA, 11 2014. Association for Computing Machinery. doi:10.1145/2666310.2666409.

[17] Hugo Lindholm. Map Matching to road segments using Hidden Markov Model with GNSS, Odometer and Gyroscope. Master's thesis, Linköping University, Linköping, Sweden, 1 2020. URL: https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-162706.

[18] Yin Lou, Chengyang Zhang, Yu Zheng, Xing Xie, Wei Wang, and Yan Huang. Map-matching for low-sampling-rate GPS trajectories. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '09, pages 352–361, New York, NY, USA, 11 2009. Association for Computing Machinery. URL: https://www.microsoft.com/en-us/research/publication/map-matching-for-low-sampling-rate-gps-trajectories/, doi:10.1145/1653771.1653820.

[19] Thomas Louail, Maxime Lenormand, Oliva García Cantú, Miguel Picornell, Ricardo Herranz, Enrique Frias Martinez, José J. Ramasco, and Marc Barthelemy. From mobile phone data to the spatial structure of cities. Scientific Reports, 4(1):5276–5287, 6 2014. doi:10.1038/srep05276.

[20] Dennis Luxen and Christian Vetter. Real-time routing with OpenStreetMap data. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '11, pages 513–516, New York, New York, USA, 11 2011. ACM Press. doi:10.1145/2093973.2094062.

[21] Xiaoyi Ma, Xiaowei Hu, Thomas Weber, and Dieter Schramm. Evaluation of Accuracy of Traffic Flow Generation in SUMO. Applied Sciences, 11(6):2584, 3 2021. doi:10.3390/app11062584.

[22] Juliette Marais, Marion Berbineau, and Marc Heddebaut. Land Mobile GNSS Availability and Multipath Evaluation Tool. IEEE Transactions on Vehicular Technology, 54(5):1697–1704, 11 2005. doi:10.1109/tvt.2005.853461.

[23] Ali Mohamed, Travis Rautman, John Krumm, and Ankur Teredesai. ACM SIGSPATIAL GIS Cup 2012. In ACM SIGSPATIAL GIS '12, November 6-9, 2012. Redondo Beach, CA, USA, SIGSPATIAL '12, 11 2012. URL: https://www.microsoft.com/en-us/research/publication/acm-sigspatial-gis-cup-2012/.

[24] Paul Newson and John Krumm. Hidden markov map matching through noise and sparseness. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '09, pages 336–343, New York, NY, USA, 11 2009. Association for Computing Machinery. URL: https://www.microsoft.com/en-us/research/publication/hidden-markov-map-matching-noise-sparseness/, doi:10.1145/1653771.1653818.

[25] Mohammed A. Quddus, Washington Y. Ochieng, and Robert B. Noland. Current map-matching algorithms for transport applications: State-of-the art and future research directions. Transportation Research Part C: Emerging Technologies, 15(5):312–328, 10 2007. doi:10.1016/j.trc.2007.05.002.

[26] Peter Ranacher, Richard Brunauer, Wolfgang Trutschnig, Stefan Van der Spek, and Siegfried Reich. Why GPS makes distances bigger than they are. International Journal of Geographical Information Science, 30(2):316–333, 2 2016. doi:10.1080/13658816.2015.1086924.

[27] Ali Rostami, Bin Cheng, Hongsheng Lu, John B. Kenney, and Marco Gruteser. A Light-Weight Smartphone GPS Error Model for Simulation. In 2019 IEEE 90th Vehicular Technology Conference, VTC2019-Fall, pages 1–5. IEEE, 9 2019. doi:10.1109/vtcfall.2019.8891089.

[28] Joerg Schweizer, Cristian Poliziani, Federico Rupi, Davide Morgano, and Mattia Magi. Building a Large-Scale Micro-Simulation Transport Scenario Using Big Data. ISPRS International Journal of Geo-Information, 10(3):165, 3 2021. doi:10.3390/ijgi10030165.

[29] Andy W. R. Soundy, Bradley J. Panckhurst, Phillip Brown, Andrew Martin, Timothy C. A. Molteno, and Daniel Schumayer. Comparison of Enhanced Noise Model Performance Based on Analysis of Civilian GPS Data. Sensors, 20(21):6050, 10 2020. doi:10.3390/s20216050.

[30] Frank van Diggelen and Per Enge. The World's first GPS MOOC and Worldwide Laboratory using Smartphones. In Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation, ION GNSS+ 2015, pages 361 – 369, Tampa, Florida, 9 2015. Tampa Convention Center. URL: https://www.ion.org/publications/abstract.cfm?articleID=13079.

[31] Hao Wu, Jiangyun Mao, Weiwei Sun, Baihua Zheng, Hanyuan Zhang, Ziyang Chen, and Wei Wang. Probabilistic Robust Route Recovery with Spatio-Temporal Dynamics. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16, pages 1915–1924, New York, NY, USA, 8 2016. Association for Computing Machinery. doi:10.1145/2939672.2939843.

[32] Adrian Wöltche. Open source map matching with Markov decision processes: A new method and a detailed benchmark with existing approaches. Transactions in GIS, 27(7):1959–1991, 10 2023. doi:10.1111/tgis.13107.

[33] Can Yang and Győző Gidófalvi. Fast map matching, an algorithm integrating hidden markov model with precomputation. International Journal of Geographical Information Science, 32(3):547–570, 3 2018. doi:10.1080/13658816.2017.1400548.

[34] Kai Zheng, Yu Zheng, Xing Xie, and Xiaofang Zhou. Reducing Uncertainty of Low-Sampling-Rate Trajectories. In 2012 IEEE International Conference on Data Engineering, ICDE 2012, pages 1144–1155. IEEE, 4 2012. doi:10.1109/icde.2012.42.

[35] Subhrasankha Dey, Martin Tomko, and Stephan Winter. Map-Matching Error Identification in the Absence of Ground Truth. ISPRS International Journal of Geo-Information, 11(11):538, 10 2022. doi:10.3390/ijgi11110538.

[36] Jan-Henrik Haunert and Benedikt Budig. An algorithm for map matching given incomplete road data. In Proceedings of the 20th International Conference on Advances in Geographic Information Systems, pages 510–513, New York, NY, USA, 11 2012. Association for Computing Machinery. doi:10.1145/2424321.2424402.

[1] A. Wöltche, "Map Matching 2", https://github.com/iisys-hof/map-matching-2

[2] A. Wöltche, "Open source map matching with Markov decision processes: A new method and a detailed benchmark with existing approaches", Transactions in GIS, vol. 27, no. 7, pp. 1959–1991, Oct. 2023, doi: https://doi.org/10.1111/tgis.13107

Efficient processing of high-volume spatial data with Spark

Alexey Egorov, Dr. Jannis Jakobi

Efficient retrieval of geospatial data is crucial but presents scaling challenges. During our transition from PostgreSQL to Apache Spark, we encountered limitations in spatial indexing. While PostgreSQL's indexing supports efficient queries, this is not directly translatable to Spark. The transition required us to create new strategies for managing and querying spatial data effectively. In this talk, we'll share the challenges we faced and the innovative solutions we implemented to address them.

We are the data team at viadukt, a start-up focused on advancing energy-efficient building modernization across Germany. To support this, we set up a comprehensive geospatial database containing building-related data and developed a processing pipeline that integrates both federal and open-source datasets.

As we transitioned from PostgreSQL to Apache Spark, we needed new, efficient approaches for spatial data processing. PostgreSQL's spatial indexing is key to its performance, using bounding boxes around geometries to pre-filter matches before exact calculations. Apache Sedona, which supports large-scale spatial processing, includes spatial indexing but encounters limitations with larger datasets. Without traditional indexing in Spark, we applied partitioning techniques based on H3 cells or Geohashes.

In this talk, we'll explore how to apply these partitioning methods effectively and why they offer advantages. Additionally, we'll dive into the specific benefits and challenges of this approach for high-volume spatial data and provide examples.

Scalable big data processing – Postgres' drawbacks eliminated with Spark

Alexey Egorov, Dr. Jannis Jakobi

PostgreSQL is often considered a standard solution for geospatial data processing. However, compute costs grow with the data volume and vertical scaling quickly becomes expensive. In contrast, distributed processing frameworks allow for horizontal scaling. In this talk, we will present our experience with Apache Spark, an open-source framework designed for high-volume data processing. We will show the benefits and highlight the challenges we faced during the implementation.

We are from the data team of viadukt, a start-up that aims to increase the energy-efficient modernization of buildings across Germany. To achieve this, we set up a comprehensive geospatial database containing building-related data and developed a processing pipeline that integrates both federal and open-source datasets. Initially, PostgreSQL with the PostGIS extension met our needs while we focused on North Rhine-Westphalia. But as our scope expanded nationwide, significant challenges arose.

PostgreSQL required costly vertical scaling, even though we only occasionally needed high processing power, resulting in an underutilized yet expensive setup. Alternatively, we could have downsized our machines whenever processing was not relevant. However, this would require downtime of our instance, which we wanted to avoid, or complex mechanisms to avoid it.

To address these issues, we implemented a solution based on Apache Spark and Apache Sedona that met specific criteria: cost-effectiveness, flexible scaling for storage and compute, consistent performance, and an easy migration path. In this talk, we'll explore how we overcame challenges we faced during the migration and the solutions we implemented.

Modellgetriebene XPlanung: von UML zur OGC API for Features und mehr

Tobias Kraft

Landesbetrieb Geoinformation und Vermessung Hamburg

Einleitung

XPlanung ist der bundesweite Datenstandard für die Inhalte raumbezogener Planwerke wie beispielsweise Bebauungspläne [1]. Er basiert auf der ISO 19100 Serie bzw. GML 3.2: die konzeptionelle Modellierung erfolgt in UML, woraus ein GML-Anwendungsschema als Encoding Standard abgeleitet wird. Da die Inhalte der Planwerke verlustfrei erfasst werden können sollen und hier weitreichende, miteinander verknüpfte Regelungsmöglichkeiten bestehen, ist das Datenmodell durch ein hohes Maß an Komplexität geprägt: es gibt zahlreiche Objektarten, zwischen denen diverse Relationen bestehen, mehrfach belegbare Attribute und komplexe Datentypen.

Diese Komplexität stellt eine Herausforderung für die Implementierung des Standards in tabellarisch bzw. Simple Features orientierten GI-Systemen und entsprechenden Datenbanken dar. Eine vollständig relationale Abbildung des Modells in SQL DDL resultiert beispielsweise in über 4000 Tabellen. Hinzukommt, dass der Release-Zyklus von XPlanung bisher vergleichsweise kurz ist, was je nach Art der Implementierung einen hohen Pflegeaufwand bedeuten kann, insbesondere, wenn das XPlanungsmodell auf eine interne Repräsentation gemappt wird.

In diesem Kontext hat die beim Landesbetrieb Geoinformation und Vermessung in Hamburg angesiedelte, für den Betrieb des Standards verantwortliche XLeitstelle Planen und Bauen begonnen, über die bestehende modellgetriebene Produktion des Encoding Standards hinaus weitere Ansätze für modellgetriebene Prozesse zu entwickeln. Resultat sind u.a. die im Folgenden erläuterten Bausteine einer modellgetriebenen Architektur zur Bereitstellung von XPlanungsdaten per OGC API for Features (OAPIF).

Bausteine der modellgetriebenen Architektur

Tools zur Generierung von Implementierungsmodellen

Ein zentrales Werkzeug ist dabei die die im OGC-Umfeld etablierte Open Source Software Shapechange [2] der Fa. interactive instruments, mit der sich verschiedene Encodings und Implementierungsmodelle – sog. Targets – aus einem konzeptionellen Modell ableiten lassen. Hiermit wird ein spezielles JSON Schema erzeugt, das möglichst viele Informationen aus dem UML Modell (u.a. Enumerationswerte, Stereotypen) transportiert und als Input für das Tool Datamodel Code Generator [3] dient, mit dem eine Python Repräsentation des XPlanungsmodells in Form von Pydantic [4] Klassen generiert wird (siehe Abb. 1).

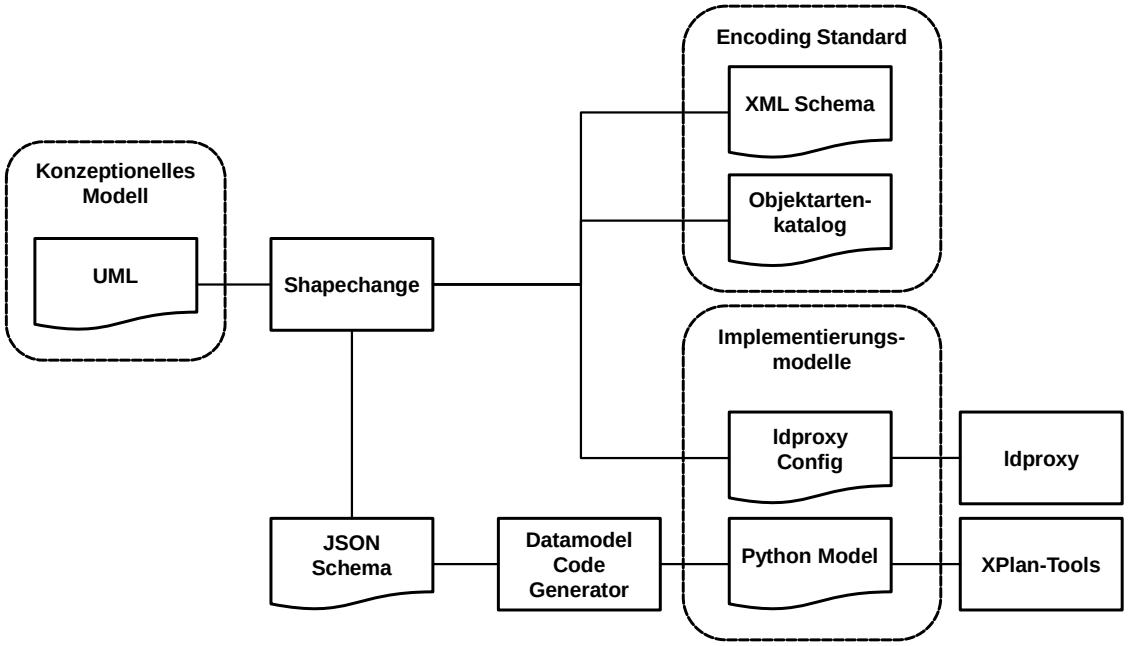


Abb. 1: Generierung von Implementierungsmodellen aus dem konzeptionellen Modell

XPlan-Tools

Diese Python Repräsentation bildet als internes Datenmodell das Fundament für die von der XLeitstelle entwickelte Python-Bibliothek XPlan-Tools [5], über die verschiedene Funktionen zur Verarbeitung von XPlanungsdaten implementiert sind, u.a.:

- Konvertierung zwischen unterschiedlichen Formaten, insbesondere XPlanGML, XPlanJSON nach der JSON-FG Spezifikation [6] und einer Datenbank-Repräsentation mit Unterstützung für PostGIS, SpatiaLite und GeoPackage
- Transformation von XPlanung nach INSPIRE-PLU auf Grundlage der offiziellen Mappings der XLeitstelle
- Migration von XPlanung v5 nach v6 (v4 nach v5 in Planung)
- CRUD-Operationen für unterstützte Datenbanken

XPlan-Tools kann als Standalone-Anwendung per CLI genutzt werden, dient aber insbesondere auch als Basisbibliothek für weiterführende Projekte (s.u.). Damit steht ein schnell und ohne Abhängigkeit von externen Dienstleistern erweiterbares, ins umfangreiche Python-Ökosystem eingebettetes Werkzeug zur Verfügung, das auf Grund des modellgetriebenen Ansatzes gewissermaßen auf Knopfdruck neue XPlanungsversionen unterstützt.

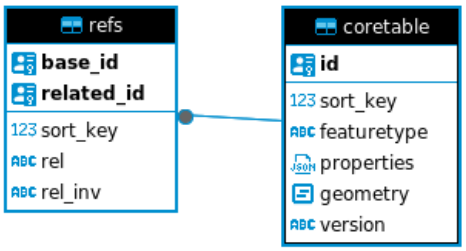


Abb. 2: Coretable-Ansatz für XPlanung

Die o.g. Datenbank-Repräsentation ist inspiriert von dem im Rahmen des AdV-Smart/Basemap.de-Projektes entwickelten Coretable-Ansatz [8], der SQL und NoSQL in einem hybriden Datenmodell kombiniert: alle Features werden in einer Tabelle, der Coretable, geführt, wobei die Properties in einer JSON(B)-Spalte gespeichert sind. Für XPlanung wurde dieser Ansatz modifiziert, indem die Relationen in eine zweite (Assoziations-) Tabelle ausgelagert wurden (siehe Abb. 2).

Dadurch ist es möglich, das vollständige XPlanungsmodell – auch versionsübergreifend – statt mit über 4000 mit 2 Tabellen abzubilden, während die referenzielle Integrität weiterhin durch RDBMS-Kernfunktionalität in Form von Foreign Keys gewährleistet wird und über PostGIS/SpatiaLite im Vergleich zu NoSQL-Datenbanken fortgeschrittene Geometriefunktionen und -kodierung vorhanden sind.

ldproxy

Das ldproxy-Framework [8] dient der Bereitstellung von OGC APIs und unterstützt auch komplexe Datenmodelle. Es wird ebenso wie Shapechange von der Fa. interactive instruments entwickelt, weswegen ein entsprechendes Shapechange Target zur Generierung einer ldproxy-Konfiguration bereits existierte. Dieses wurde in Zusammenarbeit mit der XLeitstelle jedoch so erweitert, dass eine Coretable-basierte Datenbank direkt und ohne weitere Prozessierung über (Materialized) Views o.ä. als Provider genutzt werden kann [9]. Somit können beliebige komplexe XPlanungsdaten verlustfrei per OAPIF bereitgestellt werden, vgl. [10]. Dabei sind unterschiedliche Encodings inkl. XPlanGML und XPlanJSON sowie der Abruf vollständiger Pläne mittels Stored Query möglich. Der Datenimport erfolgt mit XPlan-Tools anhand der dort implementierten Konvertierung von XPlanGML (oder anderen unterstützen Formaten) in eine Coretable DB (siehe Abb. 3).

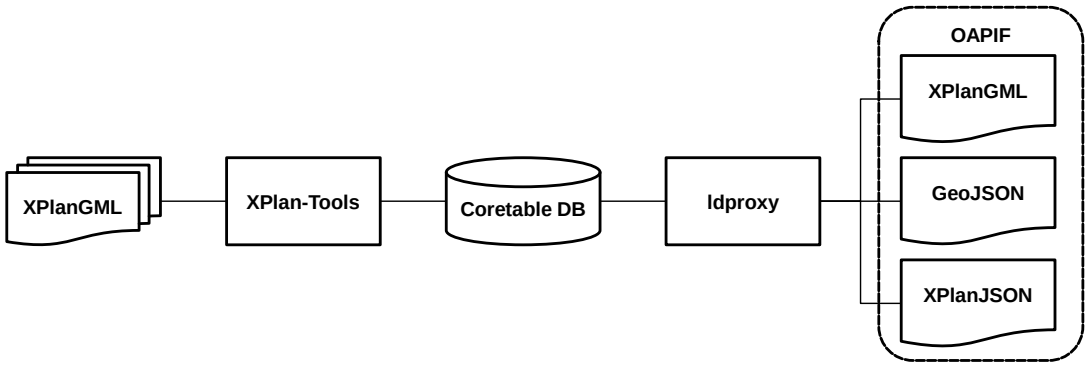


Abb. 3: OAPIF-Workflow

Ausblick

Durch den oben skizzierten Ansatz und die im Zuge dessen entwickelten Bausteine – insbesondere XPlan-Tools und die damit verbundene Python-Repräsentation des XPlanungsmodells – ergeben sich

neue Perspektiven. Ein wesentliches Ziel ist dabei die modellgetriebene Erfassung von XPlanungsdaten.

In diesem Zusammenhang wurde bereits ein Prototyp für die Generierung von HTML-basierten Attributformularen aus den XPlanungs-Pythonklassen unter Verwendung des NiceGUI-Frameworks [11] entwickelt, wofür auf Grund der im Modell enthaltenen umfangreichen Informationen u.a. zu Daten- und Stereotypen sowie Dokumentationsfeldern beste Voraussetzungen bestehen. So ist es durch die in XPlan-Tools schon implementierten CRUD-Funktionen möglich, Objekte in einer Coretable DB mit der prototypischen Webanwendung zu editieren, zu speichern und zu löschen. Derzeit wird daran gearbeitet, die Attributformulare mittels eines Plugins in QGIS zu integrieren und dort mit der Geometriedigitalisierung zu kombinieren. Eine erste „Proof-of-Concept“ Version wird für Q1 2025 angestrebt.

Darüber hinaus lässt sich der Ansatz auf andere ISO-19100-basierte Anwendungsschemas wie INSPIRE oder 3A übertragen. So wird für die INSPIRE-Transformation eine Python-Repräsentation des INSPIRE-PLU-Modells verwendet, die theoretisch auch an anderen Stellen im modellgetriebenen Workflow zum Einsatz kommen könnte, etwa für den Import von INSPIRE-PLU-Daten in eine Coretable-DB mit anschließender Veröffentlichung per OAPIF. Auch die Erfassung von 3A-Daten mittels QGIS wäre in diesem Kontext denkbar und könnte weiter untersucht werden.

Kontakt zum Autor:

Tobias Kraft
Landesbetrieb Geoinformation und Vermessung
Neuenfelder Str. 19, 21109 Hamburg
(040) 428 26 - 5912
Tobias.Kraft@gv.hamburg.de

Literatur

[1] Homepage der XLeitstelle zu XPlanung: https://xleitstelle.de/xplanung

[2] Shapechange bei Github: https://github.com/ShapeChange/ShapeChange

[3] Datamodel Code Generator bei Github: https://github.com/koxudaxi/datamodel-code-generator

[4] Pydantic bei Github: https://github.com/pydantic/pydantic

[5] XPlan-Tools bei OpenCoDE: https://gitlab.opencode.de/xleitstelle/xplanung/xplan-tools

[6] OGC Features and Geometries JSON: https://github.com/opengeospatial/ogc-feat-geo-json

[7] AdV Smart Mapping Workshop 2022, Vortrag Andres Schmidt zu internen Datenmodellen und Prozessierungen: https://basemap.de/data/website/workshops/02_2022/block2_v2_teil1_andreas_schmidt_interne_datenmodelle_und_prozessierungen.mp4

[8] ldprody bei Github: https://github.com/ldproxy/ldproxy

[9] Shapechange Dokumentation zum Coretable Approach: https://shapechange.github.io/ShapeChange/3.1.0/targets/ldproxy2.html#ldproxy2_schemaconversionrules_specificrules_coretable

[10] ldproxy-Demoinstanz mit XPlanungsdaten als Coretable: https://coretable.ldproxy.net/

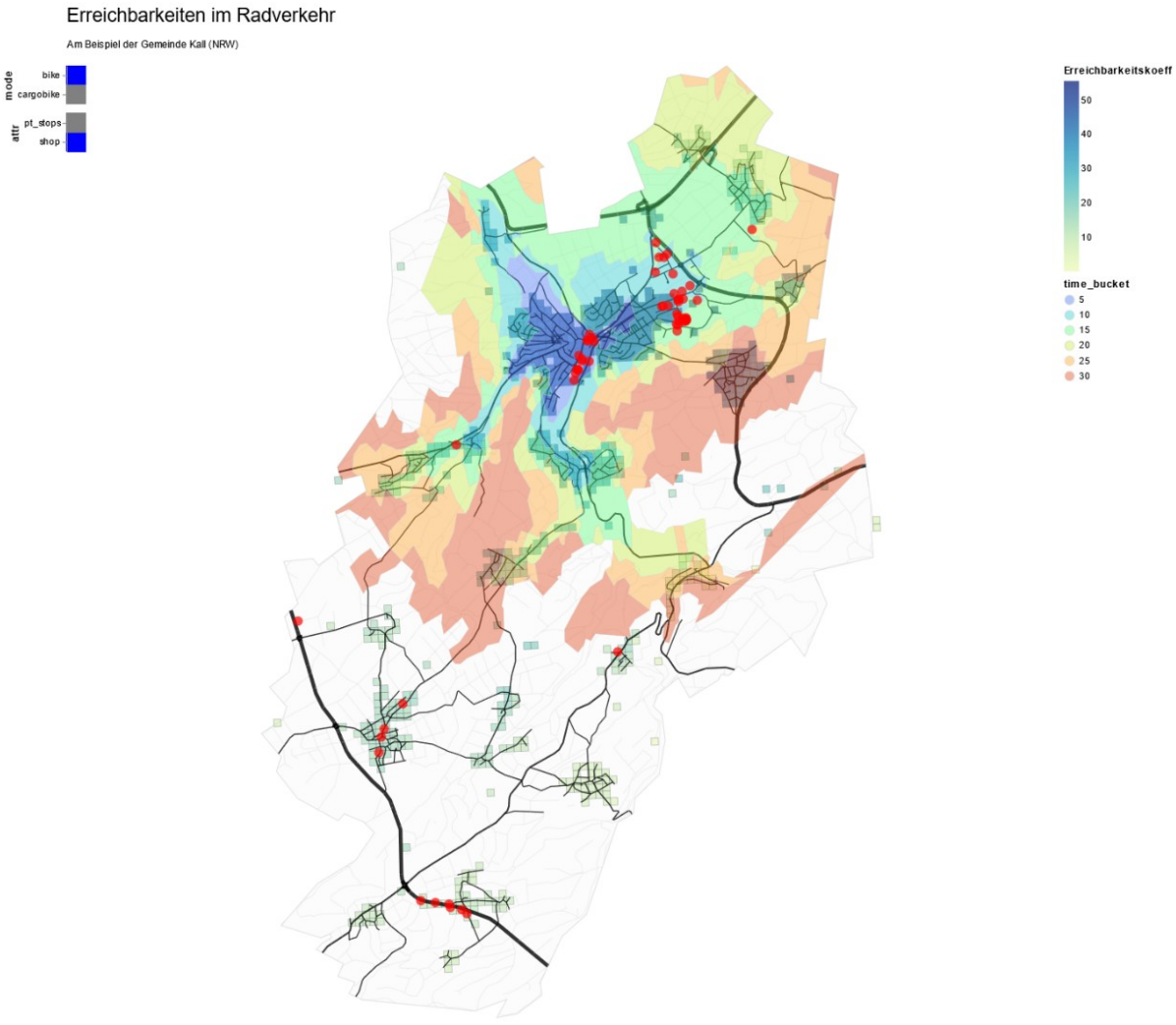
[11] Homepage von NiceGUI: https://nicegui.io/

Erreichbarkeitsanalyse in ländlichen Räumen: Potenziale des Radverkehrs

Simon Metzler

Das Titelbild des Beitrags

Der Beitrag untersucht die Erreichbarkeit von Versorgungszielen in ländlichen Räumen per Rad. Mit einer offenen Methode werden Isochronen für Startpunkte berechnet um deren Erreichbarkeit zu Attraktionspunkten zu bewerten. Die Verfahrenspipeline, entwickelt in Python, ermöglicht eine reproduzierbare Analyse und Vergleichbarkeit zwischen Routingprofilen. Die Methode wurde exemplarisch in Kall getestet und soll nun um zusätzliche Routingprofile, Attraktionspunkte und Gebiete erweitert werden.



Erreichbarkeits- und Versorgungsprobleme nehmen mit abnehmender Siedlungsdichte zu und betreffen besonders ländliche Räume. Knappe finanzielle Ressourcen der Kreise und Kommunen führen dazu, dass die Zahl der Einrichtungen der Daseinsvorsorge niedrig ist oder verringert wird. Eine strategische Planung zur Erreichung versorgungsrelevanter Einrichtungen ist daher ausschlaggebend für

die Lebenswirklichkeit und -qualität der Menschen vor Ort. Die Haushaltslage und geringere Nachfrage wirkt sich in der Regel negativ auf den ÖPNV und die Anbindung bzw. Wegelängen in ländlichen Räumen aus. Gleichzeitig ändern sich auch in ländlichen Bereichen Lebensstile, Erwerbsmuster und Lebensformen. Daraus entstehen neue Anforderungen an die Gestaltung der Alltagsmobilität und Sicherung der Erreichbarkeiten.

Dieser Beitrag adressiert konkret die Fragestellung, wie der Einfluss des Radverkehrs bzw. von Radverkehrsnetzen auf die Erreichbarkeit von Einrichtungen der Daseinsvorsorge in ländlichen Räumen sichtbar gemacht werden kann. Um dieser Frage nachgehen zu können haben wir ein offenes, nichtproprietäres Verfahren entwickelt, mit dem wir die Erreichbarkeiten mit dem Fahrrad zu den Zielen z. B. der Daseinsvorsorge darstellen können. Durch räumliche Daten und berechnete Fahrzeiten wird die Versorgungssituation von verschiedenen Standorten modelliert und eine Vergleichbarkeit zwischen Erreichbarkeiten mit bspw. Fahrrädern und Pedelecs hergestellt. Methodisch wurde hierfür folgendes Vorgehen gewählt:

Für die Bewertung der Erreichbarkeiten werden zunächst die Fahrzeiten bzw. Isochronen ausgehend von möglichen Startpunkten (100 Meter-Gitter, Bevölkerungsraster Zensus 2011 vom Statistisches Bundesamt [1]) berechnet und diese im Anschluss nach Anzahl und Gewichtung der jeweils möglichen Ziele bewertet. Zur Isochronenberechnung wird die offene Routing-Software Graphhopper [2] auf Basis von OpenStreetMap genutzt. Die Attraktionsorte bzw. Ziele stammen insbesondere aus OpenStreetMap, wobei die ÖV-Haltepunkte aus GTFS-Daten [3] stammen um für die Gewichtung die Anzahl der Abfahrten miteinfließen lassen zu können. Die Verfahrenspipeline wird in Python ausgeführt, wobei sequenziell und reproduzierbar die Analyse zusammengeführt wird.

Die Analyse wurde exemplarisch im Gemeindegebiet von Kall durchgeführt. Die Gemeinde Kall (ca. 11.000 EW) ist als Grundzentrum klassifiziert und liegt im Kreis Euskirchen (NRW) im Bereich der Nordeifel. Ein Prototyp der Analyse ist hier zu finden:

https://radverkehr.github.io/erreichbarkeiten/showcase_kall/viz/
kall_map_export_30min_6buckets_attr_access_table_modes_title.html

Der Vortrag soll zum Anlass genommen werden um die beiden Routingprofile „bike“ (konventionelles Fahrrad) und „cargobike“ (elektrisch-motorisiertes Lastenrad) um mind. ein weiteres Routingprofil bspw. für „sicheres Radfahren“ (Vermeidung von Mischverkehr mit Kfz) erweitert zu werden.

Außerdem sollen die Attraktionspunkte „pt_stops“ (Haltepunkte des öffentlichen Verkehrs) und „shop“ Einkaufmöglichkeiten um weitere Attraktionspunkte wie z.B. Bildungs- und Gesundheitseinrichtungen und Sportstätten erweitert werden. Zuletzt streben wir an, die öffentlich-verfügbare Anwendung auf mehrere PLZ-Gebiete eines Landkreises/Bundeslandes auszurollen.

In unserem Vortrag möchten wir die Entwicklungsansätze des Verfahrens vorstellen und seine Nutzbarkeit als Instrument für eine zielgerichtete Radverkehrsplanung und damit für die Verbesserung der Erreichbarkeiten bzw. die Mobilitätsoptionen diskutieren.

Inspiriert wurde das Vorgehen vom FOSSGIS-Konferenz 2023 Vortrag von Konstantin Geist „Bestimmung des Einflusses von ÖPNV-Verkehrsnetzen auf die Erreichbarkeit“ [4].

[1] https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html

[2] https://www.graphhopper.com/de/

[3] https://www.vrs.de/fuer-unternehmen/open-data-service

[4] https://pretalx.com/fossgis2023/talk/AYWSVZ/

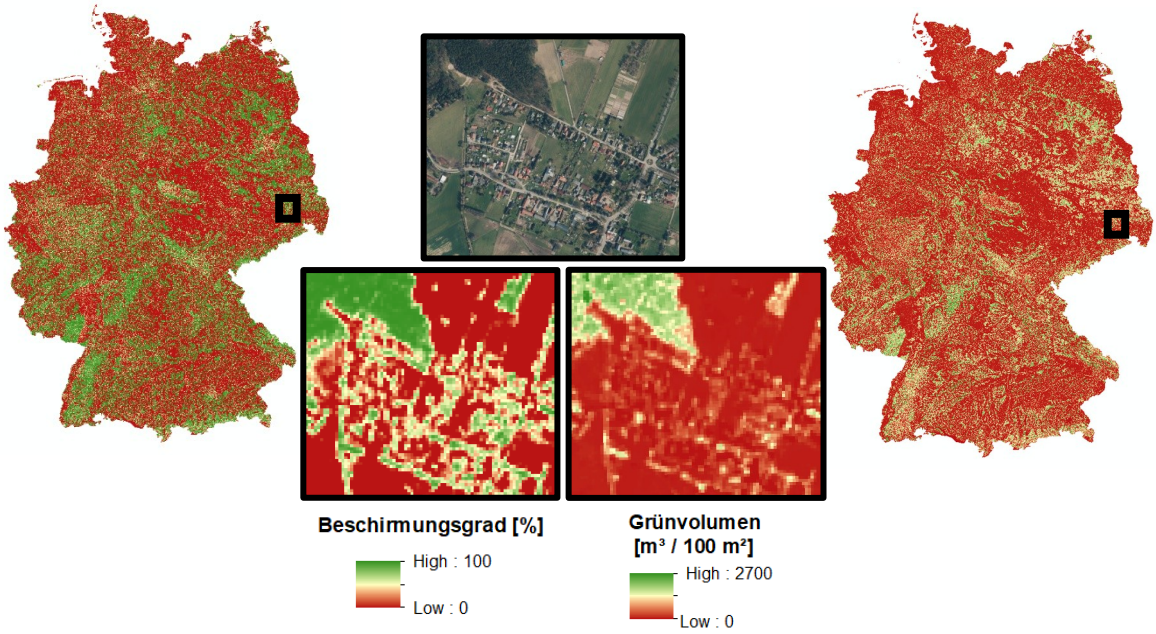
Klimaanpassungsrelevante Vegetationsindikatoren mit Hilfe von Sentinel-2 Zeitreihen

Benjamin Stöckigt[1] Dr. Annett Frick[1]

[1] Luftbild Umwelt Planung GmbH

Im Kontext des Klimawandels erkennen immer mehr Gemeinden ihre Verantwortung, Maßnahmen zur Anpassung an den Klimawandel in ihre Konzepte zur Stadtentwicklung zu integrieren. Vegetationsindikatoren wie das Grünvolumen und der Beschirmungsgrad können eine bedeutende Rolle bei der Reduzierung von bspw. städtischen Wärmeinseln spielen. Im Rahmen des UrbanGreenEye-Projekts entwickeln wir Fernerkundungsindikatoren basierend auf Satellitendaten, die ein regelmäßiges Monitoring von Gemeinden in ganz Deutschland ermöglichen. Sentinel-2 liefert dafür wöchentlich multispektrale Daten mit einer räumlichen Auflösung von bis zu 10 Metern und ist kostenlos verfügbar. Bei der Klassifizierung von Satellitenbildzeitreihen haben sich DeepLearning-Algorithmen, einschließlich Transformer-Modelle, als wirksam erwiesen. Häufig mangelt es KI-Modellen an der Fähigkeit, auf räumlich und zeitlich unterschiedliche Bedingungen übertragen zu werden. Deshalb wurden Ansätze zur Verbesserung der Generalisierungsfähigkeit wissenschaftlich untersucht und zur Herleitung der Indikatoren optimiert. Die Indikatoren ermöglichen eine Vielzahl von weiterführenden Anwendungen, auch außerhalb des kommunalen Kontexts.

Indikatoren werden als Rasterdatensätze unter CC BY-NC bereitgestellt und Algorithmen stehen unter General Public License auf GitHub zur Verfügung.



Kontakt zum Autor:

Benjamin Stöckigt (benjamin.stoeckigt@lup-umwelt.de)
Dr. Annett Frick (annett.frick@lup-umwelt.de)

SidescanTools - open source Sidescan Processing Software

Mia Schumacher, Finn Spitz

Wir stellen eine neue Open-Source-Software SidescanTools vor. Sie kann Edgetech .jsf Dateien sowie das plattformübergreifend lesbare '.xtf'-Dateiformat verarbeiten. SidescanTools erlaubt es, eine xtf-Datei zu lesen, Bodenerkennung, Schrägentfernungskorrektur sowie empirischer Verstärkungsnormalisierung (Emprirical Gain Normalisation, EGN) anzuwenden. Die Daten können dann als Geotiff/netCDF in jedes GIS exportiert werden. Das Programm ist in Arbeit und offen für Beiträge aus der community.

Seitensichtsonare sind weit verbreitete Instrumente zur Erstellung hochauflösender Bilder von Gewässerböden. Im Vergleich zu Fächerecholotsystemen, die eine geometrisch präzisere und dreidimensionale Karte liefern, sind Seitensichtsonare erschwinglicher und einfacher zu bedienen, weshalb sie z. B. von Seglern, Hobbytauchern oder NROs häufiger eingesetzt werden.

Allerdings gibt es, zumindest bis jetzt und nach unserem Wissensstand, nur begrenzte Möglichkeiten für freie oder kostengünstige Verarbeitung von Sonardaten. Tatsächlich gibt es nur ein oder zwei kommerzielle Programme, die eine vollständige Verarbeitung bieten und georeferenzierte Bilder aus den rohen Seitensichtsonar-Daten erzeugen können. Dieses Poster stellt die neue Open-Source-Software SidescanTools vor. Sie kann Edgetech .jsf Daten sowie das '.xtf'-Dateiformat verarbeiten - ein plattformübergreifendes Sonardatenformat, in das fast jedes Rohformat konvertiert werden kann. SidescanTools ermöglicht das Lesen einer xtf-Datei, die Anwendung von Bodenerkennung, Schrägentfernungskorrektur sowie empirischer Verstärkungsnormalisierung (Emprirical Gain Normalisation, EGN) und den Export als georeferenziertes Bild, das in jedes GIS importiert werden kann. Das Programm ist in Arbeit und offen für Beiträge aus der community.

SidescanTools wird im Rahmen von Ghostnetbusters entwickelt, einem Projekt zur Detektion von kleinen Objekten auf Sonardaten in Zusammenarbeit mit sonoware GmbH/Kiel und Geomar Helmholtz-Institut für Ozeanforschung, Kiel, gefördert von der WTSH Schleswig Holstein.

QFieldCloud Erweitern - Ideen und Praxisbeispiele

Michael Schmuki

In diesem Talk wird erläutert, wie QFieldCloud durch die Integration zusätzlicher Django-Apps erweitert werden kann. Dadurch lassen sich z.B. QField-Projekte generieren, auf Ereignisse aus der Feldarbeit reagieren, neue Webseiten und APIs hinzufügen und ganze QGIS-Modelle als QFieldCloud-Jobs ausführen.

Nach einer technischen Einführung werden verschiedene praxisnahe Beispiele betrachtet. Es wird erläutert, wie OpenStreetMap-Daten vollautomatisiert heruntergeladen werden können, um offlinefähige QField-Projekte zu erstellen. Lassen Sie sich inspirieren, wie mit OpenLayers ein eigenes WebGIS in QFieldCloud zum Leben erweckt wird. Zudem wird gezeigt, wie Fernerkundungsdaten heruntergeladen, in einer QGIS-Pipeline analysiert und die Ergebnisse in QField-Projekten verfügbar gemacht werden. Abschließend wird erörtert, wie diese Möglichkeiten optimal in Kombination mit QField-Plugins genutzt werden können.

Eine automatisierte FOSS-GDI zur Exploration von Erdsystem-Forschungsdaten

Peter Konopatzky

Ein Poster über den Aufbau und die Automatisierung einer containerisierten Geodateninfrastruktur in der Erdsystemforschung. Genutzt wird ein etablierter Stack aus PostGIS, GeoServer und GeoNetwork. Die gehosteten OGC-Dienste werden unter Nutzung einer Python-Bibliothek automatisiert aktuell gehalten.

Das öffentlich finanzierte Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, betreibt seit über einem Jahrzehnt eine Geodateninfrastruktur (GDI). Diese ist eingebunden in das Datenfluss-Framework "from Observations to Analysis and Archives" (O2A) [1], welches außerdem Metadatenmanagement für Sensoren ermöglicht, Near-Real-Time-Sensordaten via STA als OpenData bereitstellt, ein Repositorium für Datenpublikationen [2] anbietet und mit den Marine-Data- [3] und Earth-Data-Portalen [4] das Auffinden von Erdsystem-Forschungsdaten erleichtert.

Die von der GDI bereitgestellten OGC-Dienste werden durch einzelne Wissenschaftler:innen, in kuratierten Kartenviewern und internationalen Portalen wie dem Bathymetrie-Portal der IHO [5] genutzt. Öffentlich einsehbare Standard Operating Procedures [6] helfen Wissenschaftler:innen und Instituten dabei, ihre Forschungsdaten über diese GDI als OGC-standardkonform anzubieten und so den FAIR-Status ihrer Publikationen zu verbessern.

Unter den prominentesten Datenprodukten finden sich Tiefenprofile aus der Wassersäule, bathymetrische Visualisierungen, Daten zur arktischen Meereiskonzentration und Sammlungen von Unterwasserfotos-/videos.

Die Kernkomponenten der vorgestellten Geodateninfrastruktur – etablierte Open-Source-Technologien wie GeoServer und GeoNetwork – werden virtualisiert und containerisiert auf Linux-Systemen betrieben. Für die Automatisierung des Datenharvestings aus unterschiedlichen Quellen und des Managements der gehosteten OGC-Dienste ist die Python-Bibliothek "O2A Spatial" [7] entwickelt worden.

[1] https://o2a-data.de

[2] https://pangaea.de

[3] https://marine-data.de

[4] https://earth-data.de

[5] https://ncei.noaa.gov/maps/iho_dcdb

[6] https://hdl.handle.net/10013/927187e9-9866-4811-bc25-e7ea641d333e

[7] https://gitlab.awi.de/software-engineering/sdi/de.awi.sdi.o2a_spatial

An Inventory of Spatial Machine Learning Packages in R

Jakub Nowosad

Machine learning for spatial problems faces unique challenges, notably spatial dependence. Effective modeling requires integrating spatial information and proper validation methods to preserve spatial structure. This poster will overview spatial machine learning packages in R, focusing on tools for feature engineering, validation, and interpretation. It will also serve as a guide for comparing these tools and critically assessing their strengths and limitations.

Machine learning techniques are widely used for both spatial and non-spatial problems. As the use of machine learning models for creating maps continues to grow, several challenges have emerged that are unique to spatial problems and not encountered in non-spatial applications. Most importantly, spatial dependence poses significant challenges, requiring careful consideration to avoid overfitting and biased outcomes.

The effectiveness of machine learning models greatly hinges on the quality of input predictor data, underscoring the significance of feature engineering in the modeling process. Various approaches have been proposed to integrate spatial information into machine learning, including spatial coordinates of the observations, Euclidean distance fields (EDF), etc. Moreover, several modifications to traditional machine learning algorithms, such as RFGLS and hybrid models, have been suggested to improve the predictive performance of spatial data modeling. Regardless of the selected modeling algorithm, a proper validation approach is crucial for choosing suitable variables and hyperparameters during the tuning process and quantifying the quality of final predictions. Standard cross-validation (CV) methods tend to overestimate model transferability in spatial data, and spatial cross-validation methods, such as spatial block k-fold CV and nearest neighbor distance matching LOO CV, have been proposed to address this issue by preserving the spatial structure of data subsets. Finally, spatial data models require specific tools for model interpretation and visualization, such as area of applicability (AoA), to provide insights into the model's quality and decision-making process.

This poster aims to provide a comprehensive overview of spatial machine learning packages currently available in R, categorizing them based on their functionalities and applicability in spatial data analysis. We will not only focus on spatial extensions of three main machine learning frameworks in R (CAST for caret3, mlr3spatiotempcv for mlr3, and spatialsample for tidymodels) but also introduce specialized spatial machine learning packages for feature engineering, model validation, and interpretation, including SpatialML, RandomForestsGLS, and sits. The poster will not only serve as a navigational guide for comparing spatial machine-learning tools but also critically assess their strengths, limitations, and potential integration strategies.

OpenStreetMap-Daten zur Geolokalisierung unbemannter Luftfahrzeuge

Rebecca Schmidt

Das Poster zeigt, wie OpenStreetMap-Daten mithilfe von Open-Source-Tools für die Geolokalisierung unbemannter Luftfahrzeuge aufbereitet werden können und beleuchtet die Chancen und Herausforderungen, die mit ihrer Verwendung einhergehen.

Diese Arbeit untersucht das Potenzial von OpenStreetMap-Daten (OSM) zur Geolokalisierung unbemannter Luftfahrzeuge (UAVs). Die öffentlich zugänglichen Informationen von OSM eröffnen zahlreiche Möglichkeiten zur Weiterverarbeitung und Anwendung, insbesondere für die Automatisierung von UAVs. Allerdings bringt die Nutzung offener Datenquellen auch Herausforderungen mit sich. Aufgrund von Unterschieden zwischen den Informationen in der OSM-Datenbank und den vom UAV erfassten Daten ist eine geeignete Bewertung und Aufbereitung erforderlich. Das Poster zeigt, wie OSM-Daten mithilfe von Open-Source-Tools für die Geolokalisierung aufbereitet werden und beleuchtet die Chancen sowie die Herausforderungen, die mit der Nutzung verbunden sind.

(k)NNDM - Kreuzvalidierung im geographischen und Prädiktorraum

Jan Linnenbrink

In der räumlichen Modellierung werden oft Kreuzvalidierungsmethoden genutzt, um Modelle zu evaluieren und ihre Hyperparameter zu tunen. Es existiert eine Vielzahl an räumlichen Kreuzvalidierungsmethoden, welche jedoch nicht universell einsetzbar sind.

Mit diesem Poster werden wir verschiedene Ansätze systematisch vergleichen, sowie eine neue Kreuzvalidierungsmethode und ihre Implementierung in R vorstellen, die sowohl im Prädiktorraum, als auch dem geographischen Raum arbeitet.

Die raum-zeitliche Vorhersage von an Messpunkten gemessenen Umweltvariablen ist eine wichtige Methode in den Geowissenschaften. Hierbei werden zunehmend maschinelle Lernverfahren genutzt, welche mithilfe von Kreuzvalidierungsmethoden evaluiert und getuned werden. Die traditionelle Kreuzvalidierung, bei der Trainingsdaten zufällig in Gruppen eingeteilt werden, ist bei räumlich strukturierten Daten jedoch oft nicht angemessen. Alternativ existiert eine Vielzahl an räumlichen Kreuzvalidierungsmethoden, welche jedoch nicht universell einsetzbar sind.

Hier stellen wir (k)NNDM vor, eine neue, in R implementierte Kreuzvalidierungsmethode, die sowohl im Prädiktorraum, als auch dem geographischen Raum funktioniert und unabhängig vom Stichprobendesign eingesetzt werden kann.

Wir vergleichen die Genauigkeit der neuen Kreuzvalidierungsmethode unter verschiedenen Szenarien, und diskutieren die Eignung von Kreuzvalidierungen im Prädiktorraum im Vergleich zu Kreuzvalidierungen im geographischen Raum .

UmweltNAVI: Eine interaktive und kooperative Plattform für Umweltinformationen

Dirk Weber, Torben Fiedler, Martin Dittmann

Nds. Ministerium für Umwelt, Energie und Klimaschutz; wemove digital solutions GmbH

Die Beachtung dieser Hinweise vereinfacht uns die Erstellung des Tagungsbandes, erspart uns langwierige Formatierungsarbeiten und erleichtert die Zusammenstellung der Beiträge in der richtigen Reihenfolge. Da wir die Zusammenstellung des Tagungsbandes nicht nach Zeit bezahlt bekommen, sondern ihn ehrenamtlich in unserer Freizeit erstellen, sind wir Ihnen mehr als dankbar!

Die digitale Transformation der öffentlichen Verwaltung bleibt eine zentrale Herausforderung, der sich Bund, Länder und Kommunen in den kommenden Jahren weiterhin stellen müssen. In einer zunehmend vernetzten Gesellschaft erwarten Bürgerinnen und Bürger nicht nur Transparenz über Datenquellen und -verarbeitung, sondern auch eine einfache und direkte Möglichkeit, sich über Umweltthemen zu informieren und aktiv an deren Gestaltung teilzuhaben. Umweltdaten sind häufig nur dezentral verfügbar, und die Recherche erfordert teilweise erhebliche Aufwände. Zudem erschweren unterschiedliche Standards und technische Systeme für Abruf und Visualisierung den Zugang zu diesen Daten. Angestrebt wurde daher die Bereitstellung eines benutzerfreundlichen Zugriffs auf Umweltinformationen aus Niedersachsen und Deutschland sowie die Entwicklung nachhaltiger Prozesse und technischer Lösungen, die sich an konkreten Anwendungsfällen der Bürgerinnen und Bürger orientieren.



Abb. 1: UmweltNAVI

Die niedersächsische Landesregierung initiierte 2021 im Rahmen des „Masterplan Digitalisierung“ die Entwicklung des „Bürgerinformationsportals Umwelt Niedersachsen“, zu dem auch die Konzeption einer App zur Bereitstellung von Umweltinformationen gehörte[1]. Im März 2023 ging die App UmweltNAVI Niedersachsen planmäßig in den regulären Betrieb. Die Entwicklung erfolgte aufgrund des innovativen Charakters in einem agilen Prozess. Die UmweltNAVI-App[2] vereint frei verfügbare Umweltdaten und -informationen in einer modernen, auf Open-Source-Software basierenden Smartphone-Anwendung.

Die App nutzt unter anderem Standortdaten als Anknüpfungspunkt und zeigt regionale Umweltdaten und -informationen auf einer Karte an. Im Hintergrund arbeitet eine weitgehend automatisierte Datenaufbereitung, die die Inhalte regelmäßig objektbezogen in eine nachnutzbare Datenbasis überführt. Die Datenbasis aus amtlichen Umweltdaten wird automatisiert mit Open Data aus bekannten Quellen wie Wikipedia, Observation.org oder dem „Niedersachsen Hub“ des Tourismusnetzwerks Niedersachsen angereichert. Durch die Kombination mit anderen Anwendungen wird die Datengrundlage u. a. im Natur- und Artenschutz verbessert.

Eine Verbindung zu bestehenden erfolgreichen Angeboten schafft eine breitere öffentliche Wissensbasis, die auch als Grundlage für politische Entscheidungen auf allen Ebenen dienen kann. Ziel ist es, die vorhandenen Informationen nicht nur handhabbar zusammenzuführen und aufzubereiten, sondern sie den Nutzerinnen und Nutzern auch bürgernah und transparent zur Nachnutzung zur Verfügung zu stellen. Interaktionen mit den Nutzenden werden durch Ansätze der „Gamification“ und Teilhabe gefördert. Beispielsweise existieren ein Umwelt-Quiz, eine „Meine Orte“-Funktion und die Möglichkeit des Foto-Uploads. Die App zeigt außerdem aktuelle Veranstaltungen und (Presse-)Mitteilungen zu Umweltthemen in der Umgebung an und liefert dazu Hintergrundinformationen. Hierzu können Daten wie

Mitteilungen aus RSS-Feeds, Objekte aus REST-APIs oder Messdatenreihen aus SensorThingsAPI-Schnittstellen flexibel angebunden werden.

Im Bereich „Wissen“ gibt es kurze Zusammenfassungen und weiterführende Links zu verschiedenen Themen. Eine „Wo finde ich“-Funktion ermöglicht das Entdecken der nächstgelegenen „Umweltobjekte“, beispielsweise Naturschutzgebiete oder Messstationen, basierend auf dem gewählten Thema. Die Nutzung der App ist auch offline möglich. Inhalte werden flexibel in einem Content-Management-System, das auf dem Headless CMS Directus basiert (ehemals FOSS), gepflegt. Die Verarbeitung von Daten aus verschiedenen Datenquellen erfolgt über ein ETL-Framework auf Basis von Apache Airflow. Als Datenquellen kommen unterschiedliche Schnittstellen und Formate (u. a. WMS, WFS, SensorThings API, Shape-Dateien, Wikipedia) zum Einsatz. Ergänzend werden Kartendaten über QGIS themenbezogen aufbereitet und über einen QGIS-Server als WMS bereitgestellt. Langsame Kartendienste werden mit einem Mapproxy gecached. Das App-Framework Flutter ermöglicht die Bereitstellung einer modernen, plattformübergreifenden Anwendung. Eine Nachnutzung der angereicherten Daten über eine öffentliche Schnittstelle wie die OGC-API Features wird derzeit konzipiert.

Die App UmweltNAVI ist für Niedersachsen und Schleswig-Holstein in App- und Playstore (sowie Aurora) öffentlich verfügbar. Ein offenes Repository zur Nachnutzung und Weiterentwicklung der Anwendung steht auf Open-Code[3] zur Verfügung.

Das VKoopUIS-Projekt Nr. 67 „UmweltNAVI“ bildet seit dem 01.01.2025 den Rahmen für die technische Weiterentwicklung der UmweltNAVI-App, die in Kooperation der Bundesländer Niedersachsen, Schleswig-Holstein und Rheinland-Pfalz erfolgt. Ein gemeinsam betriebenes Backend ermöglicht sowohl die Erweiterung der Datenbasis als auch die Aufbereitung und Darstellung bundesweiter Umweltdaten unter Berücksichtigung regionaler Besonderheiten[4].

Kontakt zum Autor:

Dirk Weber
Nds. Ministerium für Umwelt, Energie und Klimaschutz
Archivstr. 2, 30169 Hannover
Referat 14 - Umweltinformation, Digitalisierung, eGovernment
Dirk.weber@mu.niedersachsen.de

Torben Fiedler
Nds. Ministerium für Umwelt, Energie und Klimaschutz
Archivstr. 2, 30169 Hannover
Referat 14 - Umweltinformation, Digitalisierung, eGovernment
Torben.fiedler@mu.niedersachsen.de

Martin Dittmann
wemove digital solutions GmbH
Hanauer Landstr. 52, 60314 Frankfurt
martin.dittmann@wemove.com

Literatur

[1] https://www.umwelt.niedersachsen.de/startseite/aktuelles/pressemitteilungen/pi-95-umweltnavi-213668.html

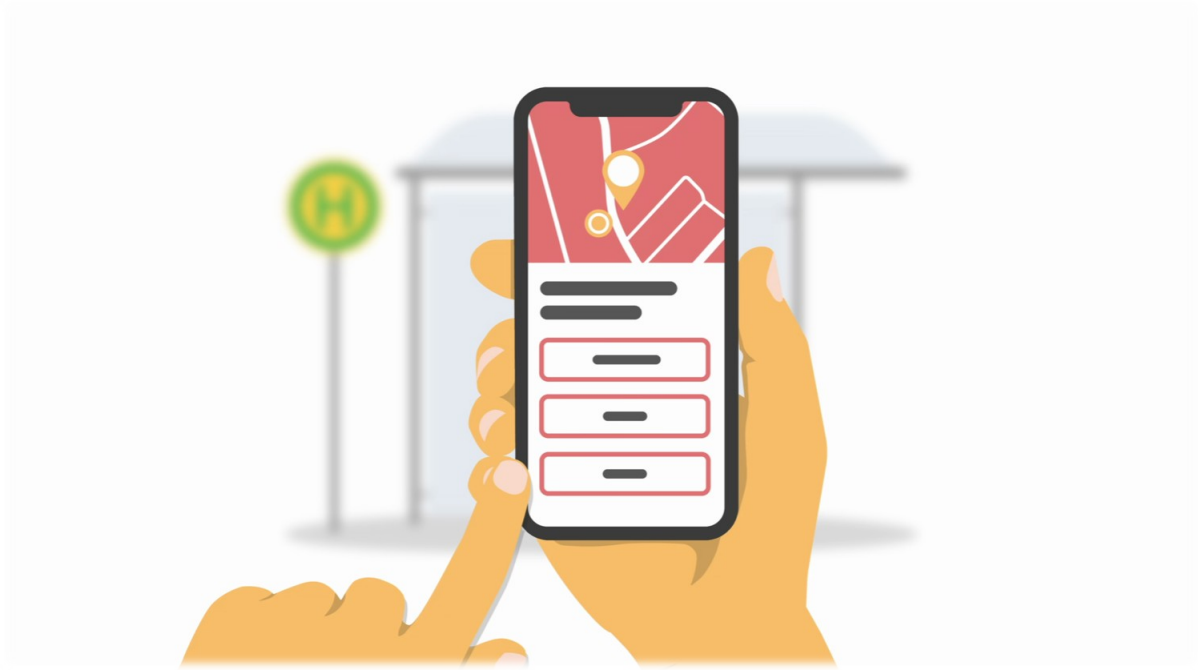
[2] https://umwelt-navi.info/

[3] https://opencode.de/umweltnavi

[4] https://www.blag-udig.de/KoopUIS.html

Offene Daten für barrierefreie Mobilität und inklusive Reiseketten im ÖPV

René Apitzsch



Das Poster thematisiert die Rolle offener Daten bei der Förderung barrierefreier Mobilität im öffentlichen Personenverkehr (ÖPV). Es zeigt auf, wie durch die Erfassung und Nutzung standardisierter Daten Barrieren durch Infrastruktur und Fahrzeuge entlang von Reiseketten verringert werden können. Dazu werden Ansätze zur strukturierten Erfassung von Hindernissen im Verkehrswesen sowie Möglichkeiten zur Integration dieser Daten in bestehende Systeme dargestellt.

Barrierefreiheit im öffentlichen Nahverkehr ist ein wesentlicher Faktor für die gesellschaftliche Teilhabe. Trotz Fortschritten bei der Ausstattung von Haltestellen und Fahrzeugen gibt es weiterhin Lücken in der Verfügbarkeit und Qualität der Informationen zur Barrierefreiheit. Eine verlässliche, durchgängige Reisekette bzw. Informationen darüber bleibt für viele mobilitätseingeschränkte Personen eine Herausforderung. Das Poster zeigt, wie offene Daten als Lösung dienen können, um diese Barrieren abzubauen und den ÖPV zugänglicher zu gestalten.

Im Mittelpunkt steht die Entwicklung bzw. Nutzung eines offenen Datenmodells, das Eigenschaften von ÖPV-Fahrzeugen standardisiert erfassen und für Auskunftssysteme nutzbar machen soll. Dabei wird beleuchtet, wie eine zentrale, öffentlich zugängliche Datenplattform dazu beitragen kann, Barrieren transparent zu machen und die Informationen in Navigations- und Planungssysteme (der Verkehrsverbünde) zu integrieren.

Zudem wird gezeigt, wie Bürger durch die Crowdsourcing-Initiative bzw. Open-Source-App OpenStop aktiv in die Datenerhebung eingebunden werden können, um eine umfassendere und aktuelle Datenlage zur Barrierefreiheit von Haltestellen zu schaffen. Dieser partizipative Ansatz verbessert nicht nur die Datenqualität, sondern fördern auch das Bewusstsein für die Bedürfnisse von Menschen mit Mobilitätseinschränkungen.

Es wird außerdem verdeutlicht, welche Herausforderungen bei der Erfassung und Nutzung von Barrierefreiheitsdaten und Fahrzeug-Eigenschaften bestehen und wie diese überwunden werden können. Das Poster bietet somit einen Überblick über den aktuellen Stand der Technik und gibt einen Ausblick auf zukünftige Entwicklungen im Bereich der barrierefreien Mobilität durch offene Daten.

Einsatz für die Energiewende

Jens Tielker, Johannes Quente

In diesem Vortrag geht es um den Aufbau und Betrieb einer Geodaten-Infrastruktur, die in diversen Projekten zur Planung der Energiewende eingesetzt wird. Dazu verwenden wir hauptsächlich Open-Source-Software wie QGIS, PostgreSQL, PostGIS, QField, R, R-Studio und Shiny Server.

Zur Bearbeitung verschiedener Projekte für Linienbauwerke der Energiewende haben wir ein Netzwerk von Open Source Technologien eingesetzt. Als Komponenten werden unter anderem QGIS als Frontend Anwendung verwendet. Im Backend haben wir eine PostgreSQL Datenbank mit PostGIS aufgebaut, in der sowohl die Geodaten als auch die QGIS-Projekte, zur parallelen Bearbeitung gepflegt werden. Auch haben wir diverse Web-Applikationen über R-Studio geschrieben und über einen firmeninternen Shiny-Server gehostet. Dieser Vortag behandelt die Erfahrungen, die wir bei dem Betrieb und dem Aufbau der Workflows gemacht haben. In Situ arbeiten wir auf mobilen Endgeräten mit QField, sowohl bei der Ausführung von Felderkundungen als auch der Überwachung von Bauleistungen.

Einbindung von OGC APIs in Webtools für Wärmestromdaten

Nikolas Ott

Für eine kompakte, clientseitige Webanwendung werden externe Daten im Weltwärmestrom Datenbankprojekt über OGC APIs bereitgestellt. Auf Nachfrage können benötigte Werte an den Client übermittelt werden. Die entwickelte Architektur wird am Beispiel des „Digitalen Bohrlochs“ veranschaulicht. Dabei wird die Temperatur an einer beliebigen Position in Abhängigkeit von der Tiefe mithilfe der Bootstrapping-Methode berechnet. Als Eingabeparameter für die Berechnung dienen der Wärmestromwert (q) an der Oberseite und am Boden der jeweiligen Schicht, die Mächtigkeit der Erdkrustenschichten sowie die Leitfähigkeit der Schichten. Bis auf die q-Werte, die aus der Datenbank selbst stammen, werden alle übrigen Daten aus globalen Rasterdatensätzen gewonnen. Diese Daten beim initialen Laden an jedes Endgerät mitzuschicken, wäre ineffizient und würde einen unnötigen Datentransfer verursachen. Daher werden die globalen Rasterdaten als OGC-API-Coverages angeboten, und lediglich der Rasterwert, der die zu berechnende Position enthält, wird bei Bedarf übermittelt.

Offene Daten für barrierefreie Mobilität und inklusive Reiseketten im ÖPV

René Apitzsch, Robin Thomas, Ulrich Heinkel

Zusammenfassung

Der Beitrag und das dazugehörige Poster thematisieren die Rolle offener Daten bei der Förderung barrierefreier Mobilität im öffentlichen Personenverkehr (ÖPV). Es zeigt Ansätze auf zur standardisierten Erfassung von Barrieren hinsichtlich Infrastruktur und Fahrzeugen sowie zur Integration dieser Daten in bestehende Auskunftssysteme.

Motivation

Barrierefreiheit im öffentlichen Personenverkehr spielt eine wichtige Rolle für die gesellschaftliche Teilhabe. Trotz Fortschritten beim Ausbau barrierefreier Haltestellen und der Ausstattung von Fahrzeugen gibt es weiterhin Lücken in der Verfügbarkeit und Qualität barrierefreiheitsrelevanter Daten. Dies gilt sowohl für Informationen über Infrastruktur als auch über Fahrzeuge, denn nur in Kombination können zuverlässige Auskünfte über durchgängig barrierefreie Reiseketten gewährleistet werden. Von den rund 240.000 Haltestellen in Deutschland [1] lagen den Verkehrsunternehmen (VU) noch im Jahr 2022 von einem Großteil keine Informationen zur im Personenbeförderungsgesetz geforderten Barrierefreiheit vor [2]. Auch über ÖPV-Fahrzeuge liegen diesbezüglich keine detaillierten offenen Informationen vor, die über die Angaben von Niederflurfahrzeugen in der VDV-Statistik [3] hinausgehen. Standardisierte, offene Daten können hier einen entscheidenden Beitrag für einen inklusiven öffentlichen Personenverkehr leisten.

Ziel

Unser Ziel ist es mittels einer eigens entwickelten App auf Basis von OpenStreetMap (OSM), Informationen über Eigenschaften und Barrieren von ÖPV-Haltestellen zu sammeln, bereitzustellen und diese in die Auskunftsysteme der Verkehrsverbünde zu integrieren. Darüber hinaus wollen wir im Rahmen einer Machbarkeitsstudie ein standardisiertes und offenes Datenmodell sowie eine Datenbank für ÖPV-Fahrzeuge konzeptionieren, um eine vollständig (barrierefreie) Reisekette auf Basis offener Daten abbilden zu können.

Datenfluss

Zentrales Element des in Abb. 1 dargestellten Datenflusses bildet OpenStreetMap. Hier werden die durch (die Erfassungs-App) erhobenen Informationen zur ÖPV-Infrastruktur standardisiert gespeichert. Über einen Export werden die Daten in die VU-Datenbanken integriert, so dass sie in der Reiseauskunft berücksichtigt werden können.

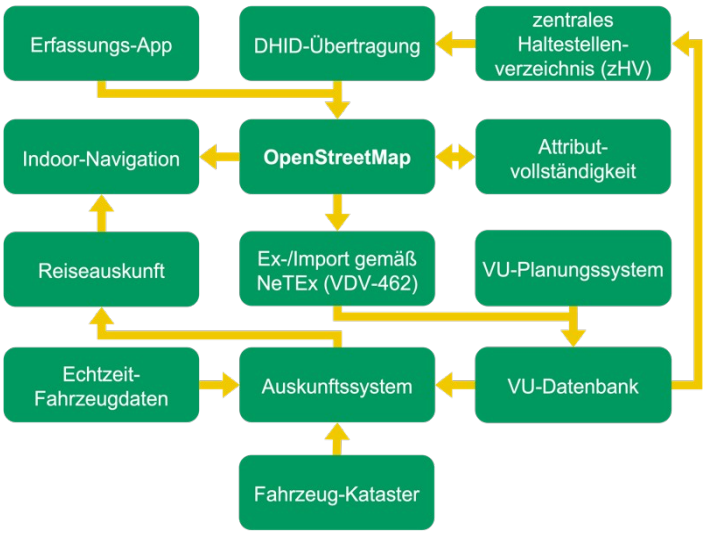


Abb. 1: Datenfluss

Standardisierung und Erfassung von ÖPV-Infrastruktur

Die Hauptquelle bei der Erarbeitung zu erfassender Attribute war [4]. Die dort rein textuell beschriebenen Attribute wurden in OSM-Schlüssel-Wert-Paare überführt. Falls ein entsprechendes Attribut mit vorhandenen OSM-Schlüssel-Wert-Paaren nicht ausreichend abgebildet werden konnte, wurden neue erarbeitet und im Rahmen des OSM-Proposal-Prozesses standardisiert. Die im Zuge des Projektes neu standardisierten OSM-Tags sind: Plattformlift, portable Rampe, Sprachausgabe an der Plattform, Anfahrtshilfe bei Bordsteinen, die Art der Erfassung von Fahrstuhl-Dimensionen und der Öffnungszustand von Türen.

Um die Last der Datenerhebung von den VU auf einer breiten Öffentlichkeit zu verteilen, zur effizienteren und bedarfsgerechteren Erfassung, wurde die App OpenStop entwickelt. Die mehrsprachige Cross-Plattform-App ermöglicht die Erfassung von ÖPV-Infrastruktur durch das Beantworten einfacher Fragen. Vorgegebene Antwortmöglichkeiten, Illustrationen und Erläuterungen sollen die Beantwortung erleichtern und Fehleingaben vermeiden.

Standardisierung und Erfassung von Fahrzeugen

Da in OSM nur georeferenzierte Objekte erfasst werden können, muss für die Erfassung und Bereitstellung von Fahrzeug-Attributen in Abstimmung mit der ÖPV-Branche eine eigene Datenbank aufgebaut werden. Das zugrunde liegende Datenmodell ist mit Transmodel bzw. NeTEx bereits definiert und wird im Rahmen des Projektes um weitere Fahrzeug-Attribute ergänzt. Analog zur Infrastruktur ist [4] hierbei die primäre Quelle.

Wer die Fahrzeug-Daten erfassen sollen, befindet sich noch in der Abstimmung. Angedacht ist, dass die Eintragung in das Fahrzeug-Kataster durch das jeweilige VU bzw. den Hersteller neu erworbener Fahrzeuge erfolgt. Dieser Prozess ist mit einem Mehraufwand verbunden. Um daher die Bereitwilligkeit der VU zu erhöhen, soll die Erfassung in verschiedenen Detailgraden möglich sein, wobei jeder Detailgrad neue Anwendungsfälle eröffnet.

Überführung der Haltestellen-Informationen in Auskunftssysteme

Zur Integration der erhobenen Daten in die Auskunftssysteme wurde eine Exporter entwickelt, der OSM-Daten in ein VDV-462-konformes XML-Format konvertiert. Dieses Format kann in die VU-Datenbank eingelesen werden, womit die Haltestellen-Informationen in der Reiseauskunft zur Verfügung stehen. Zur Referenzierung zwischen Haltestellen in der VU-Datenbank und OSM wird auf die Deut-

sche Haltestellen-ID (DHID) aus dem zentralen Haltestellenverzeichnis (zHV) zurückgegriffen. Da diese jedoch nicht für alle Haltestellen in OSM hinterlegt ist, wurde eine Algorithmus zum Abgleich von OSM und zHV entwickelt, der Grundlage für eine entsprechende MapRoulette-Challenge ist.

Ergebnisse und Ausblick

Die Nutzung von in OSM erfassten ÖPV-Infrastruktur-Daten in der Reiseauskunfts-App eines VU wurde am Magdeburger Hauptbahnhof in einem Pilotbetrieb erfolgreich getestet. Zukünftig soll ein prototypischen Fahrzeug-Kataster umgesetzt und die Daten ebenfalls in der Beauskunftung berücksichtigt werden.

Förderhinweis

Die Projekte OPENER next und GAPLESS wurden bzw. werden im Rahmen der Innovationsinitiative mFUND durch das Bundesministerium für Digitales und Verkehr gefördert.“

Gefördert durch:

Bundesministerium für Digitales und Verkehr

mFUND

aufgrund eines Beschlusses des Deutschen Bundestages

Kontakt zum Autor:

René Apitzsch, Robin Thomas
Technische Universität Chemnitz
Professur Schaltkreis- und Systementwurf
Reichenhainer Straße 70
+49 (0)371 531-33845
{rene.apitzsch, robin.thomas}@etit.tu-chemnitz.de

Literatur

[1] Zentrales Haltestellenverzeichnis (zHV). url: https://zhv.wvigmbh.de

[2] Wissenschaftliche Dienste des Deutschen Bundestages. Barrierefreiheit von Bahnhöfen und Bushaltestellen in Deutschland. Sachstand WD 5 - 3000 - 105/22. Sep. 2022.

[3] Verband Deutscher Verkehrsunternehmen (VDV). VDV-Statistik 2023. Köln: VDV, 2023.

[4] Verein zur Förderung einer durchgängigen elektronischen Fahrgastinformation (DELFI) e.V. Handbuch - Barrierefreie Reiseketten in der Fahrgastinformation (2. Aufl.). Frankfurt am Main: DELFI e.V., 2024.

Kinder, Karten, Open Source

Dr. Roland Olbricht

Wie begeistern wir die nächste Generation für Open Source und OpenStreetMap?

Der Vortrag soll eine Bestandsaufnahme bieten, was es an Bildungs- und Begeisterungsmaterial und -konzepten für verschiedene Altersgruppen unter 18 zum Raum Erfahren schon gibt. Es ist Zeit zu überlegen, welche Angebote wir machen können und wollen, um daran mit Open Source und OpenStreetMap anzuknüpfen.

Einerseits hat schon Jochen die Idee zur Mitgliederversammlung gebracht, am Maus-Türöffnertag die Themen kindertauglich zu präsentieren. Und in der Tat haben schon Vierjährige mitunter eine gute Orientierung im Raum, selbst vor dem Lesen lernen. Andererseits hat OpenStreetMap Minderjährige im Hinblick auf die umfangreichen Pflichten der DSVGO erst einmal abgewiesen.

Der Vortrag soll eine Bestandsaufnahme bieten, was es an Bildungs- und Begeisterungsmaterial und -konzepten für verschiedene Altersgruppen unter 18 zum Raum Erfahren und OpenSource schon gibt. Es ist Zeit zu überlegen, welche Angebote wir machen können und wollen, um daran mit Open Source und OpenStreetMap anzuknüpfen.

Fahrzeugortung DB Regio – mehr als GPS

Stefan Kowski

Zusammenfassung: Der Praxisvortrag gibt einen Einblick in die Fahrzeugortung von DB Regio und zeigt, welche Open Source-Software in verschiedenen Anwendungsbereichen zum Einsatz kommt.

Schlüsselwörter: Ortung, GPS

Die Fahrzeugortung DB Regio (FzO-BO) hat die Aufgabe, Zug- und Fahrzeuginformationen entgegenzunehmen, in Echtzeit zu konsolidieren und an Abnehmer auszuliefern. Verarbeitet werden u. a. GPS-Positionsdaten, bestimmte Ereignisse (z. B. Türöffnung/Boarding, Ankunft und Abfahrt) und die Wagenreihung des Zuges (z. B. für Anzeigen am Bahnsteig). Die Daten werden in ein einheitliches Format konvertiert und nachgelagerten Diensten zur Verfügung gestellt. FzO-BO DB Regio bildet damit eine zentrale Drehscheibe zwischen einer Positionsmeldung aus dem Fahrzeug und den nutzenden Verfahren wie Störfallmanagement, Reisendeninformation (RIS) oder Fahrzeugeinsatzplanung.

Situation und Anforderungen

Die Positionsbestimmung im Eisenbahnverkehr erfolgt über die Infrastruktur, d.h. im Gleisumfeld verbaute elektromagnetische Geräte und Signale. Diese Einrichtungen erzeugen u. a. Meldungen bei Durchfahrten, Ankünften und Abfahrten in einzelnen Gleisabschnitten. Die Meldung enthält als Position den Schlüssel einer sog. Betriebsstelle und die Zugnummer.

Infrastrukturmeldungen haben verschiedene Nachteile:

- da es keine mit einer Geoposition vergleichbare Positionsangabe gibt, ist die exakte Position innerhalb eines u. U. mehrere 100 m langen Gleisabschnitts unbekannt,
- Meldungen werden nur selten erzeugt, da sie nur bei bestimmten Ereignissen generiert werden (Durchfahrten etc.) und nicht periodisch, z. B. bei Stillstand,
- Meldungen werden u. U. erst zeitverzögert erstellt (mehrere Minuten möglich), so dass sie nicht echtzeitfähig sind,
- auf Strecken anderer Eisenbahninfrastrukturunternehmen und in Abstellbereichen gibt es die technische Infrastruktur nicht, so dass dort gar keine Meldungen zur Verfügung stehen,
- die Meldungen enthalten nur Zugnummern, nicht die Fahrzeuge im Zugverband.

Im Regionalverkehr enthalten Ausschreibungen und Verkehrsverträge der Bundesländer bestimmte Bedingungen, die von den Vertragspartnern/Eisenbahnverkehrsunternehmen erfüllt werden müssen. Da neuere Verträge eine Echtzeitlieferung von Ortungspositionen der bestellten Verkehre vorsahen, wurden eine Reihe von Anforderungen für ein Ortungssystem gesammelt:

- das System muss echtzeitfähig sein, d. h. die Latenz zwischen der Erfassung der Position im Fahrzeug und der Auslieferung an Abnehmer muss minimal sein,
- es muss ein zentrales Ortungssystem für alle Verkehre bei DB Regio geben,
- die Echtzeitpositionen müssen an verschiedene Abnehmer geliefert werden,
- das System muss auch Daten für andere Systeme bereitstellen können, z. B. für die Disposition (Fahrzeugeinsatzplanung) oder Reisendeninformation,

- es müssen auch Positionsmeldungen in Betriebshöfen (Abstellung) erzeugt werden, z. B. für Rangierfahrten,
- neben der Zugnummer muss eine Positionsmeldung auch die vollständige Wagenreihung enthalten.

Für das Projekt „Fahrzeugortung DB Regio (FzO-BO)“ ergaben sich daraus folgende Aufgaben:

- GPS-Positionen ermitteln, konsolidieren und an beliebige Abnehmer liefern,
- Fahrzeugnummern zu Infrastrukturmeldungen bereitstellen (u. a. für exakte Laufleistungsermittlung),
- Fahrzeugbewegungen in der Abstellung verfolgen und auswerten,
- Nicht-GPS-fähige Fahrzeuge in der Abstellung orten (z. B. per RFID oder Bluetooth),
- Messaging-Architektur für Echtzeitfähigkeit.

FzO-BO – Die Anwendung

Die Fahrzeugortung DB Regio (FzO-BO) ist das zentrale Ortungssystem für Schienen- und SEV-Busverkehre von DB Regio. Ortungsmeldungen kommen direkt aus Fahrzeugen oder von anderen Backoffice-Ortungssystemen (z. B. von Fahrzeugherstellern, die Meldungen ihrer Fahrzeuge sammeln und gebündelt an FzO-BO senden).

FzO-BO hat ca. 30 interne und externe Abnehmer, die Ortungsmeldungen erhalten. Die Ortungsmeldungen haben ein standardisiertes Format: u. a. Geoposition, Zugnummer und Fahrplaninformation, Wagenreihung und -ausrichtung, Energieverbrauch, Zugfahrtinformationen (z. B. Kundenfahrt oder Bereitstellung). Das Format wird regelmäßig um neue fachliche Daten ergänzt, die sich aus neuen Anforderungen der Abnehmerprozesse ergeben.

Als technische Plattform wird Red Hat JBoss EAP auf Red Hat Enterprise Linux (RHEL) verwendet.

FzO-BO ist als messaging-basierte Microservice-Architektur implementiert (ca. 20 Microservices). Je nach Anforderungen werden neue Services umgesetzt und integriert.

Das eingesetzte Messaging-System ist die Open Source-Software Eclipse Amlen (ehem. IBM MessageSight).

Als Programmiersprachen werden Java im Backend und Javascript im Frontend verwendet.

Kennzahlen der Fahrzeugortung:

- Leistung (24 h): ca. 27 Mio. eingehende Nachrichten, ca. 202 Mio. ausgelieferte Nachrichten (310 bzw. 2330 pro Sekunde),
- Sendeintervall: alle 5 - 30 Sekunden eine Ortungsmeldung aus einem Fahrzeug,
- Meldungslatenz Fahrzeug ✉ Abnehmer: 500 ms - 3 s, davon 7,3 ms in FzO-BO,
- Ca. 3300 sendende Fahrzeuge,
- Ca. 1000 Nutzer der fünf Web-Anwendungen im DB-Konzern.

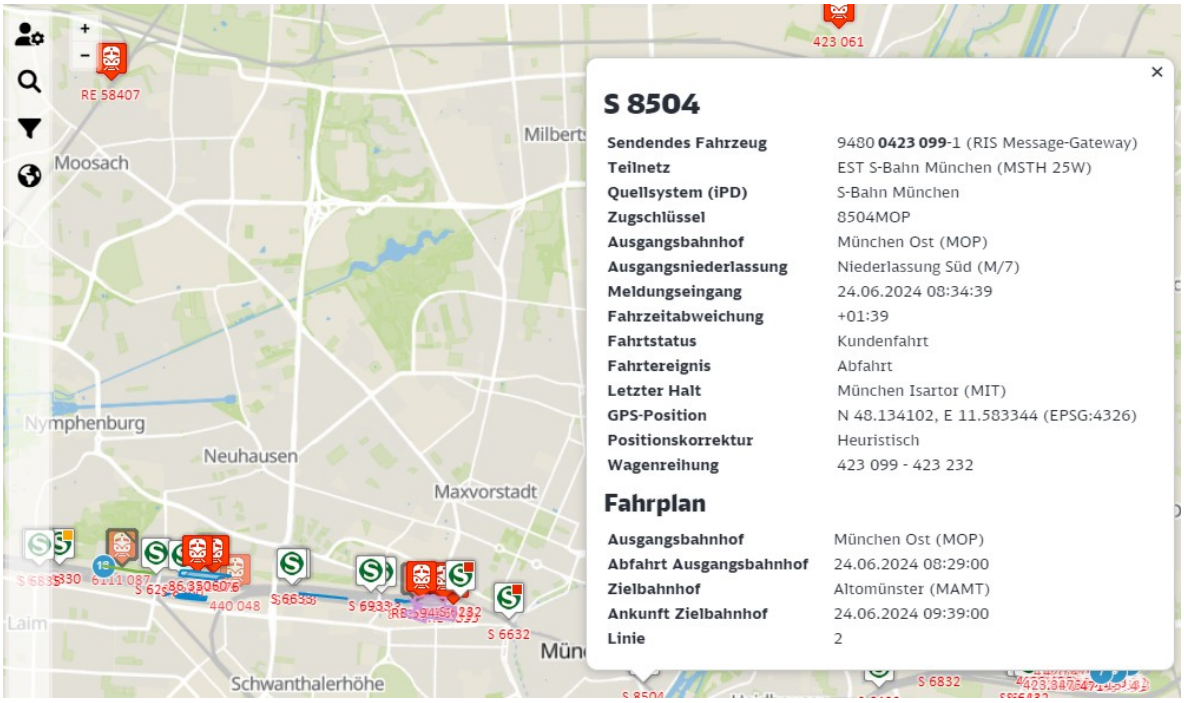


Abbildung 9: Visualisierung der Fahrzeugortung DB Regio (FzO-BO)

Open Source-Komponenten

Open Source-Komponenten werden in verschiedenen Bereichen des FzO-BO eingesetzt: Datenhaltung, Datenpflege und Softwareentwicklung.

Wichtige Kriterien für die Auswahl der Komponenten waren eine weite Verbreitung, Ausgereiftheit und gute Unterstützung durch die Community.

Im Bereich Datenhaltung gab es folgende Anforderungen:

- Für wichtige Sachobjekte gab es gar keine oder keine direkt verwendbaren Daten (wenn sie z. B. nur für Druckprodukte genutzt werden), so dass eine Datenerfassung geplant werden musste,
- Datenpflege muss regional erfolgen (wegen benötigter Ortskenntnisse oder fachlicher Zuständigkeit), daher musste die Datenhaltung zentral und mehrbenutzerfähig sein,
- Verschiedene Sachobjekte lagen nur in Dateiform vor (vor allem Daten von Lieferanten), was einzelne Änderungen und die Versionierung problematisch machte.

Das Projekt hat die Datenbank Postgres mit der PostGIS-Erweiterung ausgewählt, um seine Geodaten zu verwalten. Die damit erstellte FzO-BO Geodatenbank enthält Tabellen für Gleis- und Streckennetze, Standplätze und Betriebsstellen, RFID-Leserpositionen u. a. Sie stellt verschiedene Import- und Exportmöglichkeiten zur Verfügung.

Da das Projekt eine Datenpflege zur Erfassung von Sachobjekten benötige, wurden Anforderungen definiert, die der Pflegeprozess und die dort verwendeten Werkzeuge erfüllen müssen:

- Werkzeuge zum Editieren von Vektorgeometrien (Punkte, Linienzüge, Polygone),
- Darstellung von Karten zur Orientierung bei der Datenpflege,

- Einfache Pflege der Attribute von Sachobjekten,
- Unterstützung vieler Datenformate, Datenbanken und OGC-Standards,
- Flexible Koordinatentransformationen.

Die Anwendung QGIS erfüllte diese Anforderungen. Die Datenpflege wird durch regionale Mitarbeiter der Fachbereiche bei DB Regio durchgeführt, die durch das Projektteam FzO-BO geschult werden.

Für projektinterne Analysen und Fehlersuche wurde eine Visualisierung konzipiert. Sie sollte als Web-Anwendung umgesetzt werden, verschiedene Datenquellen einbinden, sehr performant sein auch bei großen Datenmengen sowie flexible Filter- und Auswahlmöglichkeiten bieten.

Bei der Implementierung verwendete das Projekt Open Layers als Bibliothek für die Darstellung der Geodaten im Web-Browser. DB-interne Tile Server stellen Hintergrundkarten bereit. Die Anwendung greift auf verschiedene Services zu, die Sachobjekte als Vektordaten bereitstellen, und stellt sie als konfigurierbare Layer dar. Die Darstellung von Fahrzeugen erfolgt über ein dynamisches Layer, das ständig aus Ortungsmeldungen aktualisiert wird.

Mittlerweile wird die Anwendung von mehr als 1000 Mitarbeitern im DB Konzern verwendet.

Für einige wichtige Sachobjekte gibt es keine Geodaten. Die benötigten Geometrien können aber aus anderen, vorhandenen Daten algorithmisch abgeleitet werden.

FzO-BO verwendet die Geotools-Bibliothek, um z. B. aus Linienzügen der Abstellgleise die umhüllenden Polygone zu berechnen, die das Sachobjekt Betriebspunkt darstellen.

Eigene Implementierung

Im Projekt wurden auch eigene Algorithmen entwickelt. Dazu gehört ein Map Matching-Verfahren, bei dem GPS-Positionen auf vorhandene Gleis- und Streckendaten projiziert werden. Das Verfahren kann auch bei GPS-Ausfall mit Hilfe geometrischer Algorithmen in vielen Fällen weiter korrekte Positionen liefern. Bei Ein- und Ausfahrten in Bahnhöfen werden ebenfalls Meldungen erzeugt, die von anderen Verfahren benötigt werden.

Ein Teilbereich des Map Matchings ist die Abstellungserkennung. Steht ein Fahrzeug eine bestimmte Zeit im Bereich eines Abstellgleises (Daten werden in QGIS gepflegt), so wird eine Abstellmeldung erzeugt. Diese Meldung führt dann in der Fahrzeugeinsatzplanung zu automatischen Aktionen, z. B. kann die geplante Folgeleistung des Fahrzeugs aktiviert werden.

Ein weiteres, selbst entwickeltes Verfahren ist die Berechnung der frühestmöglichen Ankunft, um potentielle Verspätungen frühzeitig zu erkennen. Dabei wird die aktuelle Position des Fahrzeugs mit der Position in einem Normalfahrtprofil verglichen. Das Normalfahrtprofil wird aus allen Zugfahrten auf einer Strecke berechnet. Liegt das Fahrzeug zu einem Zeitpunkt weiter zurück als üblich, wird eine Meldung generiert.

Das System ist selbstlernend, d. h. die Normalfahrtprofile werden regelmäßig aktualisiert, so dass auch geplante Verzögerungen (z. B. durch Baustellen) ohne Datenpflege in die Berechnung eingehen.

Bei der Entwicklung von FZO-BO haben Open-Source-Komponenten wichtige Aufgaben erfüllt und sehr zur Qualität des Systems beigetragen.

Kontakt zum Autor:

Stefan Kowski
Parks Informatik GmbH
Steinstr. 4
58452 Witten
(02302) 20290-0
stefan.kowski@parks-informatik.de

Landesweite Datenerfassungen organisieren und effizient gestalten mit QGIS

Autoren: Ingo Pfisterer, Gordon Schlolaut

Hessisches Landesamt für Naturschutz, Umwelt und Geologie (HLNUG), Gießen
geoSYS, Berlin,

Bei landesweiten Datenerhebungen, i.d.R. von Länderfachbehörden beauftragt, müssen eine Vielzahl von Aufträgen organisiert werden. Das umfasst unter anderem die Bereitstellung von Datengrundlagen, die Sicherstellung der Integrität der erhobenen Daten nach einheitlichen Kriterien und schlussendlich die zeitnahe Überführung in den zentralen, landesweiten Datenbestand.

In der Vergangenheit verursachten diese Anforderungen erheblichen Aufwand, indem Aufträge individuell vorbereitet und Kartierungsergebnisse im Einzelnen geprüft und in den zentralen Datenbestand überführt werden mussten.

Mit einem speziell für die Hessische Lebensraum- und Biotopkartierung (HLBK) entwickelten QGIS-Plugin konnte die eigentliche Datenerfassung nicht nur vereinfacht, sondern gleichzeitig auch die begleitenden Arbeitsschritte systematisiert und deutlich effizienter gestaltet werden. Dabei übernimmt das Plugin eine Vielzahl von Aufgaben, u.a.:

- unterstützt es bei der Auftragsverwaltung,
- generiert individuelle QGIS Projekte für die einzelnen Erfasser mit vorgeladenen Layern aus verschiedenen Quellen, systematisierter Stilisierung und Strukturierung,
- stellt den Erfassern verschiedene Digitalisierungswerkzeuge zur Verfügung,
- steuert durch interaktive Eingabemasken die Erhebung untereinander abhängiger Sachdaten,
- prüft mit QGIS-, Python- und Datenbanktechnologien die Daten schon bei der Erfassung und verhindert sowohl die Eingabe unzulässiger Geo- als auch Sachdaten,
- es stellt dabei dem Nutzer gleichzeitig Informationen über Mängel und Fehler in den Daten für die Behebung bereit,
- bietet vertiefte, themenbezogene Prüffunktionen für die Erfasser,
- verfügt über automatisierte Exportfunktionen für verschiedene Karten, Berichte und Daten und
- kann Datengrundlagen zentral gesteuert aktualisieren (Abb.1).

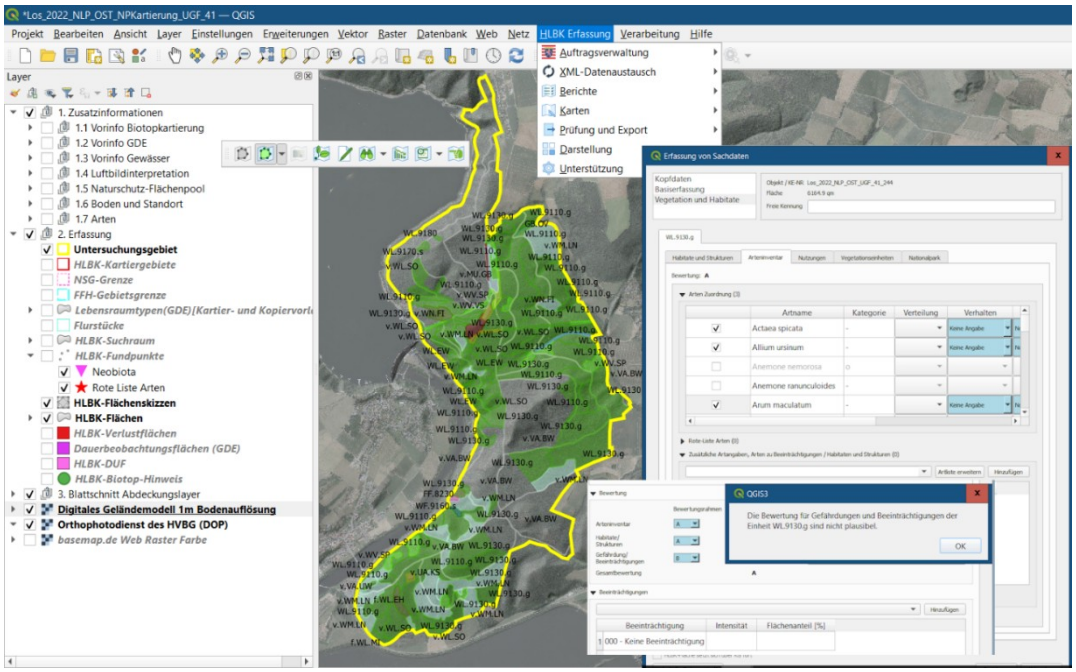


Abb. 1: Benutzeroberfläche und Funktionen der HLBK-Erfassung (Plugin für QGIS)

Bei der Konzeption der Leistungsbeschreibung zur Vergabe und der eigentlichen Entwicklung mussten einige Herausforderungen gemeistert werden. Die zentrale Steuerung von Erfassungsaufträgen, das Downloaden von Stammdaten aus einem Online-Repository und das Einbinden von verschiedenen Web-Geodiensten sind hierfür nur ein paar Beispiele.

Dabei zeigt sich allerdings auch, welches Potential maßgeschneiderte Erweiterungen für QGIS für die Automatisierung komplexer Workflows und für die Effizienzsteigerung des Datenmanagements bieten.

Kontakt zum Autor:

Ingo Pfisterer
Hessisches Landesamt für Naturschutz, Umwelt und Geologie (HLNUG)
Abteilung Naturschutz
Europastraße 10
35394 Gießen
ingo.pfisterer@hlnug.hessen.de

Dr. Gordon Schlolaut
geoSYS
Pflügerstraße 56
12047 Berlin
gordon.schlolaut@geosysnet.de

QGIS fernsteuern: Kickstart mit PyQGIS zur Automatisierung

Dr.-Ing. Helge Städtler, PlanBlue GmbH, Bremen

PyQGIS Schnittstelle

QGIS ist eine Applikation, die ein Python-basiertes **Application Programming Interface (API)** namens PyQGIS zur Verfügung stellt. Die Nutzung des API erlaubt z.B. das Automatisieren von sich oft wiederholenden Aufgaben, beispielsweise Datenmanipulation, Kartengenerierung oder Druckausgabe.

Da QGIS auf der Basis des Qt-Frameworks erstellt wurde, lassen sich per API auch Funktionen von Qt ansprechen. Das bedeutet, Funktionen der grafischen Benutzerschnittstelle (GUI) von QGIS kann man ebenfalls nutzen.

Die Basis der Anwendung auf Qt sorgt zum einen für eine gewisse Plattformunabhängigkeit (Windows, Linux, Mac), zu anderen aber auch für eine einheitliche Ansprache wesentlicher GUI-Funktionen, die man per PyQGIS API einsetzen kann. Viele Plugins für QGIS machen von diesen Möglichkeiten ausgiebig Gebrauch und integrieren sich so nahtlos in die Benutzerschnittstelle von QGIS, z.B. als Werkzeugleiste.

Automatisierung per Eingabezeile

Das Aufgabenspektrum möglicher Automatisierungen hängt von den eigenen Fähigkeiten ab, einzelne Arbeitsschritte in PyQGIS-Programmcode zu formulieren und diesen im QGIS-Kontext auszuführen.

Ein Weg, wie man erste Schritte der Automatisierung ausprobieren kann, ist die Nutzung der **integrierten, interaktiven Python-Konsole** in QGIS. Wie man diese aufruft, lässt sich in der zugehörigen Dokumentation nachlesen [1]. Während der aktiven Entwicklung einer Automatisierungslösung kann dies ein sinnvoller Weg sein, schnelle Ergebnisse zu erhalten und neue Dinge auszuprobieren.

Um eine möglichst flexible Automatisierungslösung zu erstellen, soll im Folgenden allerdings ausschließlich die **Kommandozeile des Betriebssystems** zum Einsatz kommen.

Gründe für die Kommandozeile:

- Automatisiert man von der Kommandozeile aus, braucht es keine laufende QGIS-Instanz mit interaktiver Python-Konsole, die wiederum eine Desktopumgebung mit GUI benötigt.
- Eine Automatisierung von der Kommandozeile aus erlaubt die Einbindung in jeglichen per Skript steuerbaren Kontext, z.B. in eine Continuous Integration (CI) oder eine Reihe von Automatisierungen per Bash-Skript.
- Der Aufruf einer „headless“-Instanz von QGIS verringert den Hauptspeicherbedarf und eliminiert die Gefahr, dass man aus Versehen mit der Applikation auf dem Desktop interagiert, während ein Vorgang abläuft.

Die schwierigste Hürde bei der Nutzung der Kommandozeile ist der Start von QGIS. Das liegt daran, dass QGIS bereits eine eigene und in sich konsistente Python-Umgebung mitbringt, um einwandfrei zu funktionieren.

Um die in QGIS integrierte Python-Umgebung im eigenen Skript zu nutzen, sind einige händische Pfadangaben im eigenen Skript notwendig, die je nach Installationsort von QGIS etwas variieren.

Die ‚magische' Eingabezeile

Das Geheimnis für einen guten Start ist also der Kommandozeilenaufruf, der das eigene Skript im Kontext der Python-Umgebung von QGIS ausführt. Da dieser Aufruf so wichtig, aber zugleich nicht unbedingt offensichtlich ist, soll er hier am Beispiel gezeigt werden:

```
/Applications/QGIS.app/Contents/MacOS/bin/python3 src/mein-script.py
```

Diese Eingabezeile verweist explizit auf das Python-3-Binärprogramm, das innerhalb der QGIS-Applikation (unter macOS) liegt. Durch den direkten Aufruf mit absoluter Pfadangabe wird automatisch die Python-Umgebung (Virtual Environment) von QGIS benutzt. Es wird das Skript im Ordner `src/mein-script.py` ausgeführt. Der Kommandozeilenaufruf ist für andere Betriebssysteme ähnlich.

Das ‚magische' Skript

Das Python-Skript in Abbildung 1 wurde als „standalone"-Skript entwickelt, um mit QGIS in einer nicht-interaktiven (headless) Weise auf einem macOS-System zu interagieren [2]. Es dient als Vorlage für die Automatisierung von QGIS-Aufgaben ohne Benutzerinteraktion und ist geeignet für die Stapelverarbeitung oder skriptgesteuerte Arbeitsabläufe.

Es ist ein wenig Vorarbeit notwendig, um QGIS mit einem passenden PREFIX_PATH und einem passenden PROJECT_PATH zu konfigurieren (Abbildung 1). Beide Pfade basieren auf dem Installationspfad für QGIS, der spezifisch für jedes Betriebssystem ist.

Zunächst wird eine QGIS-Instanz erstellt, die keine grafische Benutzerschnittstelle (GUI) starten soll (Zeile 23). Der eigentliche Start der Anwendung findet durch Aufruf der Methode „initQgis()" statt (Zeile 26). Oft benötigt bereits die Instanziierung von QGIS einige Zeit, daher kann es etwas dauern, bis auf der Kommandozeile die ersten „print"-Ausgaben erscheinen.



```python
import sys
import os

sys.path.append("./src")  # OPTIONAL IF SCRIPT IS STORED IN A src-DIRECTORY
if (sys.version_info.major * 1000 + sys.version_info.minor) < 3006:  # COMPATIBILITY CHECK
    print(f"Python 3.6 or higher required.")
    sys.exit(1)
else:
    print(f"LAUNCHING, PLEASE WAIT ...")  # GIVE EARLY FEEDBACK ON LAUNCH

from qgis.core import QgsApplication, QgsProject, QgsProjectMetadata

# PATH CONSTANTS NEEDED FOR macOS
QGIS_INSTALL_PATH = "/Applications/QGIS.app/Contents/"
QGIS_APP_BINARY_PATH = os.path.join(QGIS_INSTALL_PATH, "MacOS/QGIS")
QGIS_APP_PREFIX_PATH = os.path.join(QGIS_INSTALL_PATH, "MacOS")
QGIS_APP_LIB_PATH = os.path.join(QGIS_INSTALL_PATH, "MacOS/lib")
QGIS_APP_RESOURCES_PATH = os.path.join(QGIS_INSTALL_PATH, "Resources")
QGIS_APP_PROJECTION_PATH = os.path.join(QGIS_INSTALL_PATH, "Resources/proj")
QGIS_APP_PLUGINS_PATH = os.path.join(QGIS_INSTALL_PATH, "Resources/python/plugins")

LAUNCH_WITH_GUI = False
app = QgsApplication([], LAUNCH_WITH_GUI)
app.setPrefixPath(QGIS_APP_PREFIX_PATH, True)
os.environ["PROJ_LIB"] = QGIS_APP_PROJECTION_PATH
app.initQgis()
print(f"HELLO {app.osName()} WORLD!")

# DO THINGS HERE...
project: QgsProject = QgsProject.instance()
try:
    metadata: QgsProjectMetadata = QgsProjectMetadata()
    metadata.setTitle("Kickstart mit PyQGIS")
    metadata.setAuthor("Dr.-Ing. Helge Staedtler")
    project.setMetadata(metadata=metadata)
    project.write(filename="mein-projekt.qgz")
    print(f"PROJECT CREATED & SAVED.")
except Exception as e:
    print(f"ERROR CREATING & SAVING PROJECT. {e}")

app.exitQgis()
print(f"GOODBYE!")
```

Abb. 1: Einfaches PyQGIS Skript „mein-script.py“ zum Erstellen und Schreiben eines Projektes

Zu Demonstrationszwecken wird ein QGIS-Projekt instanziiert (Zeile 30), anschließend werden einige Metadaten für das Projekt definiert und dann der Projektinstanz zugewiesen (Zeile 35).

Zum Abschluss wird das neu erstellte Projekt unter dem Dateinamen „mein-projekt.qgz“ gespeichert bzw. geschrieben (Zeile 36).

Durch einen Aufruf von „exitQgis()“ in wird die QGIS-Instanz ordnungsgemäß beendet (Zeile 41).

TL;DR

QGIS bietet eine Python-basierte API namens PyQGIS, mit der man die Funktionen erweitern und automatisieren kann.

Um eine flexible Automatisierungslösung zu erreichen, nutzt man am besten die Kommandozeile des Betriebssystems:

- Es ist keine laufende QGIS-Instanz mit interaktiver Python-Konsole nötig.
- Die Einbindung in Skripts, CI-Umgebungen etc. ist möglich.
- Der Speicherbedarf ist geringer als mit vollständiger Desktop-Umgebung.

Der Schlüssel ist der Kommandozeilenaufruf, der das eigene Skript im Kontext der Python-Umgebung von QGIS startet, z.B.:

```
/Applications/QGIS.app/Contents/MacOS/bin/python3 src/mein-script.py
```

Das Beispielskript in Abbildung 1 zeigt, wie man QGIS in einer nicht-interaktiven Weise unter macOS steuert. Es konfiguriert QGIS, erstellt ein neues Projekt, setzt Metadaten und speichert es.

Kontakt zum Autor:

Dr.-Ing. Helge Städtler
Holbeinstr. 10, 28209 Bremen
hs@appdoctors.de

Literatur

[1] PyQGIS Cookbook (QGIS 3.34), “Scripting in the Python Console”: https://docs.qgis.org/3.34/en/docs/pyqgis_developer_cookbook/intro.html#scripting-in-the-python-console, Abruf im Feb 2025

[2] PyQGIS Cookbook (QGIS 3.34) – “Using PyQGIS in standalone scripts”: https://docs.qgis.org/3.34/en/docs/pyqgis_developer_cookbook/intro.html#using-pyqgis-in-standalone-scripts, Abruf im Feb 2025

QGIS Server per REST API konfigurieren

Jakob Miksch

Das Projekt QSA (QGIS Server Administrator) ermöglicht es QGIS Projekte über eine REST-Schnittstelle zu konfigurieren.

QGIS Server hat den Vorteil dass alle Einstellungen mit der Benutzeroberfläche von QGIS Desktop getätigt werden können. Manchmal kann es aber notwendig sein die veröffentlichten QGIS Projekte automatisiert zu verändern. Da kommt das Projekt QSA (QGIS Server Administrator) ins Spiel. Die in Python geschriebene Software stellt eine REST-Schnittstelle bereit mit der gängige Einstellungen an der QGIS Projekten durchgeführt werden können. Beispiele sind das Hinzufügen, Entfernen und Ändern von Layern, das Ändern von Stilen und das komplette Erstellen von neuen Projekten.

OpenStreetMap ist doch vollständig …

Michael Reichert

OpenStreetMap ist letztes Jahr 20 Jahre zwar alt geworden. Dennoch ist die Datensammlung weder vollständig noch fertig. Der Erfassungsgrad in den einzelnen Themenfeldern ist regional äußerst unterschiedlich. Der Vortrag zeigt auf, wie man Gegenden und Themen findet, in denen OpenStreetMap noch Aufholbedarf hat. Zielgruppe des Vortrags sind Mapper, die vor Ort Daten erheben möchten. Die gezeigten Methoden eignen sich auch, um spannende Reiseziele zu finden.

Das OpenStreetMap-Projekt feierte letztes Jahr sein 20-jähriges Jubiläum. Dennoch ist die Datenerfassung weder vollständig noch abgeschlossen. Der Erfassungsgrad in den einzelnen Themenfeldern ist regional sehr unterschiedlich unterschiedlich. Während in manchen Städten fast nur noch Datenpflege möglich ist, fehlen in manchen Dörfern noch die Briefkästen und Straßenbeläge.

Ausgehend von einem Bikepacking-Kurzurlaub des Autors in der Eifel zeigt der Vortrag, wie man Reiseziele für Mapping-Urlaube auswählt. Der geübte Blick auf gerenderte OpenStreetMap-Karten, OSM-Daten, und die Versionsgeschichte zeigt Datenlücken und Pflegebedarf auf. Manuelle Abgleiche mit Luftbildern und Fremdquellen ergänzt das Bild. Schwerpunkt des Vortrags sind Straßen und Wege, markierte Wanderrouten und Points of Interest.

DFS Deutsche Flugsicherung: Open Source und Sicherheit bei UAS? Ein Erfahrungsbericht

Vera Werner

Unmanned Aircraft Systems (UAS) etablieren sich als neue Verkehrsteilnehmer im Luftraum. Die DFS unterstützt deren sichere Integration durch die Bereitstellung relevanter Geodaten und setzt dabei auf freie Open Source Software (FOSS) im Geodatenmanagement. Fünf Jahre nach Gründung des Geodatenmanagement-Teams bieten wir Einblicke in Herausforderungen bei der Datenbereitstellung, die Akzeptanz und den Mehrwert von FOSS sowie die Entwicklung der gesetzlichen Rahmenbedingungen.

Seit einigen Jahren etablieren sich neben dem klassischen Luftverkehr immer mehr „Unmanned Aircraft Systems“ (UAS), oder auch einfach Drohnen genannt, als neue Verkehrsteilnehmer im Luftraum. UAS lassen sich nahezu unbegrenzt einsetzen und finden Anwendung in Bereichen wie der Bauinspektion, sowie im Rettungs- und Transportwesen, wobei sich die Einsatzmöglichkeiten kontinuierlich erweitern. Die DFS ist ein privatrechtlich organisiertes Unternehmen im vollständigen Besitz des Bundes. Unsere Aufgabe ist es, den Luftverkehr in Deutschland zu kontrollieren, die Integration neuer Verkehrsteilnehmer zu unterstützen und gleichzeitig die Sicherheit im Luftraum für alle Teilnehmer zu gewährleisten.

Drohnen dürfen nicht überall fliegen. Einheitliche Regeln für den Betrieb, Sicherheitsanforderungen und Zertifizierungen werden durch einen europäischen Gesetzesrahmen vorgegeben und zusätzlich durch die deutsche Luftverkehrsordnung ergänzt. Hier werden u.a. geografische Gebiete aufgelistet, in welchen ein Drohnenflug eingeschränkt bzw. nur unter bestimmten Voraussetzungen gestattet ist. Dazu zählen etwa Flughäfen und Flugplätze, Krankenhäuser, Naturschutzflächen und Anlagen zur Energieerzeugung. Die DFS ist für die Bereitstellung dieser Geodaten verantwortlich und veröffentlicht diese im Auftrag des BMDV auf der Digitalen Plattform für Unbemannte Luftfahrt (www.dipul.de).

Die geografischen Gebiete sind im klassischen Sinne Geodaten, idealerweise Polygone, die das auszuweisende Objekt so realitätsgetreu wie möglich darstellen (möglichst bis zur Grundstücksgrenze). In der Realität passen die gesetzlichen Anforderungen, das Geodatenangebot und die Möglichkeiten der Prozessierung nicht immer überein. Daher wurde im Jahr 2019 ein Geodatenmanagement-Team bei der DFS gegründet, um dem Thema zu begegnen. GIS, Geodaten, OpenData und insbesondere freie und Open Source Software waren bis dahin Randthemen und wurden eher sicherheitskritisch betrachtet. Der etablierte Luftverkehr unterliegt seit Jahrzehnten einer nach und nach gewachsenen Regulierung, die z.B. auch für die Bereitstellung und Verwaltung von aeronautischen Daten und der dazu verwendeten Soft- und Hardwarekomponenten strenge Regeln und Zertifizierungen vorgibt (z.B. das „Vier-Augen-Prinzip“). Der junge Drohnenmarkt hingegen ist noch unreguliert und bot somit die Gelegenheit Open Source Software insbesondere aus dem GIS-Umfeld (QGIS, PostgreSQL, GeoServer) erstmals umfänglich zu nutzen.

Fünf Jahre nach der Gründung des Geodatenmanagement-Teams in der DFS möchten wir einen Einblick in die aktuellen Entwicklungen und Herausforderungen geben. Anhand eines konkreten Beispiels zeigen wir auf, welche spezifischen Hürden bei der Bereitstellung geografischer Gebiete für den UAS-Einsatz auftreten – von der Datenintegration bis zur kontinuierlichen Aktualisierung der Datengrundlage. Des Weiteren beleuchten wir den Einsatz von FOSS im Geodatenmanagement und zeigen, wie diese intern genutzt und wahrgenommen werden. Dabei geht es nicht nur um technische Aspekte, sondern auch um die Akzeptanz und den Mehrwert, den FOSS innerhalb der Organisation bietet.

QGIS im Glasfaser-Ausbau der Deutschen Telekom

Felix von Studsinske, Torsten Drey

Der Anschluss von Haushalten mit Glasfaser durch die Deutsche Telekom ist zu einem hohen Grad automatisiert und wird massiv durch den Einsatz von Geodaten und freier Software gestützt. Ein wichtiger Schritt im Planungsprozess ist die sog. Detailplanung für ein Ausbaugebiet bei der mittels QGIS und dem entwickelten Plugin “Plan[Goo]” der Trassenverlauf, die Kosten und der benötigte Materialeinsatz bestimmt werden. Der Vortrag gibt einen Überblick über den gesamten Prozess und Plan[Goo].

2017 hat sich die Deutsche Telekom auf den Weg gemacht, den Glasfaserplanungsprozess zu revolutionieren und, wo immer möglich, zu automatisieren. Das Ziel damals: ab 2021 sollten 2 Millionen Haushalte pro Jahr mit Glasfaser versorgt werden. Mittlerweile sind über 8 Mio. Haushalte angeschlossen und bis 2030 sollen es alle 41,5 Mio. Haushalte sein. Der Planungsprozess umfasst die Aufnahme von Daten vor Ort (Bilder, Laserscandaten), die Berechnung der Trassenführung, die Erstellung von Unterlagen für Städte und Gemeinden, eine detaillierte Planung inklusive der Material- und Kostenberechnung sowie eine 3D-Aufnahme der gebauten Trassen und Überführung in die Bestandsdokumentation. Das alles verlangt einen massiven Einsatz von Geodaten, deren Prozessierung und Visualisierung für Planer, Genehmiger und Tiefbauer.

Nach einer allgemeinen Übersicht zu dem Planungsprozess, wird der Schwerpunkt des Vortrags auf der Verwendung von QGIS und dem Plugin “Plan[Goo]” liegen, womit wichtige Schritte im Planungsprozess bedient werden. Neben der grundlegenden Nutzung zur Visualisierung von Planungsständen, wird eine Detailplanung durchgeführt, die zwar nicht automatisiert werden kann, bei der aber Aufgaben effizient durchgeführt werden können. Mit Plan[Goo] werden Plantrassen gezeichnet, Plantrassen mit ihren Inhalten der benötigten Rohre und Glasfaserkabel erweitert, Glasfaser-Netzverteiler positioniert und vieles mehr. Plan[Goo] wird sowohl für den eigenfinanzierten Ausbau der Deutschen Telekom sowie für den geförderten Ausbau des Bundes und anderen Kooperationen zur Planung genutzt. Zur Bauausführung werden Exporte angeboten, die der „Bauakte“ angehangen werden können. Zusätzlich wird Plan[Goo] in der Netzdokumentation verwendet, um die tatsächlich gebaute Infrastruktur einfacher und schneller zu zeichnen. Über Schnittstellen kann der dokumentierte Bereich zu anderen Systemen übertragen werden.

Neben der reinen Nutzung und Erweiterung von QGIS bestehen Aktivitäten, Entwicklungen für das Plugin als Open Source zu veröffentlichen, hierzu wird abschließend ein Stand und Ausblick gegeben werden.

Zaubern mit dem Geopackage

Claas Leiner

Das Geopackage kann nicht nur Daten verpacken sondern ganz schnell im Hintergrund Arbeit verrichten, die anschließend im QGIS wie Zauberei sichtbar wird. Die Verschneidung zweier Layer aktualisiert sich, wenn die Quell-Layer geändert werden, ohne das ein weiterer Prozess aufgerufen werden muss. Aggregierungen und Auswertungen sind stets auf dem Stand.

Schließlich ist das Geopackage eine SQLite-Datenbank, in welcher der SQL-Zauberstab zum effektiven GIS-Helferlein wird.

Overpass Turbo goes PostGIS

Frederik Ramm

Der Vortrag stellt einen Dienst vor, der es erlaubt, das weit verbreitete "Overpass Turbo"-Webfrontend mit einer PostGIS-Datenbank anstelle einer Overpass-Datenbank zu verwenden und stellt einige Feature- und Performance-Vergleiche an.

"Overpass Turbo" ist aus der OpenStreetMap-Welt kaum wegzudenken - jeder greift zu diesem Web-Dienst, wenn es darum geht, schnell einen Blick auf OpenStreetMap-Daten zu werfen und dabei auch hinter die Fassade der vorberechneten Karten zu blicken. Zeige mir alle italienischen Restaurants in Rom, welche Spielplätze in Kassel haben eine Altersbeschränkung - so etwas geht mit "Overpass Turbo" fix und, weil es auf der unter anderen vom FOSSGIS betriebenen Datenbank "Overpass" fußt, ohne jede lokale Softwareinstallation.

Weil das "Rapid Prototyping" mit Overpass Turbo so leicht ist, entstehen auch viele kleine und größere Community-Projekte auf Basis von Overpass und seinen zwei etwas kryptischen Abfragesprachen. Viele Threads in Community-Foren, viele Seiten im OSM-Wiki drehen sich darum, wie man mit Overpass dieses oder jenes löst - und einiges davon wäre durchaus mit einer simplen PostGIS-Datenbank genauso gut oder besser möglich. Um "Rapid Prototyping" auch auf PostGIS-Basis zu ermöglichen, fehlt aber ein PostGIS-Backend zu Overpass Turbo.

In diesem Vortrag werden ein auf PostGIS basierendes Backend für Overpass Turbo sowie die notwendigen Änderungen an Overpass Turbo zur Nutzung dieses Backends präsentiert. Das Backend wird vom Autor in einer öffentlichen "Ausprobierversion" betrieben und könnte, wenn es Gefallen findet, künftig auch vom FOSSGIS betrieben werden. Backend sowie Overpass-Patches sind Open Source und werden bis zur Konferenz auf GitHub zur Verfügung stehen.

Geodatenmanagement in einer Baubehörde

Niny Zamora[1]

[1] Wasserstraßen-Neubauamt Nord-Ostsee-Kanal

Zusammenfassung: Bauprojekte binden große Mengen an zeitlichen und personellen Ressourcen. Geodatenmanagement kann dabei helfen, diese effizient zu verwalten. Im fachlichen Umfeld einer Baubehörde ist das Bewusstsein dafür nicht an allen Stellen gleich stark ausgeprägt, da es oft an GIS-Kenntnissen fehlt. Dieses Potenzial auszuschöpfen, ist eine Herausforderung. In diesem Vortrag sollen die QGIS-Erweiterungen ‚Baufeldverwaltung' und ‚Kampfmittelverwaltung' vorgestellt werden, die im Kontext der Erweiterung des Nord-Ostsee-Kanals, Nutzergruppen benötigte Informationen bedarfsgerecht und niederschwellig bereitstellen.
Schlüsselwörter: Kampfmittelräumung, QGIS-Erweiterung, Projektmanagement, Karte, Anwendung

Einleitung

Erweiterungen in QGIS sind ein mächtiges Werkzeug, um komplexe Funktionen bereitzustellen. Sie können aber auch für Nutzer eingesetzt werden, die weniger GIS-affin sind und so dabei helfen, dass Potenzial von Geodaten besser auszuschöpfen. Der Vortrag stellt anhand zweier Beispiele dar, wie anwendungsbezogene QGIS-Erweiterungen eingesetzt werden können, um die Qualität der Planung bei komplexen Bauvorhaben zu steigern, indem sie den Projektbeteiligten die benötigten Informationen bedarfsgerecht und einfach zugänglich präsentieren.

Die Autorin arbeitet als Bauingenieurin im Bereich Geodatenmanagement des Wasserstraßen-Neubauamtes Nord-Ostsee-Kanal (WNA NOK) und beschäftigt sich dort u.a. mit der GIS-Unterstützung von Baumaßnahmen im Bereich des Nord-Ostsee-Kanals.

Erweiterung ‚Kampfmittelverwaltung'

Die Massenbombardierungen deutscher Städte während des 2. Weltkrieges liegen für die meisten von uns bereits lange zurück. Die damit verbundenen Gefahren für Menschen und Umwelt sind jedoch auch heute präsent. Explosive Kampfmittel lagern noch immer im Boden, altern, korrodieren, setzen giftige Stoffe frei und werden mit jedem Jahr gefährlicher. Werden sie verlagert oder erfahren mechanische Einwirkungen können sie detonieren, sogar spontane Zündungen ohne äußere Beeinflussung sind bei bestimmten Munitionsarten möglich.

Schätzungen gehen davon aus, dass sich in Deutschland noch 100.000 bis 300.000 Tonnen Kampfmittelaltlasten im Boden befinden. Die große Menge erklärt sich unter anderem daraus, dass die Zündvorrichtungen der alliierten Bomben unzuverlässig waren und abhängig vom Modell etwa 10 bis 20 Prozent der abgeworfenen Bombenlast nicht explodierten [1]. Schwerpunkte der alliierten Luftangriffe waren deutsche Großstädte wie Bochum, Hamburg und Dresden. Aber auch Kiel als bedeutender Flottenstützpunkt und Standort großer Rüstungsbetriebe war zwischen dem 30.04.1940 und dem 02.05.1945 insgesamt 387-mal Ziel von Angriffen, darunter ca. 120 Großangriffe mit jeweils mehr als 20 Flugzeugen. Insgesamt wurden ca. 75 Prozent der Gebäude der Stadt Kiel zerstört oder beschädigt [2]. Auch die Schleusenanlagen in Kiel-Holtenau als Endpunkt des Nord-Ostsee-Kanals zur Ostsee waren aufgrund ihrer strategischen Bedeutung Ziel von Angriffen.

Ausgehend vom Zerstörungsgrad kann man grob auf die Schwerpunkte der zu erwartenden Kampfmittelbelastung schließen [3]. Für eine genauere geographische Bestimmung und eine erste Gefahrenprognose lassen sich Luftbilder der alliierten Streitkräfte heranziehen. Für die Schleusenanlagen ergab die historische Erkundung, dass dort 90 Prozent der Gesamtfläche kampfmittelgefährdet sind, was auch die 38 Prozent Wasserfläche des Untersuchungsgebiets miteinschließt [2].

Der hohe Anteil potenziell belasteter Flächen macht es notwendig, im Vorfeld von Bau- und Erweiterungsmaßnahmen im Bereich des Nord-Ostsee-Kanals Kampfmittelsondierungen und -räumungen durchzuführen. Diese werden nach Vorgaben des Kampfmittelräumdienstes hinsichtlich der Freigabetiefe durchgeführt, die im Laufe der Zeit Änderungen unterworfen sind. In der Gesamtschau ergibt sich ein sehr unübersichtliches Bild aus vielen, sich teilweise überlagernden Untersuchungen mit verschieden Freigabetiefen auf der gesamten Länge des Nord-Ostsee-Kanals.

Um diesen unübersichtlichen Informationsstand nutzergerecht aufzubereiten und zu präsentieren wurde die QGIS-Erweiterung ‚Kampfmittelverwaltung' entwickelt. Mit ihr werden Daten so aufbereitet, dass dem Nutzer in Echtzeit die jeweils maximale Freigabetiefe für ein Gebiet präsentiert wird und er sich bei Bedarf einen Bericht über alle vorliegenden Untersuchungen in diesem Bereich erstellen lassen kann.

Die einzelnen Untersuchungen werden in der Regel von Dienstleistern durchgeführt und die Ergebnisse an das WNA NOK übermittelt. Dort werden die Daten hinsichtlich topologischer Konsistenz und der Entfernung von Kleinstflächen, die für den Prozess irrelevant sind; aufbereitet, generalisiert und in den Gesamtdatensatz übernommen. Anschließend durchläuft der Datensatz einen mehrstufigen, iterativen Prozess, in dem überlappende Flächenteile in weitere Segmente zerlegt werden, damit dem Nutzer eine übersichtliche Darstellung der Untersuchungsergebnisse präsentiert werden kann.

Der Verarbeitungsprozess ist in Abb. 1 schematisch dargestellt und wird nachfolgend erläutert:

- In der ersten Iteration werden die aufbereiteten Daten in den ‚Target Layer' geladen und auf Überlappungen überprüft (1). Flächen, die keine Überlappungen aufweisen, werden direkt in das Ergebnis übernommen (2). Flächen, die Schnittmengen mit anderen Flächen haben, werden darauf geprüft, ob sie einander teilweise oder vollständig überlagern (3).
- Liegen teilweise Überlappungen vor, werden sie zum Layer ‚Überlappungen n' hinzugefügt und doppelte Flächen mit gleicher Flächengröße und ID entfernt. Aus der Differenz mit den ‚Target Layer' ergeben sich die Flächenteile, die nicht überlappen und mit den vollständig überlappenden Polygonen in den Layer ‚Polygone n' übernommen werden (4), der zum Layer ‚Vereinigung-Layer' addiert wird.
- Der Layer ‚Überlappungen n' wird für die zweite Iteration in den ‚Target Layer' übernommen und durchläuft den Prozess ab Schritt 1 erneut. Dies erfolgt so lange, bis keine Überlappungen mehr auftreten. Dann wird der Layer ‚Überlappungen n' zum ‚Vereinigung-Layer' addiert (5) und doppelte Polygone mit gleicher Flächengröße und Schwerpunkt entfernt und in ‚Layer Final' (6) geschrieben.

Als Ergebnis wird dem QGIS-Projekt ein neuer Layer mit den einzelnen Flächen hinzugefügt, deren Symbologie den aktuellen Status der Kampfmitteluntersuchung repräsentiert (7). Zusätzlich können die Nutzer alle Untersuchungen für eine Fläche über eine Sachdatentabelle abfragen (8).

Bei der Umsetzung des Projektes galt es Lösungen für einige Probleme zu entwickeln:

Im Hinblick auf die *Bereinigung der Information* wurde eine Methode zur Identifizierung und Filtrierung von Objekten des Geometrietyps ‚GeometryCollection' entwickelt, die bei der Verschneidung entstehen. Bei der geometrischen Verarbeitung mussten *Methoden zur Behandlung Überlappungen*, die Erstellung neuer Polygone und dem Umgang mit verbleibenden Splittflächen entwickelt werden. Bei der relationalen *Datenhaltung* galt es, 1: n Beziehungen abzubilden und die Synchronisierung mit den Geometrien umzusetzen, z.B. Speichern der Beziehungen zwischen ‚neue_id' (Polygonen) und zugehörigen Tabelleneinträgen bei gleichzeitiger Entfernung von Duplikaten. Weiterhin mussten Überschneidungen zwischen drei und mehr Polygonen durch einen *iterativen Prozess* behandelt werden, ohne dabei Duplikate zu erzeugen und eine Möglichkeit geschaffen werden, alphanumerische Daten nach mehreren Kriterien wie Tiefe, Datum und Synchronisierung mit den Geometrien zu *sortieren*.

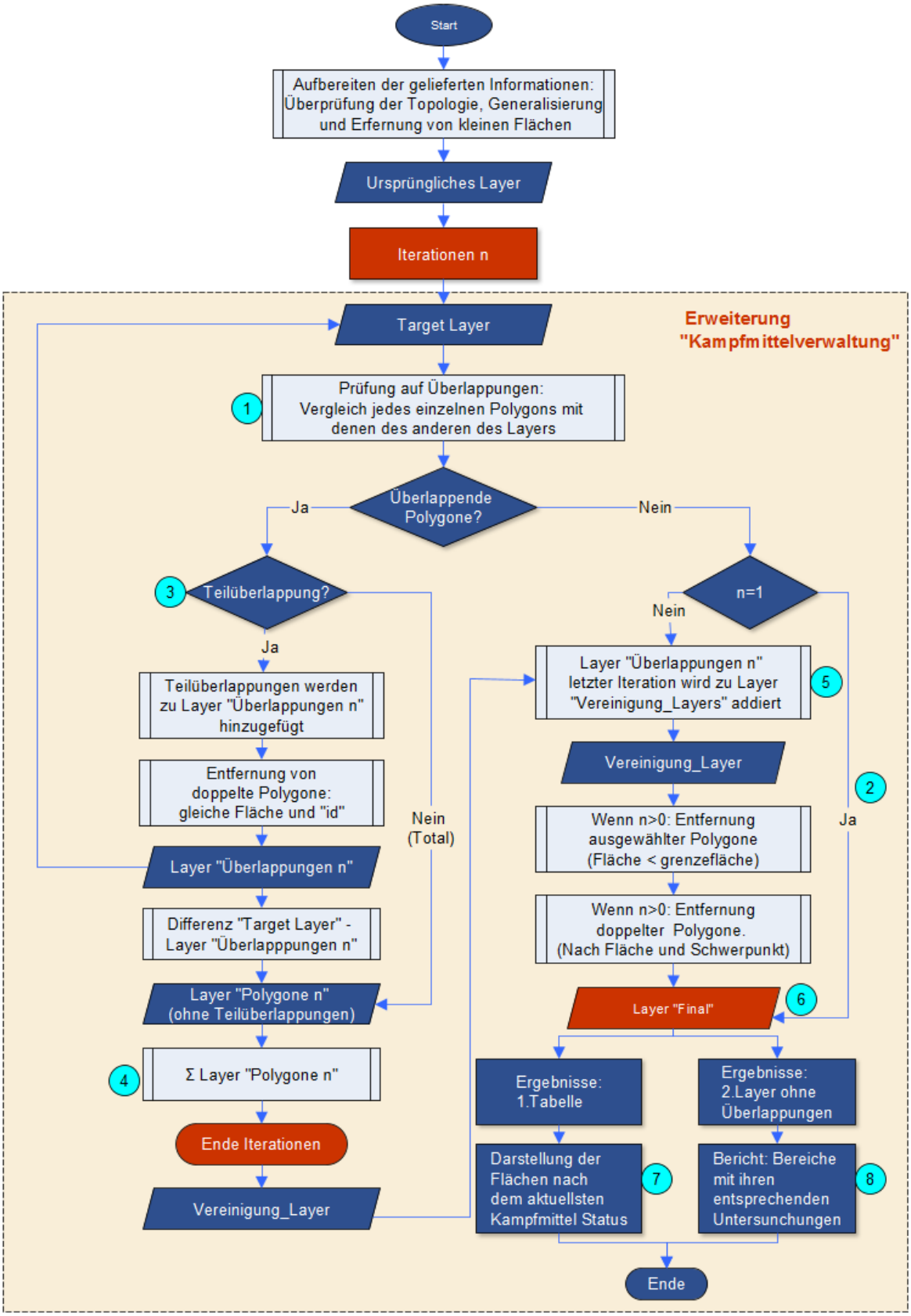
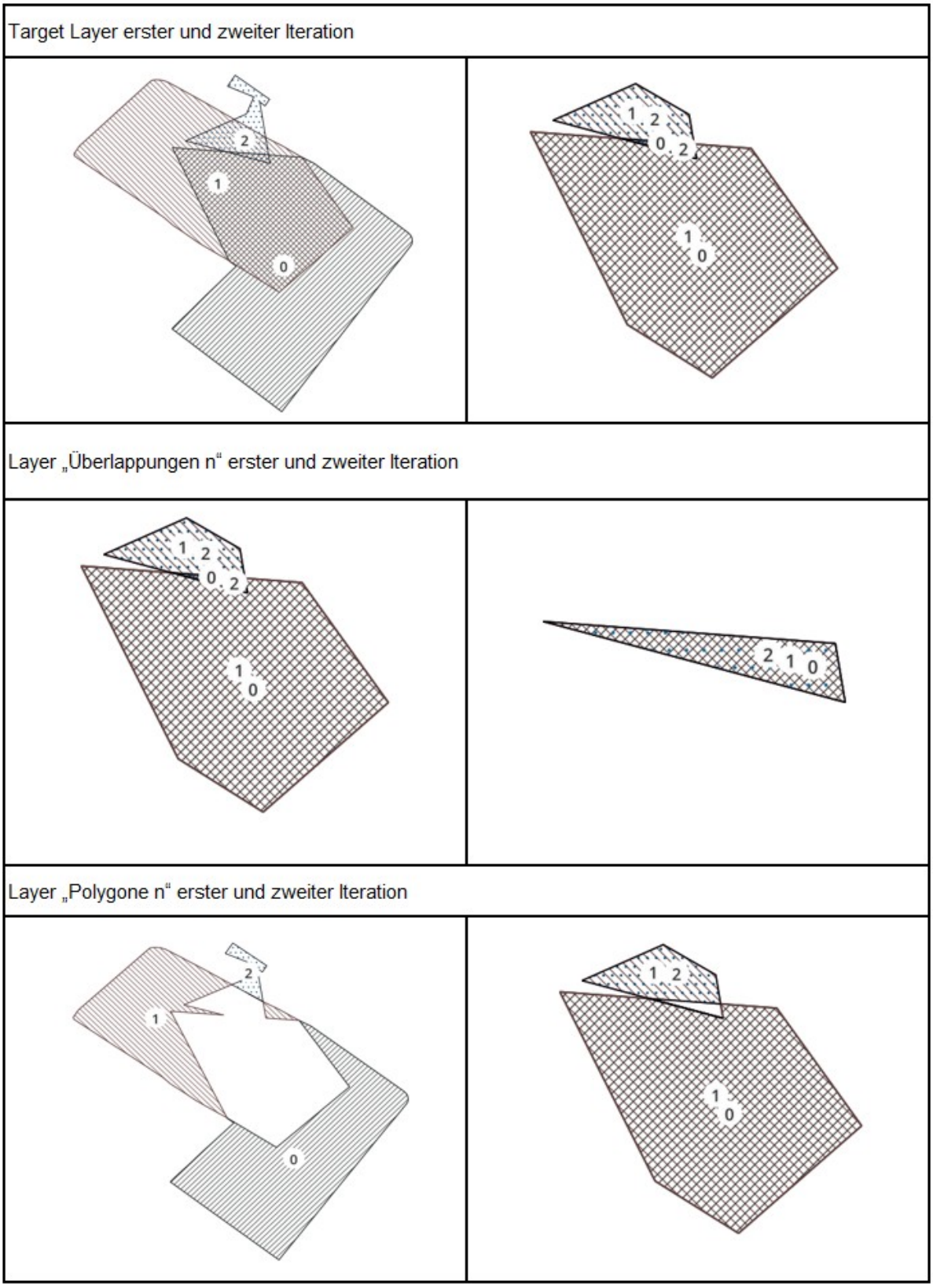
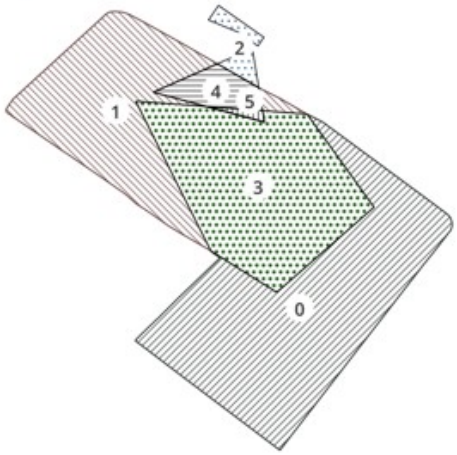


Abb. 1: Prozessschema der Erweiterung ‚Kampfmittelverwaltung'

Beispielhaft werden die Ergebnisse für zwei Iterationen in Abb. 2 dargestellt:



Ergebnisse: Layer ohne Überlappungen und Tabelle mit allen Informationen (Siehe „neu_id“ und „alte_id“ Spalten)



neue_id	alte_id	auftragsnr	tiefe	status
0	0	2017-0001	3	2
1	1	2017-0002	3	3
2	2	2017-0003	2	1
3	0	2017-0001	3	2
3	1	2017-0002	3	3
4	1	2017-0002	3	3
4	2	2017-0003	2	1
5	0	2017-0001	3	2
5	1	2017-0002	3	3
5	2	2017-0003	2	1

Abb. 2: Teilergebnisse und Endergebnis der Verarbeitung der Informationen Erweiterung ‚Kampfmittelverwaltung‘

Erweiterung ‚Baufeldverwaltung‘

Bei einem komplexen Projekt, wie der Sanierung der Schleusenanlage Kiel Holtenau werden verschiedene Baumaßnahmen parallel durchgeführt. Eine gute Koordination der einzelnen Maßnahmen vermeidet Behinderungen und Verzögerung und sorgt damit für einen flüssigen und sicheren Ablauf.

Zur Unterstützung der Koordination wurde die Erweiterung ‚Baufeldverwaltung‘ entwickelt, indem sie den Maßnahmenträgern wesentliche Informationen zu den einzelnen Baustellen (Wo? Wann? Was? und Wer ist zuständig?) bereitstellt. Der Verarbeitungsprozess ist in Abb. 3 dargestellt. Die Erweiterung dient damit als raumbezogenes Werkzeug für das Projektmanagement.

Ziel war es, eine intuitive, einfach zu bedienende Anwendung, die keine GIS-Kenntnisse voraussetzt zu entwickeln. So erhält der Nutzer aktiv Hinweise, für die optimale Verwendung und kann die Ergebnisse als Bericht oder Karte im ausgewählten Maßstab ausgeben.

Bei der Umsetzung des Projektes galt es Lösungen für einige Probleme zu entwickeln:

Erstellung von Karten im optimalen Maßstab: es wurde eine Methode entwickelt, die es ermöglicht, die Polygone zentriert und im optimalen Maßstab unabhängig von der Bildschirmgröße darzustellen.

Zusammenfassung

Mit dem Beitrag sollte demonstriert werden, welches Potenzial QGIS Erweiterungen auch für Nutzer entfalten können, die bisher kaum GIS-Erfahrungen haben. In den Beispielen wurde gezeigt, wie einerseits komplexe Datenbestände aufbereitet und präsentiert werden oder bestehende Prozesse für den Nutzer vereinfacht werden können.

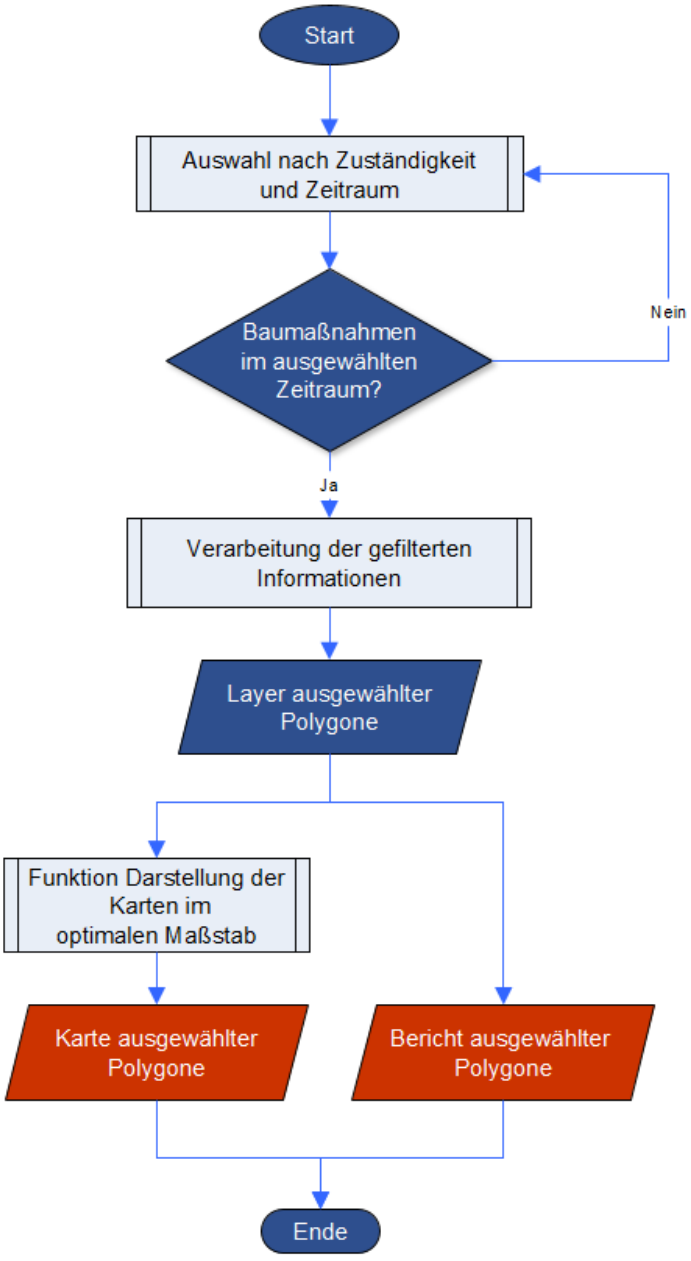


Abb. 3: Prozessschema der Erweiterung ‚Baufeldverwaltung'

Kontakt zum Autor:

Niny Zamora
WNA Nord-Ostsee-Kanal
24159
0431 3603 366
niny.zamora@wsv.bund.de

Literatur und Quellen

[1] https://www.bundestag.de/resource/blob/909216/a5448dd84ac14a5b36fb93f400a4dff1/WD-2-032-22-pdf-data.pdf.

[2] Mull und Partner Ingenieurgesellschaft mbH: Luftbildauswertungen für Liegenschaften des Bundes, Hannover, 2017

[3] https://en.wikipedia.org/wiki/Strategic_bombing_during_World_War_II

Web trifft Desktop

Matthias Kuhn

Mit realen Anwendungsfällen in QGIS und QField zeigen wir, wie das Django-Framework konsumierbare Geodatenebenen als OGC API - Features-Endpunkte erzeugen kann. Indem sowohl das Datenmodell als auch die Geschäftslogik in Python mit dem Django-ORM definiert werden, lassen sich Herausforderungen umgehen, die häufig bei herkömmlichen Datenbankansätzen auftreten. Auf diese Weise demonstrieren wir, wie die Nutzung von Django zu interessanten Perspektiven für solche Anwendungen führen kann.

In der GIS-Welt bedeutet der Aufbau einer Geschäftsanwendungslogik oft die Verwendung eines "Thick Database"-Ansatzes. Das heisst, die Geschäftslogik wird in SQL geschrieben und in der Datenbank ausgeführt oder im Frontend implementiert. Während dieser Ansatz einfach einzurichten ist, hat er insbesondere für Entwickler und Wartungsverantwortliche solcher Anwendungen Schwächen gezeigt. Versionskontrolle, Änderungsmanagement, Modularisierung oder Anpassung, Unit-Tests und die Lesbarkeit des Codes sind einige der Hauptprobleme.

Wir werden die Verwendung von Django, einem Python-Web-Framework, demonstrieren, um konsumierbare Geodatenebenen zu erstellen und Django-Modelle zur Definition des Datenmodells und der Geschäftslogik zu verwenden. Dazu haben wir eine Bibliothek namens Django-OAPIF entwickelt, die auf dem Django REST Framework basiert. Mit einem einfachen Decorator um ein Django-Modell herum wird sie OGC API - Features-Endpunkte bereitstellen, den OGC-Standard zur Veröffentlichung von Vektordaten über eine REST-API. Die Bibliothek ermöglicht CRUD-Operationen, Filterung und kann in einem Desktop-Client wie QGIS oder in Webmapping Bibliotheken wie OpenLayers verwendet werden.

Wir werden einen Vergleich der Ansätze mit Django-OAPIF und PostGIS nebeneinander anstellen und die Herausforderungen, denen wir derzeit begegnen, insbesondere in Bezug auf die Leistung, sowie die Chancen, die die Nutzung von Django für GIS-Anwendungen eröffnet, diskutieren.

Automatische Veredelung von offenen Nahe-Echtzeit-Daten

Benedikt Gräler, Stephan Holl

Vorgestellt wird eine Plattform zur Echtzeit-Verarbeitung von Erdbeobachtungs- und Wetterdaten, die in der Bewältigung von Extremereignissen unterstützt. Open-Data-Quellen wie EFAS, GLOFAS und DWD-Radardaten werden automatisiert auf einer skalierbaren Kubernetes-Plattform verarbeitet und z.B. den Versicherungsunternehmen nahezu in Echtzeit zur Verfügung gestellt. Der Vortrag erklärt das Plattformkonzept und den Mehrwert für verschiedene Branchen.

Automatische Veredelung von offenen Nahe-Echtzeit-Daten zu einem hydro-meteorologischen Informationsprodukt

Ausgehend von der verheerenden Sturzflut im Ahrtal im Juli 2021 hat die Versicherungsbranche den dringenden Bedarf festgestellt, auf derartige Ereignisse besser vorbereitet zu sein. Durch den Klimawandel wird das Auftreten von Extremereignissen wahrscheinlicher, so dass das Risiko von Versicherungsrelevanten Ereignissen steigt.

Daher hat die Versicherungsbranche das Ziel ausgegeben, eine Plattform zur Verarbeitung beliebiger Erdbeobachtungsdaten für die kurzfristige Integration von Echtzeitdaten von aktuellen Überflutungs- und Katastrophenereignissen zu erarbeiten und in Betrieb zu nehmen. In der ersten Ausbaustufe wurden Open Data Satelliten- und Wetter-Daten von Flood Awareness Systeme (wie EFAS, GLOFAS) und Flood Monitoring Systeme (GFM, EMSR), sowie Radardaten des DWD mit ihren jeweils unterschiedlichen Charakteristika untersucht, bewertet und als Branchen-Informationsprodukt bereitgestellt wird.

Das Ziel ist es, schnellstmöglich und vollautomatisiert eine kontinuierliche Bereitstellung der genannten Inputdaten zu gewährleisten. Das resultierende Informationsprodukt beinhaltet die klassifizierten Datenquellen sowie die verwendeten Rohdaten und wird in nahezu Echtzeit („Near-Real-Time“ (NRT)) im WebGIS für weitere Analysen den Endkunden bereitgestellt.

Das gezeigte Plattformpattern implementiert die Datenstromveredelung mit umfangreichen Prozessierungen auf einer maximal skalierbaren Kubernetes-Plattform. Mithilfe von automatisierten Argo-Workflows wird täglich ein Informationsprodukt aus Vorhersage-, Überflutungs- und Verdachtsflächen erzeugt und über eine API für die Weiterverarbeitung angeboten.

Der Vortrag stellt das Vorgehenspattern der NRT-Plattform und ihre eingebundenen Datenquellen sowie Nutzen und Mehrwerte für die Konsumenten der Datenaggregate vor und erlaubt einen tiefen Blick in die OpenSource-Maschinenhalle.

QField: Neue Strategie und Anwendungspotentiale

Berit Mohr

Mit über 1 Million Downloads und 350.000 aktiven Benutzern wird QField als digitales öffentliches Gut anerkannt, dass die UN-Ziele für nachhaltige Entwicklung unterstützt.

Anwendungsfälle aus der Praxis zeigen, wie FeldarbeiterInnen aus aller Welt, Ihre Datenlücken schließen, um qualifizierte und bewusste Entscheidungen zu treffen für das Wohl unserer Lebensgrundlagen und für eine nachhaltige Zukunft

Erfahren Sie im Detail, wie QField die Feldarbeit für Einzelpersonen und Organisationen weltweit verändert. Mit über 1 Million Downloads und 350.000 aktiven Benutzern wird QField als digitales öffentliches Gut anerkannt und ist in jedem Kontext anwendbar. In dieser Sitzung präsentieren wir Anwendungsfälle aus der Praxis von Ozeanien bis Europa, wie gezielt fehlende Daten erhoben und in die weiteren Arbeitsschritte integriert werden. Wir geben einen Einblick in die Strategie von QField und wohin die weitere Reise geht.

Lernen Sie aus Erfolgsgeschichten aus verschiedenen Branchen und entdecken Sie, wie die nahtlose Integration von QField 3 mit QFieldCloud die Feldarbeit auf der ganzen Welt spürbar beeinflusst. Werden Sie Teil der Community und nehmen an der weiteren Reise teil.

Wie können OpenStreetMap und QGIS einen Wegewart unterstützen?

Harald Hartmann

Zusammenfassung: OpenStreetMap und QGIS unterstützen Wegewarte bei der Pflege von Wegen durch präzise Kartierung, Visualisierung und Analyse. OSM ermöglicht das Erfassen und Aktualisieren von Wegen sowie das Hinzufügen von Wegedetails und touristischem Inventar entlang des Weges. Mit QGIS lassen sich diese Daten analysieren und mit weiteren GIS Daten verknüpfen. Weitere Tools erleichtern zudem den mobilen Einsatz und fördern die Zusammenarbeit durch Datenaustausch und Crowdsourcing.

Schlüsselwörter: OpenStreetMap, QGIS, Wegewart

Betrachten wir zunächst das Touristische Inventar entlang des Weges, wie z.B. Wegweiser, Sitzbänke, Rastplätze, Schutzhütten, usw. Viel zu oft wurden solche Daten ohne GIS Kenntnisse erst einmal ganz einfach in einer Excelliste inventarisiert, mit dem Nachteil, dass man die Standorte nicht geografisch auf einer Karte sieht. Der Webdienst uMap [1] bietet hier bereits einen sehr einfachen Einstieg, um das Inventar geografisch zu erfassen, während man bei der Erfassung der Eigenschaften flexibel ist. Natürlich können diese Daten auch mit JOSM [2] erfasst werden. Möchte man diese allerdings zu OpenStreetMap hochladen, sollte man sich an die im OSM Wiki dokumentierten Tags der Objekte halten; alternativ kann man diese Daten aber auch lokal als .osm Datei abspeichern. Wer bereits ist sich richtig einzuarbeiten, kann natürlich auch direkt eine GIS Software, wie z.B. QGIS [3] verwenden.

Wenn wir nun auf die Erfassung der Wege selbst schauen, ist die OpenStreetMap Datenbank – der „digitale Zwilling“ der realen Umwelt – gerade zu prädestiniert dafür. Der Weg, welcher bereits von Anfang an geografisch erfasst wird, kann mit vielen Eigenschaften angereichert werden, wie z.B. die Art des Weges, dessen Oberfläche(nbeschaffenheit), Gangbarkeit, Sichtbarkeit, u.v.m. Eine sehr gute Übersicht bietet hier der Artikel „Wandern“ [4] im OSM Wiki.

Da es sich bei OpenStreetMap um ein Community und CrowdSourcing Projekt handelt, ist ein Großteil der Daten bestenfalls schon vorhanden, oder können auch jederzeit von Anderen beigetragen werden. Des Weiteren helfen viele Tools auch Jemanden, der sich nicht im Speziellen mit OpenStreetMap auskennt, sein Wissen beizutragen, wie z.B. StreetComplete [5], oder „Hinweis/Fehler melden“ [6]. Die Möglichkeit, diverse RSS-Feeds [7] zu abonnieren (z.B. direkt in Thunderbird), helfen dem Wegewart dabei, über solche Hinweise oder Bearbeitungen in seiner Umgebung benachrichtigt zu werden. Mitunter kann es schon einmal vorkommen, dass eine Sitzbank, welche sich beim Bauhof in Winterschlaf befindet, aus den OSM Daten entfernt wird.

Als Wegewart wird man u.U. aber auch mit anderen Themen wie z.B. der Verkehrssicherungspflicht konfrontiert. Hierbei muss ermittelt werden, auf welchem (Flur-) Grundstück ein touristisches Inventar steht. In diesem Fall kann man dann auf die umfangreichen Werkzeuge von QGIS zugreifen. In fast allen Bundesländern werden die Flurstücksdaten (ALKIS) als WFS Layer im Rahmen der Open-Data-Initiative über die „Datenlizenz Deutschland – Namensnennung – Version 2.0“ zur freien Nutzung bereitgestellt. Diesen Layer kann man nun in QGIS laden und die geografischen Standorte des Touristischen Inventars über eine räumliche Abfrage (Geospatial Query) den Grundstücken automatisiert zuordnen.

Die bis hierhin aufgeführten Möglichkeiten sind nur einige Beispiele, wie OpenStreetMap und QGIS einen Wegewart unterstützen können, und sie sind definitiv nur ein sehr kleiner Auszug aus den Möglichkeiten, die einem der FOSSGIS Kosmos bietet.

Kontakt zum Autor:

Harald Hartmann
osm@haraldhartmann.de

Literatur und Quellen

[1] https://umap.openstreetmap.de/

[2] https://josm.openstreetmap.de/

[3] https://www.qgis.org/

[4] https://wiki.openstreetmap.org/wiki/DE:Wandern

[5] https://streetcomplete.app/

[6] https://wiki.openstreetmap.org/wiki/DE:Notes

[7] https://wiki.openstreetmap.org/wiki/Feeds

Leerstandsmelder & A Thousand Channels - Countermappings aus der Zivilgesellschaft

Ulf Treger, Sebastian Fuchs

Leerstandsmelder und A Thousand Channels sind zwei unkommerzielle, zivilgesellschaftliche Plattformen, die Openstreetmap und freie Software nutzen. Zwei Projekt-Beteiligte stellen die Entwicklung ihrer Plattformen vor, berichten von der Praxis mit OSM und FLOSS Software und von den jeweils spezifischen Ansätzen und Fragestellungen, die beim Visualisieren von Inhalten und Geoinformationen aufkommen.

Leerstandsmelder ist eine interaktive Plattform zur Sammlung von Wohnraum-Leerständen in den beteiligten Städten und Regionen. A Thousand Channels entwickelt ein Counter-Mapping Werkzeug, das urbanen Akteur*innen und LGBTIQ+-Projekten die Möglichkeit bietet, von queeren Räumen, Ereignissen und Perspektiven zu berichten und diese Erzählungen räumlich zu visualisieren. Beide Projekte basieren auf Openstreetmap-Karten bzw. Karten, die aus den OSM Geodaten erzeugt werden.

Beide sind Softwareprojekte, denn sie entwickeln freie Software für die Verwaltung wie auch Ausgabe und sie sind Projekte der Kartierung, denn sie nutzen Karten für die Visualisierungen ihrer Gegenerzählungen. Diese sollen – je nach Fokus – zu einer Problematisierung von gesellschaftlichen Defiziten wie der Wohnungskrise, oder der Sichtbarmachung marginalisierter Perspektiven und kollektiver Erinnerungen beitragen.

Sebastian Fuchs für den Leerstandsmelder und Ulf Treger für A Thousand Channels stellen die Entwicklung ihrer Plattformen und der spezifischen Herausforderungen beider Projekte bei der Arbeit mit Geodaten und ihrer Visualisierung vor.

“Wir sehen unsere Projekte als Beispiele für eine Erprobung und Bottom-up Nutzung von digitalen Technologien, wie sie die durch offene Geodaten und Freie Software möglich werden. Wir wollen mit unserem Input anregen, zivilgesellschaftliche Projekte mit OSM zu realisieren, aber auch darauf hinweisen, dass solche Projekte und besonders die ihnen zugrundeliegenden Infrastrukturen und Bibliotheken Unterstützung und Förderung brauchen, damit es mehr und nachhaltig funktionierende Civic Code Projekte geben kann."

https://leerstandsmelder.de/

https://www.a-thousand-channels.xyz/

Amtliche Satellitenpositionierung 2.0

Andreas Gerschwitz

Landesamt für Vermessung und Geoinformation Schleswig-Holstein

Einleitung:

Der neue Satellitenpositionierungsdienst des Bundes und der Länder ermöglicht die Positionierung mit einer Genauigkeit weniger Zentimeter als niederschwelliges Infrastrukturangebot für jeden als open data in Echtzeit.

PPP-RTK-Dienst der AdV:

Die stetig steigenden Nutzerzahlen, insbesondere außerhalb der klassischen Vermessungsbranche machen Satellitenpositionierungsdienste zu einem sehr wichtigen Infrastrukturbeitrag der Geoinformationsverwaltungen und sind ein positives Beispiel für die Digitalisierung. Diese positive Entwicklung wird sich in Zukunft fortsetzen[1]. Mit der Open Data Bereitstellung in den vergangenen Jahren verzeichnet SA*POS* eine immer noch stetig hohe Zunahme der Nutzerzahlen.

Damit die amtlichen GNSS-Korrekturdatendienste diesen Entwicklungen folgen und einen deutlich größeren Nutzerkreis bedienen können, hat die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) die Einführung eines neuen Bund-Länder-Dienstes beschlossen. Die in diesem Dienst verwendete PPP-RTK-Technologie hat das Potential einer Bereitstellung von Korrekturdaten an eine nahezu unbegrenzte Anzahl von Nutzern und öffnet sich somit perspektivisch auch für Massenmarkt-Anwendungen. Um eine hohe Verfügbarkeit zu erreichen, erfolgt die Berechnung der Korrekturdaten redundant sowohl durch das Bundesamt für Kartographie und Geodäsie (BKG) als auch durch die von der Zentralen Stelle SA*POS* (ZSS) betriebene Instanz der Länder (ZSS-Instanz). Beide Instanzen verwenden als Vernetzungssoftware die Software GNSMART der Firma Geo++ GmbH, sowie das offen gelegte Format SSRZ[2] zur Übertragung der Korrekturparameter. In der konkreten Auswertearbeit werden die deutschlandweiten SA*POS*®- und GREF-Stationen in Teilnetze aufgeteilt, die SSRZ-Daten aller Teilnetze zentral berechnet und zu einem deutschlandweiten Datenstrom vereint. Der Dienst steht seit Januar 2025 als open data online zur Verfügung (https://gepos.sapos.de).

Dieser Dienst wird per Internet via Ntrip[3] bereitgestellt. Bisherige Untersuchungen des DAB+ (Digital Audio Broadcasting) haben gezeigt, dass sich dieses Medium technisch für die Aussendung der PPP-RTK-Korrekturdaten ebenfalls eignet. Daher wird der Dienst in einer Testphase ebenso per DAB+ ausgesendet. Alternativ zu DAB+ könnten zur Aussendung der Korrekturdaten auch geostationäre Satelliten und alternative Telekommunikationsstandards, wie z.B. 3GPP LPP (3rd Generation Partnership Project, LTE Positioning Protocol) in Betracht gezogen werden. Die AdV wird diese Entwicklungen zu bewerten haben.

Die Projektgruppe PPP der AdV hat im Rahmen von ersten Belastungstests im November 2023 und im Februar 2024 die deutschlandweiten SSRZ-Korrekturdatenströme beider Instanzen auf insgesamt 23 SA*POS*®-Monitorstationen in der Bundesrepublik Deutschland geprüft. Mit diesem gelungenen Tests hat die AdV erstmals gezeigt, dass die bisher nur in wissenschaftlichen Studien getroffene Aussage einer Positionierungsgenauigkeit mittels PPP-RTK-Technik von wenigen Zentimetern auch über die sehr große Ausdehnung eines gesamten Landes, der Bundesrepublik Deutschland, realisierbar ist.

Diese ausführlichen Tests zeigen deutlich, dass der PPP-RTK-Dienst der AdV bereits an die Genauigkeiten des SA*POS*®- HEPS-Dienst heranreicht. Es wurde in verschiedenen Bundesländern Positions-

abweichungen von unter 1,5 cm für die Lage- und unter 2,5 cm für die Höhenkomponenten erreicht. Standardabweichung und RMS-Werte bewegen sich in ähnlichen Bereichen. Somit konnte die generelle Eignung der PPP-RTK-Technologie mit den bereits realisierten Komponenten des neuen Dienstes für die Übertragung von GNSS-Korrekturdaten gezeigt werden. In Abb. 1 werden beispielhaft die Ergebnisse der SAPOS®-Monitorstation Freising gezeigt. Die Netze beider Instanzen unterscheiden sich in Auswahl und Dichte der Referenzstationen, sind aber in den Einstellungen zur Berechnung und Modellierung gleich konfiguriert. Während dieser Testphase Anfang Februar 2024 zeigt die Ionosphäre eine im Vergleich zu vorherigen Tests erhöhte Aktivität.

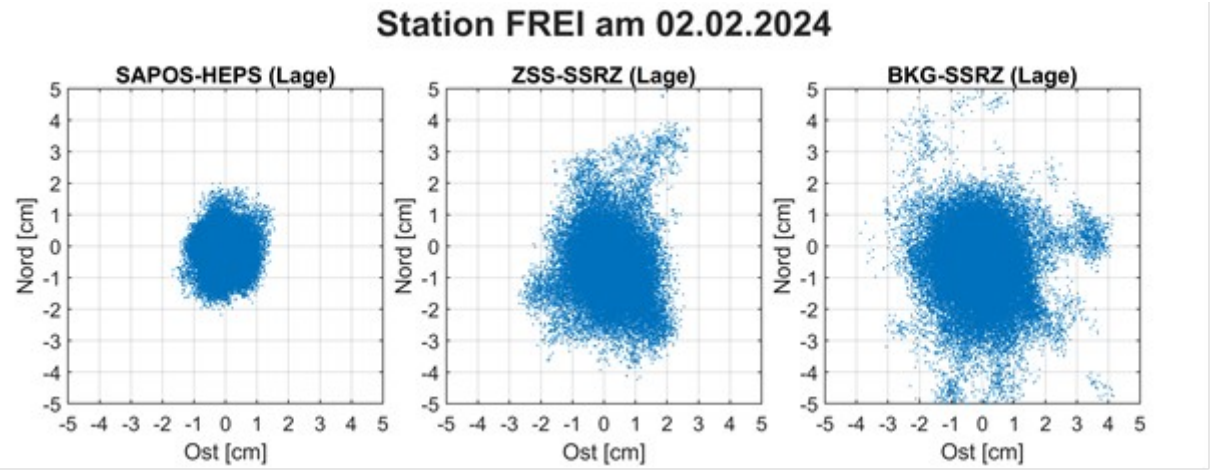


Abb. 1: Lageabeweichung der Monitorstation Freising bei Verwendung von SAPOS®-HEPS bzw. ZSS- und BKG-Korrekturdatenstrom am 02. Februar 2024

Ausblick:

Die kommenden zwei Jahren sollen als Optimierungsphase genutzt werden, um die Positionsgenauigkeit möglichst im laufenden Betrieb zu verbessern. Dies ist wichtig für Anwendungen, die genaue Positionsdaten erfordern, wie Vermessung, präzise Landwirtschaft oder autonomes Fahren.

Des Weiteren soll das Redundanzkonzept noch robuster aufgestellt werden, damit der Dienst bei technischen Problemen einer Instanz kontinuierlich und zuverlässig arbeitet.

Abschließend erlaubt die Optimierungsphase neue Technologien wie z.B. neue Mobilgeräte und weitere Kommunikationsmedien, noch zu entwickelnde Standards und sich ändernde Nutzeranforderungen zu prüfen und zu integrieren, um den Dienst noch performanter für eine breite Nutzung aufzustellen.

Ein DANK gilt den weiteren Projektbeteiligten: Lena Albert | Martin Freitag | Thomas Goetze | Ilka von Gösseln | Hans-Jürgen Goldan | Franziska Hartmann | Cord-Hinrich Jahn | Mathias Meißies | Christoph Knöfel | Rebecca Reiling | Jens Riecken | Franziska Riedel | Christian Rost | Axel Rülke | Volker Schneider | Jürgen Siebert | Timon Stockhaus | Markus Vennebusch | Vitaly Winter

Kontakt zum Autor:

Andreas Gerschwitz
Landesamt für Vermessung und Geoinformation
Mercatorstraße 1, 24106 Kiel
0431/3832075
Andreas.Gerschwitz@LVermGeo.landsh.de

Literatur

[1] EUSPA (2019): https://www.euspa.europa.eu/sites/default/files/calls_for_proposals/rd.03_-_ppp-rtk_market_and_technology_report.pdf, letzter Zugriff 12/2023.

EUSPA (2022): EUSPA EO and GNSS Market Report ISSUE 1, EU Agency for the Space Programme, https://www.euspa.europa.eu/sites/default/files/uploads/euspa_market_report_2022.pdf, letzter Zugriff 01/2024.

[2] SSRZ (2023): State Space Representation Format. http://www.geopp.de/wp-content/uploads/2022/11/gpp_ssrz_v1_1_2.pdf , letzter Zugriff 12/2023.

[3] Weber, G., Dettmering, D. (2005): Networked Transport of RTCM via Internet Protocol (Ntrip) – IP-Streaming for Real-Time GNSS Applications, Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2005), Long Beach, CA, September 2005, 2243-2247.4

[4] Vennebusch et al., Zwischenziel erreicht: Neuer amtlicher PPP-RTK-Dienst der AdV erreicht SA*POS*®-HEPS-Genauigkeit, zfv 2/2024

[5] Gerschwitz, Klesen, PPP-RTK – neue Technik für die massentaugliche SA*POS*®-Nutzung, Forum 3/24

SMASH, Stand der Technik der digitalen Feldkartierungs-app

Antonello Andrea

Diese Präsentation gibt einen Einblick in den Stand der Technik des SMASH-Ökosystems und seine aktuelle Roadmap. SMASH, die "Smart Mobile App for Surveyor’s Happiness", ist eine schnelle App für die digitale Feldkartierung.

SMASH, die "Smart Mobile App for Surveyor’s Happiness", ist eine schnelle App für die digitale Feldkartierung.

Die quelloffene Flutter-App für Android, IOS (und auf Anfrage auch für Linux, Macos und Windows) ist vollgepackt mit Funktionen, wie zum Beispiel: Geopackage- und PostGIS-Bearbeitung, Kalman-Filter für GPS-Logs, nativer Geotiff und Shapefile Support, Geojson-Visualisierung, SLD-Styling für Vektordatensätze.

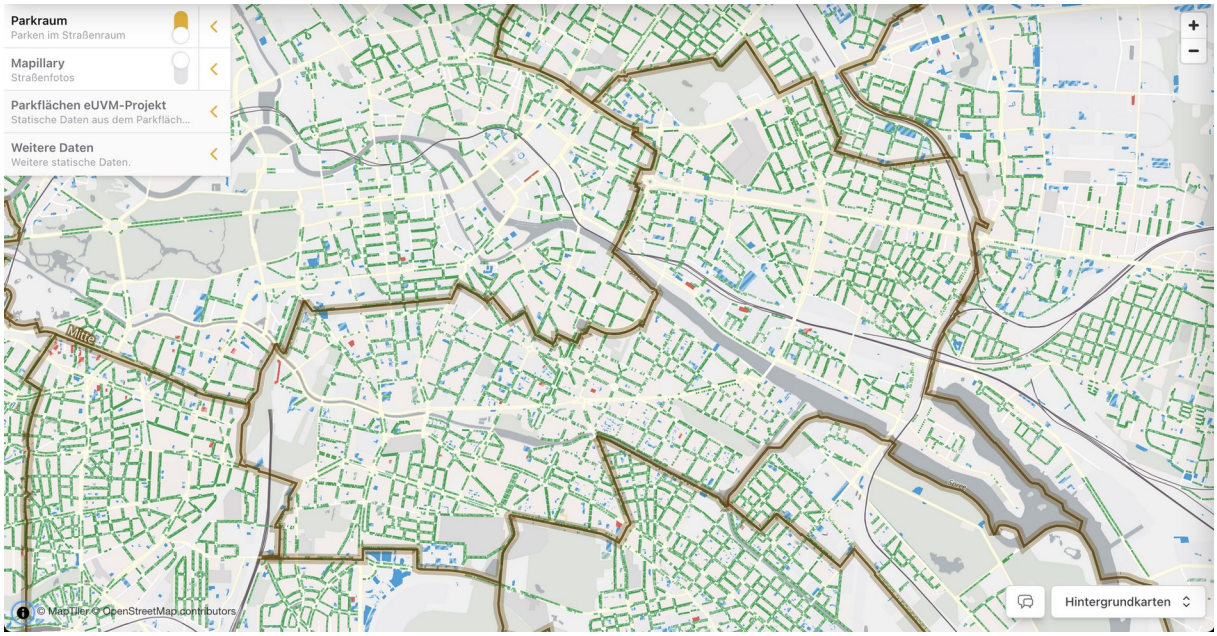
Das Web-Gegenstück zu SMASH ist der Survey Server, eine Webanwendung, die es Gruppen von Vermessern ermöglicht, die Datenerfassung zu zentralisieren. Die Benutzer können die Daten aus der App synchronisieren, aber auch spezielle Formulare und Projekte sowie Basiskarten und Datensätze herunterladen. Der Server basiert auf der gleichen Technologie wie die mobile App und visualisiert die Daten mit dem gleichen Look and Feel.

Im letzten Jahr gab es einige wichtige Erweiterungen, wie z.B. die bidirektional synchronisierten Vektorebenen und den neuen Formbuilder, der sowohl auf dem Server als auch auf der mobilen Seite verfügbar ist.



Parkraumdaten aus OSM mit der Verwaltung pflegen (Praxisbericht)

Alex Seidel, Tobias Jordans



Seit zwei Jahren kooperiert die Berliner OSM-Community mit dem Bezirksamt Friedrichshain-Kreuzberg, um hochwertige Daten zum Parkraum bereitzustellen. Dieser neuartige Ansatz zeigt Potenziale auf, wie Herausforderungen in der Nutzung von OSM-Daten für amtliche Zwecke. Nach der initialen Datenerhebung geht es darum, Prozesse zu etablieren, um die Daten aktuell zu halten. In diesem Praxisbericht berichten wir über unsere Erfahrungen in der Kooperation sowie bei der gemeinsamen Datenpflege.

Seit 2021 arbeitet die OSM-Verkehrswende-Gruppe daran, hochwertige und aktuelle Daten zum Parken im öffentlichen Raum auf Basis von OpenStreetMap zu erzeugen (parkraum.osm-verkehrswende.org).

In dieser Zeit konnten wir verschiedene Erfahrungen zur Datenerfassung sammeln – bspw. über die Community, über Apps wie StreetComplete und über beauftragte Mapping-Projekte. Eine besondere Stärke der Datenhaltung in OpenStreetMap gegenüber herkömmlichen Verfahren, wie regelmäßigen Neubefahrungen, ist jedoch die fortlaufende und einfache Aktualisierung der Daten. Im Rahmen unserer Kooperation mit dem Bezirksamt Friedrichshain-Kreuzberg in Berlin haben wir verschiedene Ansätze erprobt, um die Aktualisierung der Daten zu systematisieren und in die dazugehörigen Verwaltungsabläufe zu integrieren.

In diesem Praxisbericht wollen wir Erfahrungen aus verschiedenen Projekten besprechen, wie OpenStreetMap und Verwaltungen kooperieren können, um OpenStreetMap-Daten für alle Nutzungsgruppen aktuell zu halten.

Wie mache ich eine gute Karte mit QGIS?

Mathias Gröbe

Eine Karte mit QGIS zu erstellen ist einfach – aber wie erstellt man eine gute Karte? Die Demosession soll Grundlagen der Kartengestaltung vermitteln: wie man gute Farben wählt, wie groß oder klein Schriften sein darf und welche Liniendicke gut aussieht. Nach einem kurzen Teil mit Theorie folgt die praktische Umsetzung in QGIS: ausgehend von Daten, über die Signaturen bis zum Drucklayout. Dabei werden nützliche Tipps und Tricks gezeigt, wie man Daten gut lesbar und verständlich mit QGIS Funktionen visualisiert. Der Fokus liegt dabei darauf, alltägliche Funktionen (Farben, Schriften, Legende, Maßstab, Kartenrahmen) möglichst gut zu nutzen.

Neue Methoden zur autarken Indoor-Positionsbestimmung

Benjamin Würzler, Philipp Fiedler

Wie kann die eigene Position in einem Gebäude bestimmt werden, wenn jede bekannte Infrastruktur fehlt? Im Projekt „Denied Areas“ erforscht das Bundesamt für Kartographie und Geodäsie innovative Wege zur autarken Positionsbestimmung – unabhängig von WLAN und Bluetooth. Mithilfe tragbarer Sensoren werden potenzielle Lösungen untersucht und erprobt. Erleben Sie spannende Einblicke in aktuelle Ansätze und Technologien im Bereich der Indoor-Positionsbestimmung und -Navigation.

Wie können sich Personen schnell und präzise im Inneren eines Gebäudes orientieren? Welche Wege führen am effizientesten zum nächsten Ausgang oder zu einem bestimmten Raum? Diese Fragen sind nicht nur für die alltägliche Orientierung, sondern vor allem in Notfallsituationen wie Feuerwehreinsätzen oder im Katastrophenschutz von zentraler Bedeutung. Seit 2021 beschäftigt sich das Bundesamt für Kartographie und Geodäsie (BKG) im Projekt „Denied Areas“ mit dem Thema der Positionsbestimmung im Innenbereich. Die Besonderheit hierbei: Es soll nicht auf bestehende Infrastruktur wie WLAN oder Bluetooth-Beacons zurückgegriffen werden.

Unser Ansatz konzentriert sich stattdessen auf die Nutzung tragbarer Sensoren, die eine eigenständige Orientierung in Gebäuden ermöglichen. Hierzu untersuchen wir verschiedene Technologien und Methoden, um die Vorteile der jeweiligen Sensoren zu nutzen und eine Sensorfusion durchzuführen. Bei dieser Technik werden Daten von mehreren Kleinstsensoren, darunter Barometer, Beschleunigungssensoren, Gyroskope oder Kameras, kombiniert, um die individuellen Schwächen der einzelnen Sensoren zu kompensieren und präzisere Positionsdaten zu erhalten.

Im Projekt untersuchen und testen wir eine breite Palette an Systemen und Geräten, von handelsüblichen Smartphones über speziell entwickelte Prototypen auf Grundlage von Mikroprozessoren bis hin zu fortschrittlichen Inertialnavigationssystemen. Zusätzlich profitieren wir vom Austausch mit wissenschaftlichen Partnern, welche wichtige Impulse für neue Ansätze und Technologien liefern. Unsere bisherigen Forschungsergebnisse haben uns ein detailliertes Verständnis über die Stärken und Schwächen der verschiedenen Sensoren und Methoden in der Praxis vermittelt.

In unserem Vortrag möchten wir Einblicke in unsere bisherigen Erkenntnisse geben und die Potenziale der untersuchten Technologien für konkrete Anwendungen wie Rettungseinsätze und Katastrophenschutz erörtern. Darüber hinaus werden wir die praktischen Anforderungen diskutieren, die für einen erfolgreichen Einsatz erfüllt sein müssen, und einen Ausblick auf mögliche zukünftige Lösungen für die autarke Indoor-Navigation geben.

OGIS in der Öffentlichen Verwaltung

David Arndt

FOSSGIS e.V.

QGIS ist ein benutzerfreundliches Open-Source geographisches Informationssystem (GIS), das unter der GNU General-Public-License steht. QGIS ist ein offizielles Mitglied der Open-Source-Geospatial-Foundation (OSGeo). Es läuft unter Linux, Unix, Mac OSX, Windows und Android und unterstützt eine Vielzahl von Vektor-, Raster- und Datenbankformaten und -funktionen.

In diesem Anwendertreffen wollen wir die QGIS Anwender, oder die es werden wollen in einem Netzwerktreffen zusammenbringen. Ziel ist es Kontakte zu knüpfen, Erfahrungen auszutauschen und gemeinsam QGIS in der Öffentlichen Verwaltungen zu etablieren.

Kontakt zum Autor:

Dipl. Geogr. David Arndt
FOSSGIS e.V.
Bundesallee 23
10717 Berlin
david.arndt@fossgis.de

Literatur

[1] QGIS-Projektseite: https://qgis.org

[2] QGIS-DE Anwendergruppe Deutschland e.V.: https://qgis.de

MerginMaps - Mobile Geodatenerfassung beim RVR

Stefan Overkamp

Zusammenfassung: Seit 2024 setzt der Regionalverband Ruhr (RVR) die Software MerginMaps ein. Auswahlkriterium war u.a. die gute Integrationsmöglichkeit in die GDI des RVR.

Wesentliche Funktionen

- mobile (Geo)Datenerfassung im Gelände und Synchronisieren ins Backend
- Weiterbearbeiten in QGIS mittels Plugin
- Arbeiten auf einem gemeinsamen Datenbestand mit Erkennung von Konflikten
- Versionierung aller Änderungen und dabei detaillierte Dokumentation

Schlüsselwörter: Mobile Datenerfassung, QGIS

Der **Regionalverband Ruhr (RVR)** mit seinem Eigenbetrieb Ruhr Grün ist nicht nur einer der größten kommunalen Waldbesitzer Deutschlands sondern auch für die Pflege und Unterhaltung vielfältiger Infrastrukturen (z.B. dem regionalen Radwegenetz) verantwortlich. Neben der Unterstützung an mehr als 150 QGIS-Desktop-Arbeitsplätzen spielt die Verfügbarkeit der Geodaten vor Ort also mobil „im Gelände“ zunehmend eine Rolle.

Im Jahr 2024 hat das Referat Geoinformation und Raumbeobachtung hierzu in seiner Docker-Umgebung die Software Mergin Maps [1] eingerichtet. Der RVR setzt beim Thema Geoinformation durchgängig auf OpenSource-Software und digitale Souveränität.

Aktueller Schwerpunkt der mobilen Datennutzung ist die Verkehrssicherung. Hier geht es darum, Gefahren u.a. von Bäumen frühzeitig zu erkennen, Schäden (auch fotografisch) zu dokumentieren und Maßnahmen zur Gefahrenminimierung einzuleiten.

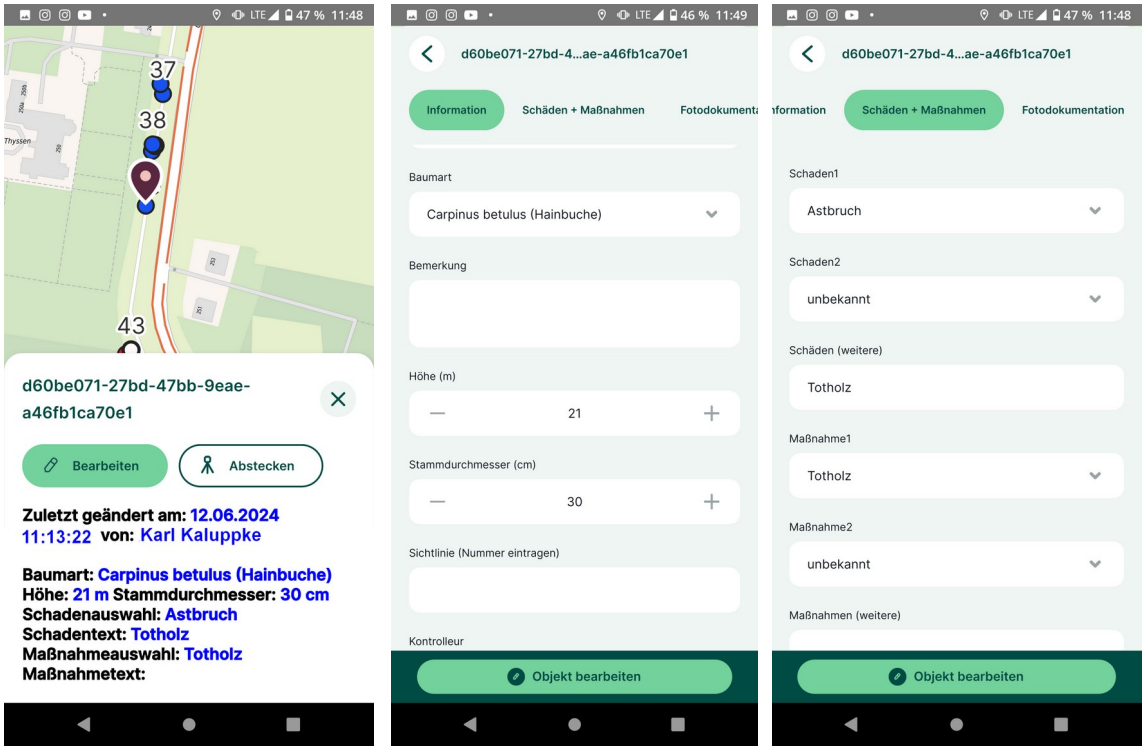


Abbildung 1: Bedienoberfläche Mobilgerät: Karte und Formular

Die im Gelände auch ohne Internetzugang (offline) bereitstehenden Fachgeodaten werden in Mergin Maps durch eine QGIS-Projektdatei konfiguriert und über Geopackages zwischen Mobilgeräten und Büroarbeitsplätzen ausgetauscht (wenn wieder eine Onlineverbindung steht).

Bei den Geopackages wird unterschieden in nur-lesbare Hintergrund-Layer (z.B. Prozessschutzflächen) und bearbeitbare Layer (z.B. Baumkontroll-Punkte). Für letztere werden alle Änderungen (GeoDiffs), also neue, geänderte oder gelöschte Baumkontrollen erkannt und beim Synchronisieren als Änderung an das Mergin Maps-Backend gesendet.

Durch automatische befüllte Feldinhalte ist jederzeit intern nachvollziehbar, welcher Mitarbeitende wann mit welchem Gerät wo Baumkontrollen erfasst oder verändert hat.

Die Fotodokumentation ist dabei ein wichtiger Aspekt und erlaubt die Zuordnung mehrerer Fotos zu einer Kontrolle. Fotos verschiedener Erfasser sind im Büro unter QGIS für alle einsehbar. Auf dem Mobilgerät sehen diese (aufgrund des begrenzten Speicherplatzes) optional nur die selbst aufgenommenen Fotos.

Die Eingabemasken auf den Mobilgeräten sind über das QGIS-Attributformular vielfältig gestaltbar. Verschiedenste Typen von Bedienelementen können ausgewählt und mit Restriktionen und Vorgabewerten konfiguriert werden.

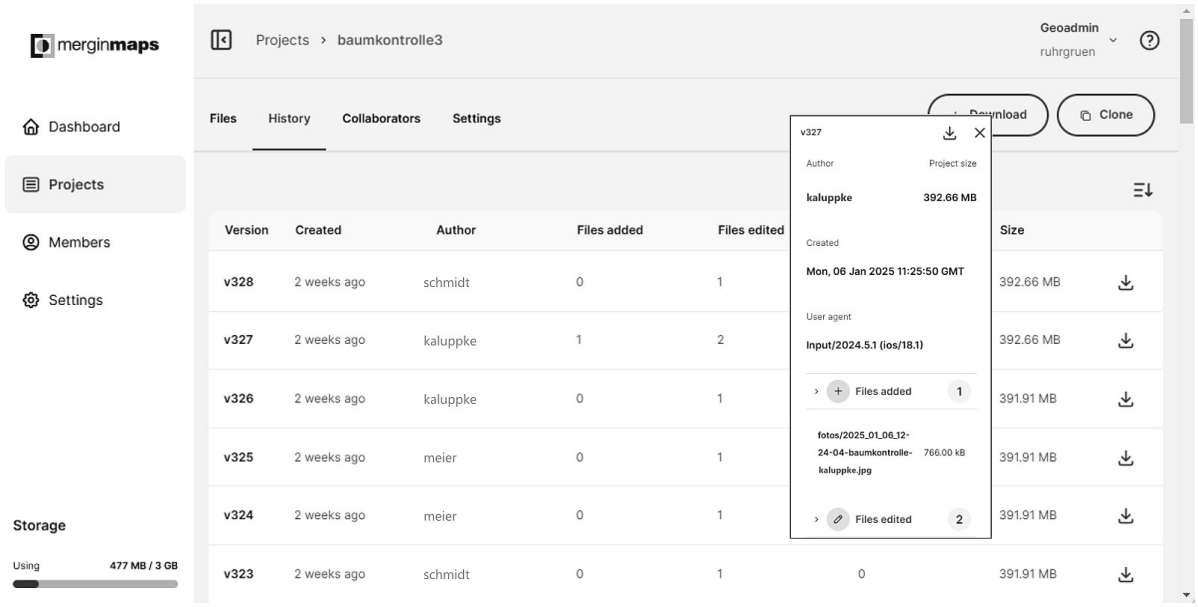


Abbildung 2: Verwaltungs-Backend (Dashboard) mit Versionen und Details einer Version

Die (kostenpflichtige) Enterprise Edition von Mergin Maps erweitert die freie Community Edition um eine bessere Strukturierung der Mobilprojekte in sog. Arbeitsbereiche (bei uns z.B. ruhrgruen und radrevier) sowie die Administration der Zugriffsrechte auf dieser Ebene. Eine zusätzliche Rolle „Editor“ verhindert, dass Datenerfasser die Konfiguration über das QGIS-Projekt ändern können.

Die Softwareinstallation besteht aus einem Docker-Stack mit MerginMaps Frontend und Backend Containern ergänzt um eine PostgreSQL-Datenbank, sowie Redis und einen nginx-Proxy.

Bei der Erfassung können **Konflikte** auftreten, wenn zwei verschiedene Personen innerhalb einer Version die gleichen Eigenschaften an einer Baumkontrolle unterschiedlich ändern und dann synchronisieren. Das Backend erzeugt dann eine JSON-Datei, die die Unterschiede (z.B. unterschiedliche Schadensbeschreibung) dokumentiert. Auf dieser Basis kann der Konflikt inhaltlich gelöst und die Baumkontrolle entsprechend korrigiert werden.

Das **QGIS-Plugin** bietet neben der Datenbearbeitung und Synchronisation u.a. einige Verarbeitungswerkzeuge. So können z.B. Änderungen zwischen zwei Versionen analysiert oder ein Bericht der aggregierten Änderungen an mobil bearbeiteten Layern erstellt werden.

Darüber hinaus ermöglicht das Plugin das Paketieren eines vorhandenen QGIS-Projektes für die mobile Datenerfassung. Dabei werden für die einzelnen Layer des Projektes wahlweise die Daten in ein Geopackage in das Projektverzeichnis konvertiert oder sie werden inklusive ihrer Datenquelle beibehalten oder sie werden ignoriert. Für größtmögliche Kontrolle bietet es sich an, selbst ein entsprechend vorbereitetes Projekt mit allen Datenquellen in einem lokalen Verzeichnis anzulegen. Dann kann dieses Projekt mittels Plugin zu einem Mergin Maps-Projekt gemacht und in das Backend hochgeladen werden.

Fotos zu mobil erfassten Daten sollten zur Wahrung der Übersicht in einem Unterverzeichnis abgelegt werden. Dies unterstützt Mergin Maps ebenso wie die Umbenennung der Fotodateien (z.B. mit Zeitstempel, Layername und Erfasser) und die Reduzierung der Auflösung und somit Dateigröße.

Über das **Dashboard** der MerginMaps-Webanwendung können für jede Version und jeden Erfassungslayer (= Geopackage) im Detail die durchgeführten Änderungen auf Objekt- und Feldebene nachvollzogen werden. Durch Nutzung der MerginMaps-API kann so auch ein automatisiertes Einspielen der Änderungshistorie in eine Geodatenbank eingerichtet werden.

Die mobile Datenerfassung mit MerginMaps ermöglicht beginnend im Bereich Verkehrssicherung eine kontinuierliche Datenerfassung ohne Brüche und eine detaillierte Dokumentation. Weitere Projekte wie die Mängelerfassung durch unsere Radwegepatinnen und -paten stehen bereits in den Startlöchern.

Kontakt zum Autor:

Stefan Overkamp
Regionalverband Ruhr, Team 9-2 Geodaten-Technik
Kronprinzenstraße 6, 45128 Essen
0201 / 2069-700
overkamp@rvr.ruhr

Literatur und Quellen

[1] Mergin Maps Dokumentation: https://merginmaps.com/docs/

[2] Mergin Maps Server auf GitHub: https://github.com/MerginMaps/server
Docker Images Community Edition (CE): https://hub.docker.com/u/lutraconsulting

[3] Collecting Field Data with QGIS and Mergin Maps, 132 Seiten 08/2024, Kurt Menke & Alexandra Bucha Rášová,, https://locatepress.com/book/mergin-maps

[4] Öffentliche Projekte zum Abschauen: https://app.merginmaps.com/projects/

BRouter und BRouter-Web für Anfänger und Fortgeschrittene

Arndt Brenschede, Serge Bellina

BRouter ist das flexibelste Werkzeug für Streckenplanung auf Basis von OSM-Daten und seit 10 Jahren besonders bei Radfahrern und Wanderern beliebt. Wir geben zunächst eine Einführung in die wichtigsten Verwendungsarten, sowohl mobil als auch am heimischen PC, und richten uns dabei auch an Neu-Einsteiger. Im zweiten Teil des Vortrags geht es dann um spannende Neu-Enwicklungen, insbesondere um die Berücksichtigung von Umgebungseigenschaften für noch bessere Routing-Ergebnisse.

Seit einem Jahr stellt der Fossgis e.V. eine leistungsfähige und hochverfügbare IT-Infrastruktur für die Datenvorverarbeitung und den Betrieb des BRouter-Servers und verwandter Dienste zur Verfügung. Dadurch wurde die aufwändige Berechnung der "Pseudotags", auf denen die Berücksichtigung der Umgebungseigenschaften beruht, möglich gemacht. Ausserdem ist BRouter damit fit für die weiter wachsende Benutzer-Basis, die auch wegen der nun erreichten Servicequalität und Hochverfügbarkeit weiter wächst. Auch andere Angebote wie "bikerouter" oder "LoRouter" bieten Routing-Dienste auf Basis von BRouter, Darüber hinaus wird BRouter auch in grossem Umfang komplett offline auf mobilen Geräten verwendet.

FOSS-GIS+FOSS in der Wind- und Solarbranche (Praxisbericht)

Lars Roskoden[1], Björn Kilian[1]

[1] rkDrei GmbH

Zusammenfassung: Mit QGIS und PostgreSQL / PostGIS / PostNAS am Anfang eines Wind- oder Solarparkprojekts kann die Projektierung auch „nach der Karte“ schneller, genauer, übersichtlicher, also effizienter gestaltet werden. Denn auch nach der Nutzung vom FOSS-GIS-Paket kann es mit FreeOpenSourceSoftware noch lange weitergehen. Dieser Praxisbericht zeigt, wie eine professionelle Zukunftsbranche für die Energieversorung eines Industrielands von FOSS profitieren kann.

Schlüsselwörter: #QGIS #QGIS-Server #PostgreSQL #PostGIS #ALKIS #PostNAS #Leaflet #php

QGIS wird in der Projektierung von Wind- und Solarparks häufig schon verwendet. Doch nach der Identifizierung neuer Projektgebiete und der Erstellung von Karten ist meistens schon Schluss.

Basierend auf einer zentralen PostgreSQL/PostGIS-Datenbank werden ALKIS-Daten, die früher oder später für die Projektgebiete notwendig sind, in erster Linie nur als Geodaten mit QGIS dargestellt und bearbeitet. Es ist aber auch möglich und sinnvoll weitere Fach-Daten, die keinen zwingenden geographischen Bezug haben, auf der selben Datenbank mittels Web-Browser und Web-GIS darzustellen und mit weiteren projektrelevanten Informationen anzureichern. So können z.B. personenbezogene Daten von Grundstückseigentümern oder daraus abgeleitete Vertrags- und Zahlungsdaten erfasst und verwaltet werden. Die Ergebnisse lassen sich wiederum in QGIS oder einem Web-GIS darstellen. Im konkreten Fall der Energieprojekte können so Sicherungsstatus, Vertragsstatus, aber auch Stand der Pacht- oder Nutzungsentgelte auf der Karte tagesaktuell abgerufen werden.

Die Vorteile hinsichtlich Schnelligkeit, Genauigkeit und eben zentraler Datenhaltung entbürokratisieren viele innerbetriebliche Prozesse. Mit QGIS kann einerseits die gesamte Beplanung mehrerer Projektgebiete inkl. Wegeführung und Kabeltrassen durchgeführt und zusätzlich mit Nutzung von Geodaten aus der zentralen PostgreSQL/PostGIS-Datenbank optimiert werden. Die verschiedenen Fachabteilungen, welche im Laufe der Projekte zeitlich versetzt aktiv werden, können so ohne größere Übergabeverluste besser zusammenarbeiten.

Wir konnten z.B. die Erstellung von Nutzungsverträgen mit den Grundtstückseigentümern inklusive aller Karten und Anlagen automatisieren. Somit sind (proprietäre) Office-Programme zur Erstellung von Dokumenten oder zur Verwaltung von Projektübersichten - meist mit redundanter Datenhaltung - nicht mehr zwingend notwendig.

Darüber hinaus stehen mittlerweile sehr viele verschiedene datenbankbasierte FOSS für die Büroprozesse der Projektierungsunternehmen zur Verfügung, die ein erstaunliches Potential zur Effizienz-Steigerung haben: DokumentenManagementSysteme (DMS), Arbeitszeiterfassung, Kalender, Kontakte, CustomerRelationManagement (CRM), EnterpriseResourcePlaning (ERP).

Das bedeutet, dass von der ersten Projektidee für einen Wind- oder Solarpark bis zu dessen späteren Betrieb die Nutzung von FOSS das Potential hat, nicht nur alle Büroprozesse abzubilden sondern auch viele Planungsarbeiten durchzuführen und beide zu dokumentieren.

Kontakt zum Autor:

Lars Roskoden, Björn Kilian
rkDrei GmbH
David-Gilly-Str. 1, 14469 Potsdam
+49 331 97910101
info@rkdrei.de

QWC2 als WebGIS in der kommunalen Verwaltung und Einblick in die 3D Funktionalität

Sandro Mani, Daniel Cebulla



Dieser Vortrag befasst sich mit der Web-GIS-Anwendung QGIS Web Client 2 (QWC2). Seit Herbst 2024 wird die bisherige 2D-Anwendung um eine 3D Ansicht erweitert. In diesem Vortrag wird zum einen der aktuelle Stand sowie ein Ausblick der Entwicklung der 3D Ansicht vorgestellt. Außerdem wird der Vortrag die Verwendung des QWC2 exemplarisch an ausgewählten Portalen präsentieren und dabei aufzeigen, welches Potential das System zur Anwendung in kommunalen Verwaltungen bietet.

Der QGIS Webclient 2 (QWC2) ist ein umfangreiches und mächtiges Web-GIS, was durch seinen Funktionsumfang die verschiedensten Anwendungsfälle abbilden kann. Ob als reines Auskunftssystem oder als Werkzeug zur Erfassung und Bearbeitung von Geodaten, vieles ist damit möglich.

Mit der Anwendung können Dank des QGIS Servers Projekte mit demselben Rendering wie QGIS Desktop im Internet veröffentlicht werden. Die Umgebung besteht aus einem modernen responsiven Frontend, das in JavaScript auf Basis von ReactJS und OpenLayers geschrieben wurde. Mit den QWC-Services, ein Ökosystem von serverseitigen Python/Flask-Microservices, kann der Funktionsumfang zudem erweitert werden, um damit zum Beispiel Benutzerrechte steuern und in der Webanwendung Geodaten editieren.

Durch die Veröffentlichung des Webclients als Open Source ist die Hürde zum Einsetzen des Web-GIS relativ gering. Es ist umfangreich nach den eigenen Vorstellungen konfigurierbar und auch mit etwas Grundwissen ohne externe Hilfe zu implementieren. Professionelle Unterstützung ist dennoch durch verschiedene Akteure gegeben.

Der QWC2 bietet seit vielen Jahren eine ausgereifte 2D Kartenwebanwendung für die Publikation von QGIS Projekten im Web. Dank der Kooperation der Städte Detmold, Frankfurt, Jena und Wolfsburg wird die Anwendung seit Herbst 2024 um eine 3D Ansicht erweitert. Diese nutzt die auf THREE.js basierende 3D-Geovisualisierungsbibliothek Giro3D zum Darstellen von 3D-Tiles und Geländemodellen. Eine Tile-Prozessierungs-Pipeline erlaubt zudem das Generieren von 3D-Tiles aus verschiedenen gängigen 3D Formate wie CityGML oder CityJSON.

In diesem Vortrag wird zum einen der aktuelle Stand der 3D Ansicht des QWC2 vorgestellt, mit einem Einblick in dessen Architektur, einigen Anwendungsbeispielen und einem Ausblick auf künftige Funktionen.

Die Weiterentwicklung des Systems lebt von dessen Nutzern, sodass jeder die Möglichkeit hat, Funktionen nach den eigenen Wünschen im Projekt beizutragen bzw. von Dritten entwickeln zu lassen. Dadurch eignet sich das Web-GIS unter anderem auch sehr gut zur Anwendung in kommunalen Verwaltungen! Ob als Auskunftssystem einer Gemeinde für dessen Bürger oder verwaltungsintern als Werkzeug in den eigenen Prozessen, vieles ist mit dem QWC2 möglich.

Der Vortrag soll daher außerdem die Verwendung des QGIS Webclient 2 exemplarisch an ausgewählten Portalen präsentieren und dabei aufzeigen, welches Potential das System zur Anwendung bietet.

Nicht von neuen Algorithmen überflutet: OSM-Daten für WaterwayMap verarbeiten

Amanda McCann

Heutzutage gibt es viele komplizierte Tools und Datenbanken. WaterwayMap.org zeigt, dass alle Flüsse und Bäche miteinander verbunden sind und verschiedene QA-Datensätze für den gesamten Planeten erstellt werden.

In diesem Vortrag werden einige der grundlegenden Algorithmen erläutert, die zur Erstellung dieser Daten verwendet werden, sowie die sehr komplizierte Speichermethode „einfach in den Speicher einlesen“.

Unterstützung des regionalen Naturschutzes in Luxemburg mit Open Source GIS-Software

Tobias Mosthaf

Der regionale Zweckverband für Umwelt- und Naturschutz SIAS in Luxemburg (https://www.sias.lu/) plant Naturschutzmaßnahmen für seine 21 Mitgliedsgemeinden und begleitet deren Umsetzung.

Die 16 Mitarbeiter des SIAS führen dabei z.B. Kartierungen des Großen Feuerfalters (Tagfalter) durch oder planen das Anlegen von Streuobstwiesen. Wie sie dabei von Open Source GIS-Software unterstützt werden, soll in diesem Vortrag erläutert werden.

Im luxemburgischen regionalen Zweckverband für Umwelt- und Naturschutz SIAS hat sich die Anzahl der Mitgliedsgemeinden in den letzten 15 Jahren mehr als verdreifacht und die Anzahl der Mitarbeiter in den letzten fünf Jahren mehr als verdoppelt. Mit diesem Wachstum haben sowohl die Menge als auch die Vielfalt der Aufgaben stark zugenommen, was bei den Mitarbeitern zu dem Wunsch geführt hat eine Datenbank aufzubauen, um Informationen einfacher austauschen und Maßnahmen besser koordinieren zu können.

Im Jahr 2022 wurde diese Aufgabe von den Mitarbeitern selbst gestartet. Dazu wurde nach dem Prinzip „Learning by Doing“ eine PostgreSQL-PostGIS-Datenbank sowie ein WebGIS zur einfachen Bearbeitung der Daten mit dem Webclient Lizmap aufgebaut. Inzwischen werden etwa 1600 Flächen, auf welchen der SIAS Maßnahmen umsetzt, über die Datenbank verwaltet. Dazu zählen das Pflanzen von Obstbäumen und Hecken, die Pflege von Streuobstwiesen oder das Anlegen von Gewässerrandstreifen. Auch für die Kartierung gefährdeter Tier- und Pflanzenarten wird die Datenbank genutzt, zum Beispiel für die Kartierung des Großen Feuerfalters auf ca. 300 Flächen.

Im Vortrag wird u.a. auf die folgenden Punkte eingegangen: das komfortable Vorbereiten der Karten für das WebGIS mit dem QGIS-Plugin Lizmap oder die Nützlichkeit des engen Austauschs zwischen Nutzern der Anwendungen (PostgreSQL-PostGIS Datenbank, Lizmap WebGIS) und dem internen Management dieser Anwendungen (z.B. Erweiterung von Dropdown-Listen im WebGIS, Anpassen von Trigger-Funktionen der Datenbank). In der Zukunft würden wir gerne eine aktuell genutzte proprietäre Software zum Erstellen einer weiteren Web-Anwendung durch Open Source Software ersetzen und sind dazu sehr an Ideen zu möglichen Open Source Alternativen der FOSSGIS Besucher interessiert.

State of SHOGun: Ein flexibles Web-GIS Framework

Daniel Koch



SHOGun ist ein Anwendungsframework zum Aufbau von einfachen bis komplexen Geodateninfrastrukturen. Dabei kann SHOGun direkt und ohne spezifische Anpassungen verwendet oder aber auch nach Bedarf stark angepasst werden, um die Anforderungen spezifischer Projekte zu erfüllen.

SHOGun ist ein Open Source Anwendungsframework zum Aufbau von Geodateninfrastrukturen. SHOGun ist in seiner Architektur dabei so ausgelegt, dass es sich zwischen den Extremen „off-the-shelf“ (Standardprodukt von der Stange) und „fully handcrafted“ (spezifische Individualentwicklung) bewegen kann.

Funktional stellt SHOGun die gewohnten Standardfunktionen eines modernen WebGIS bereit. Diese sind neben der Möglichkeit der Konfigurierbarkeit über eine Nutzeroberfläche z.B. um Tools wie die Editierungsfunktionen oder die Möglichkeit zur Vergabe von Berechtigungen auf Clients und OGC-Diensten auf Basis von externen Nutzerverwaltungssystemen ergänzt.

In dem Vortrag werden wir den aktuellen Stand des Projekts vorstellen und dabei das modulare Zusammenspiel der einzelnen Komponenten präsentieren. Wir werden auf die vorhandenen Funktionalitäten eingehen und dabei den Fokus auf die individuelle Erweiterbarkeit legen. Denn alle Komponenten können, z.B. durch Plugin-Systeme, auf spezifische Fragestellungen dynamisch erweitert und angepasst werden. Weiter sollen auch einige konkrete Projektbeispiele aus unterschiedlichen Branchen aufgezeigt werden.

OSM2World: Updates für den 3D-Pionier

Tobias Knerr

OSM2World gehört zu den Pionieren der 3D-Darstellung von OpenStreetMap-Daten. Neue Features wie die Erzeugung von 3D Tiles und ein modernes Web-Frontend stellen sicher, dass der freie Renderer auch in Zukunft Maßstäbe setzt.

OSM2World verwandelt OpenStreetMap-Daten in 3D-Modelle. Das Projekt wurde bereits 2010 gegründet und hat seither nicht nur enorm an Funktionalität hinzugewonnen, sondern sich auch den Veränderungen in der Softwarelandschaft angepasst und von den neuen Möglichkeiten profitiert.

In den letzten beiden Jahren gab es wieder zahlreiche Neuerungen, u.a. auch dank der finanziellen Förderung durch den Prototype Fund. Hierzu zählen das neue GeoDesk-Backend, Level-of-Detail-Unterstützung und umfangreiche Verbesserungen beim Import und Export von glTF-Modellen. Highlights sind eine browserbasierte Darstellung mit Babylon.js und die Erzeugung standardkonformer 3D Tiles. Diese Veränderungen bilden neben einem kurzen Überblick über die Möglichkeiten des freien 3D-Renderers den Schwerpunkt des Vortrags.

Kartografie-Verbesserungen, Tips und Tricks in QGIS

Andreas Neumann

In dieser Demo-Session sollen viele der jüngeren Verbesserungen im Bereich der QGIS Symbologie und der automatischen Beschriftung anhand von Kartenbeispielen aus der Praxis gezeigt werden. Des weiteren werden einige ältere, vermutlich wenig bekannte Möglichkeiten demonstriert, die schon länger in QGIS bestehen: selektives Maskieren von Beschriftungen und Symbolen, übersteuerbare Symbolebenen, komplexe Linientypen und Tips und Tricks rund um Expressions und daten-definierte Eigenschaften.

Die folgenden Techniken sollen gezeigt werden:

- Verschachtelte Diagramme (stacked diagrams, neu in QGIS 3.40)
- Halbseitige Diagramme (neu in QGIS 3.40)
- Komplexere Beschriftungen mit HTML-Auszeichnung und Bildern im Text (neu in QGIS 3.40)
- Beschriftungen im Zusammenhang mit linearer Referenzierung (neu in QGIS 3.40)
- Selektives Maskieren (freistellen) von Beschriftungen und Symbolen
- Kartografisch korrekt strichlierte Linien (volle Striche bei Linienbeginn und -ende, sowie bei spitzen Winkeln)
- Symbolebenen und Zeichnungsreihenfolge getrennt von Ebenenreihenfolge
- Datendefinierte Eigenschaften und Tips und Tricks mit Expressions
- Blend modes und Kombinationen mit Raster- und LiDAR Daten
- Mehrere Stile per Ebene und Abspeichern von Kartenthemen (Kombinationen von Stilen und Ebenensichtbarkeiten)
- Animierte Darstellungen mit dem Time-Manager und datendefinierten Eigenschaften.

Die oben erwähnten Techniken werden gezeigt anhand von konkreten Kartenbeispielen in der Praxis, wie sie z.B. im Kartografie-Unterricht gezeigt werden oder beim Atlas der Schweiz eingesetzt werden.

Mapbender - Neuigkeiten aus dem Projekt

Astrid Emde

Im Mapbender-Projekt ist im letzten Jahr wieder viel passiert. Es soll von den Neuigkeiten im Projekt berichtet werden und neue Features sollen vorgestellt werden.



- Es liegt ein neues Element zum Routing mit Mapbender vor.
- Stile von WMS Layern können nun im Client im Ebenenbaum ausgewählt werden.
- Die Usability von Mapbender im Backend wurde an einigen Stellen verbessert.
- Mapbender verfügt nun über eine REST API. Befehle, die bisher nur auf dem Server auf der Konsole ausgeführt wurden, z.B. zum Laden von Diensten, können nun auch über die REST API angesteuert werden.
- Auch im Bereich der Datenquellen geht es voran. Das Projekt arbeitet daran, weitere Datenquellen zu unterstützen.
- In QGIS gibt es nun ein Plugin QGIS2Mapbender, das die Interaktion mit Mapbender aus QGIS ermöglicht. QGIS Server Projekte können über das Plugin leicht veröffentlicht werden und direkt in Anwendungen in Mapbender eingebunden werden.

Außerdem wollen wir einen Blick auf bestehende Mapbender-Lösungen richten und deren besondere Anforderungen oder Lösung vorstellen.

OpenLayers - Neues und Nützliches

Marc Jansen, Andreas Hocevar

OpenLayers, das Urgestein der Open Source Web Mapping Libraries, hat weiterhin steigende Nutzerzahlen und eine aktive Entwickler-Community. In diesem Vortrag zeigen wir neben neuen auch altbekannte nützliche Funktionen, aber auch spannende Anwendungsmöglichkeiten, die sich aus der Kombination mit anderen Libraries ergeben.

Mit QGIS zum digitalen Prozess 1 - Gebäudeeigenschaften

Marius Schäfer[1]

[1] Kreis Unna

Zusammenfassung: Der Prozess der außendienstlichen Erfassung von Gebäudemerkmalen für die Wertermittlung war früher zeitintensiv und hat somit viele materielle sowie personelle Ressourcen beansprucht.

Dank einer detaillierten Erfassung der diversen Arbeitsschritte konnte QGIS so konfiguriert werden, dass die Programmierung eines Plugins nicht notwendig ist. Der große Umfang von Grundfunktionen in QGIS bietet ausreichend Mittel, um die zuvor erfassten Schritte optimiert abbilden zu können.

Dieser tatsächlich digital stattfindende Ansatz spart Zeit ein und bietet Mitarbeitenden mit weniger Berufserfahrung eine unkomplizierte und effektive Möglichkeit zur Datenerfassung. Gleichzeitig werden der Verbrauch von Papier und die Nutzung von Dienstfahrzeugen deutlich minimiert.

Schlüsselwörter: QGIS, Digitalisierung, QGIS-Plugin, Gutachterausschuss

„So hat das Bearbeiten sogar richtig Spaß gemacht“

In der Ära der Digitalisierung rückt oft die Technologie, insbesondere digitale Dateiformate wie Word-Dokumente, Excel-Tabellen oder PDFs, in den Vordergrund. Doch bei genauerem Hinsehen wird deutlich, dass die eigentliche Revolution nicht nur in der bloßen Digitalisierung von Dateien liegt, sondern vielmehr im Durchdenken von Prozessen und der Implementierung von innovativen Methoden.

Ein solches Durchdenken fand im Sachgebiet 62.4 Wertermittlung und Gutachterausschuss des Kreises Unna statt. Denn die Geschäftsstelle des Gutachterausschusses erstellt halbjährlich (automatisiert) einen Trendbericht des Immobilienmarktes. Darin werden verschiedene Kennzahlen und Faktoren des Immobilienmarktes berechnet und visualisiert. Ein wichtiger Bestandteil ist der Sachwertfaktor (SWF). Dessen Ermittlung soll den zu modernisierenden Prozess darstellen. Die bestimmenden Einflüsse des SWF sind unter anderem die verschiedenen Eigenschaften eines Hauses, z.B. Haus- & Dachtyp, Anbauten, usw. Um verlässliche Aussagen über diese Attribute treffen zu können, wurden die dazu benötigten Häuser außendienstlich von Mitarbeitenden der Geschäftsstelle begutachtet, ohne dabei die Innenräume zu betreten.

Durch diese Methode mussten innerhalb von zwei bis drei Wochen halbjährlich um die 350 Gebäude im Kreisgebiet (ausgenommen der Städte Unna und Lünen, 2. Halbjahr 2022) begutachtet werden. Die wahrgenommenen Eigenschaften wurden auf zwei analogen Formularen notiert. Im Innendienst wurden die handschriftlichen Vermerke in eine Datenbank übertragen. In fast 50 % der Fälle wurde auf den Formularen im Außendienst vermerkt, dass im Innendienst Nachkontrollen über das Luftbild oder die Schrägluftbilder nötig sind, da z.B. eine Hausseite nicht einsehbar war und somit kein Wert für Dachgauben, einen Balkon oder ähnliches geschätzt werden konnten. Die sachbearbeitenden Personen wurden dementsprechend mit Außen- und Innendienstarbeiten über mehrere Wochen zeitlich sehr stark eingebunden. Zudem entstanden mehrere Medienbrüche (analog ↔ digital) bei der Übertragung von Daten und Notizen. Diese führen zu einer höheren Fehleranfälligkeit.

Zu Beginn wurde eine Ablaufs- und Anforderungsanalyse durch das Geodatenmanagement durchgeführt. Die Ziele der Überführung sind die Eliminierung dieser Medienbrüche und Optimierung der Arbeitsabläufe, also eine Digitalisierung.

Durch die daraus gewonnen Erkenntnisse konnte das Geodatenmanagement in Zusammenarbeit mit dem Fachdienst 16 (Zentrale Datenverarbeitung) den sachbearbeitenden Personen Werkzeuge (QGIS und weitere digitale Geodaten) bereitstellen.

Die Hilfsmittel werden von den sachbearbeitenden Personen genutzt, um einen stetigen Prozess zu etablieren. Die Verstetigung führt dazu, dass sich die hohe Arbeitsaufkommen nicht auf wenige Wochen im Jahr konzentriert, sondern in den Arbeitsalltag integriert werden kann und sich über das ganze Jahr verteilt. Gleichzeitig gibt es keine Medienbrüche mehr, wodurch weniger Arbeitsaufwand durch kleinteilige Arbeitsschritte und mehrere Beteiligte entsteht. Außerdem sinkt die Anfälligkeit für Übertragungsfehler aus Abschriften.

Die Standorte der zu begutachtenden Häuser werden über ein Symbol in einer sich stetig aktualisierenden Karte visualisiert. Über wenige Klicks lässt sich das entsprechende Gebäude in Luftbildern, Schrägluftbildern, Apple Look Around (ähnlich zu Google Street View) und Google Street View betrachten. Über Messwerkzeuge lassen sich Teilflächen oder z.B. Giebelhöhen bestimmen.

Über die Karte können Informationen zum Bearbeitungsstand angepasst werden, ähnlich zu einem Ak-

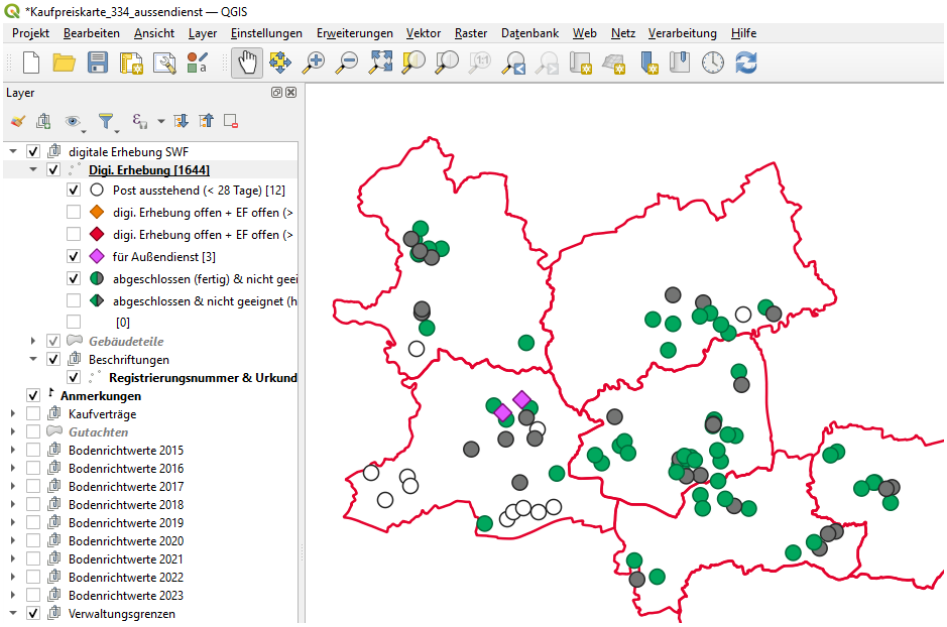


Abb 1:Übersicht nach Bearbeitungsstand über Kauffälle in QGIS

tenvermerk. Wird beispielsweise vermerkt, dass das Gebäude aus fachlichen Gründen doch „ungeeignet“ ist, wird das Symbol ausgegraut. Damit ist eindeutig, dass dort keine weitere Arbeit nötig ist. Falls die digitale Ansicht keine Erkenntnisse bringt, kann der Standort des Gebäudes entsprechend für den Außendienst markiert werden. Diese Häuser werden vierteljährlich innerhalb eines Tages begutachtet.

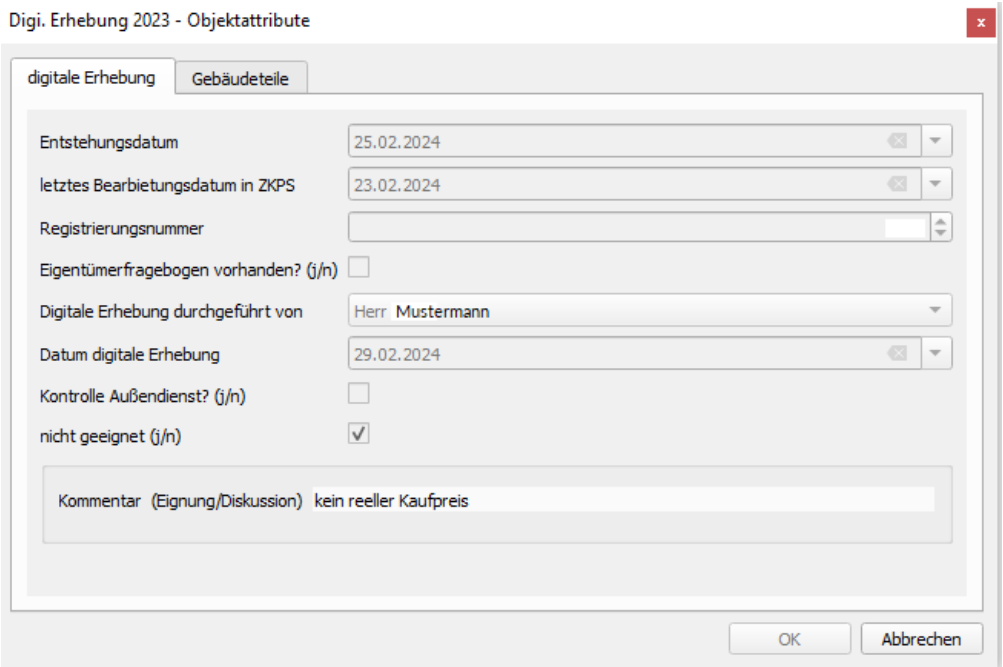


Abb 2: Formular zum Festhalten des Bearbeitungsstandes

Über diese Realisierung ist eine Integration in die alltägliche Arbeit gelungen. Dadurch wird die Arbeitsbelastung besser verteilt. Trotzdem wird weiterhin eine hohe oder sogar verbesserte Güte der erfassten Daten gewährleistet, denn Medienbrüche werden vermieden. Des Weiteren lassen sich die erhobenen Daten zum Prozessablauf auswerten.
Weitere Einsichten in die Daten können durch Auswertungen gewonnen werden, die von der abgeschotteten Statistikstelle vorgenommen werden können. Diese ermöglichen Weiterentwicklungen des Prozesses oder andere fachliche Maßnahmen, hin zu einem verbesserten Arbeitsablauf, fachlich besseren Ergebnissen oder nützlichen Daten.

Kontakt zum Autor:

Marius Schäfer
Kreis Unna – Fachbereich 62, Geoinformation und Kataster
59425 Zechenstraße 51
0 23 03 27-42 62
marius.schaefer@kreis-unna.de

GeoStyler - ArcGIS Unterstützung und weitere neue Features

Jan Suleiman, Benjamin Gerber

Dieser Vortrag zeigt die aktuellsten Neuerungen des OSGeo Community Projekts GeoStyler. Hier wird besonders auf die Unterstützung von ArcGIS Stilen eingegangen, und wie diese in offene Formate wie QGIS Styles, SLD, etc. konvertiert werden können. Zusätzlich zeigen wir die Weiterentwicklungen am GeoStyler-CLI, und was sonst noch u.a. im letzten GeoStyler Code Sprint umgesetzt wurde.

Der Weg aus einem Vendor Lock-in ist besonders im Bereich der Kartographie oft nicht leicht. Abgesehen von neuen Arbeitsabläufen und neuen Benutzeroberflächen, mit denen sich Anwender:innen auseinandersetzen müssen, ist ein Umstieg oftmals mit hohem Aufwand verbunden; bspw. bedingt durch das Nachbilden von Kartenstilen in andere Formate. Genau in diese Lücke tritt GeoStyler und bietet durch die neue Unterstützung von ArcGIS Stilen die Möglichkeit existierende Stile automatisiert in offene Formate wie SLD, QML (QGIS), und weitere zu konvertieren. Dies kann interaktiv über die GeoStyler-UI, oder mittels verschiedener Schnittstellen wie die GeoStyler-CLI oder GeoStyler-Rest geschehen.

In diesem Vortrag zeigen wir den aktuellen Stand des GeoStyler ArcGIS Parsers, sowie weitere Ergebnisse aus dem letzten GeoStyler Code Sprint.

Text und Data Mining in der OpenStreetMap-Datenbank aus rechtlicher Sicht Text

Falk Zscheile

Einleitung

Seit der Nutzung von künstlicher Intelligenz (KI) insbesondere im Zusammenhang mit generativen KI-Modellen hat der Bedarf nach Daten als Rohstoff noch einmal zugenommen. Dabei müssen Datensammlungen jedoch grundsätzlich als rechtlich geschützt angesehen werden. Durch rechtliche Neuregelungen ist in den letzten Jahren rechtlich vieles in Bewegung geraten. Die rechtswissenschaftliche Diskussion hierzu dauert noch an bzw. nimmt – aufgrund der immensen wirtschaftlichen Bedeutung solcher Daten – deutlich zu.

Schutz und Schranken des Schutzes von Datenbanken

Der Schutz von wirtschaftlich wertvollen Informationen im weitesten Sinne (Informationsgüter) durch den Gesetzgeber folgt in der Regel den gleichen Kanon: Das Gesetz definiert was geschützt und welche Rechte damit verbunden sind. Im Anschluss daran werden – trotz des eigentlich bestehenden Schutzes – vom Gesetz Nutzungen definiert, die trotz des Schutzes und ohne Zustimmung der durch den Schutz begünstigten Person zulässig sind. Man spricht hier von den sogenannten gesetzlichen Lizenzen oder Schranken des Schutzes beziehungsweise Schranken des Eigentums. Schließlich haben auch diese Schranken ihren Grenzen. Hier geht es darum, dass die durch vom Gesetz erlaubten Nutzungen ohne oder sogar gegen den Willen der begünstigten Person, nicht ein solches Ausmaß annehmen, dass vom Schutzrecht für die begünstigte Person nichts Essenzielles mehr verbleibt. Um diesen Aspekt entstehen häufig Rechtsstreitigkeiten, weil die Grenzen, insbesondere bei neuen Technologien durch das Gesetz selbst meist nicht klar definiert sind und so durch die Rechtsprechung erst näher bestimmt werden müssen.

Die Entstehung des Datenbankschutzes

Die Europäischen Union (zum damaligen Zeitpunkt noch Europäische Gemeinschaft) erkannte den wirtschaftlichen Wert von Daten sehr früh. Zunächst wurde dieser Mehrwert “nur” in der Datensammlung an sich gesehen. Aus der Überzeugung heraus, dass Datensammlungen von wirtschaftlichem Wert nur dann entstehen, wenn der Hersteller einer solchen Sammlung diese auch rechtssicher verwerten kann, wurde analog zu anderen immateriellen Schutzrecht, wie z. B. dem Urheberrecht oder dem Schutz von Lichtbildern, unter anderem ein neues Schutzrecht für den Datenbankhersteller geschaffen.
Mit dem Erlass der Datenbankrichtlinie 96/9/EG im Jahr 1996 schuf die Europäische Union ein einheitliches Schutzrecht für alle Mitgliedsstaaten. Die Vorgaben der Richtlinie waren bis zum Jahr 1998 umzusetzen. Der deutsche Gesetzgeber kam dieser Aufforderung mit dem Einfügen der §§ 87a–87e ins Gesetz über Urheberrecht und verwandter Schutzrechte (UrhG) nach. Seit dem Wirksamwerden der Reglung sind in der Europäischen Union Datensammlungen bei Vorliegen der Tatbestandsvoraussetzungen durch das Datenbankherstellerrecht (sui-generis-Recht für Datenbanken) geschützt.

Im umfassenden Schutz von wirtschaftlich wertvollen Informationen im weitesten Sinne liegt der Grund für die Entstehung von freien und offenen Lizenzen. Der Schutz von Software durch das Urheberrecht ließ am Ende der 80er Jahre die GNU General Public License entstehen. Am Anfang der 2000er Jahre folgten dann die Creative Commons Lizenzen allgemein für urheberrechtlich geschützte Werke und am Ende der 2000er Jahre die Open Data Commons Open Database License (ODbL) für durch das Datenbankherstellerrecht geschützte Datensammlungen.

Der rechtliche Schutz durch das Gesetz entsteht in diesen Fällen immer automatisch. Das heißt ein entgegenstehender Wille der vom Schutz begünstigten Person ist unerheblich. Wenn man möchte, dass die Allgemeinheit ein rechtlich geschütztes Werk nutzen darf, so ist hier für eine freie oder offene Lizenz erforderlich. Diese räumt, versehen mit bestimmten Bedingungen, den an einer Nutzung interessierten alle hierfür notwendigen Lizenzrechte ein.

Die Lizenzierung der OpenStreetMap-Datenbank unter der ODbL ist also die Konsequenz aus dem automatisch entstehenden Schutz durch das Datenbankherstellerrecht, um dem Wunsch zu entsprechen, dass alle interessierten Kreise die gesammelten Geodaten im Rahmen der Lizenz frei nutzen dürfen.

Anpassungsbedarf durch Digitalisierung der Gesellschaft

Die fortschreitende Digitalisierung verändert die gesellschaftlichen Rahmenbedingungen in bisher nicht gekannter Geschwindigkeit. Der Gesetzgeber ist hier aufgefordert, mit diesen geänderten Rahmenbedingungen Schritt zu halten und den rechtlichen Rahmen zu schaffen. Für den Bereich des Urheberrechts unternahm der europäische Gesetzgeber diesen Versuch mit der Richtlinie über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt (Directive on copyright and related rights in the *D*igital *S*ingle *M*arket, DSM-RL, 2019/79/EU) vom 17. April 2019. Diese Richtlinie war seinerzeit hochumstritten und medial sehr Präsent, unter anderem wegen der darin enthaltenen Regelungen für Upload-Filter zum Aufspüren von urheberrechtlich geschützten aber nicht lizenzierter Inhalte und wegen des Leistungsschutzrechts für Presseverleger.
Daneben wurde weitere zunächst unscheinbare und völlig im medialen Schatten der eben geschilderten Regelungen stehende Änderungen vorgenommen. In Art. 14 wurde die Gemeinfreiheit für Werke der bildenden Kunst geschaffen (§ 68 UrhG)[10], eine Regelung, die auch große Bedeutung für die Digitalisierung historischer Landkarten hat. In Art. 3 erblickte eine neue Schranke des Urheberrechts das Licht der Welt, die hier näher zu besprechende Regelung des Text und Data Minings. Die Vorgaben der DSM-Richtlinie waren bis 2021 im nationalen Recht umzusetzen.

Text und Data Mining als neue Schranke

Die Umsetzung der von der DSM-Richtlinie neue geschaffenen Schranke für Text und Data Mining erfolgte für das Urheberrecht in § 44b UrhG, für die wissenschaftliche Forschung in § 60d UrhG und für Datenbanken in § 87c UrhG. Für Datenbanken wird im Wesentlichen auf die §§ 44b, 60d UrhG verwiesen.
Durch § 87b Abs. 1 UrhG werden Datenbanken durch das Datenbankherstellerrecht geschützt. Der gewährte Schutz sorgt dafür, dass eine entsprechende Datenbank grundsätzlich nur mit der Zustimmung des Datenbankherstellers durch Dritte genutzt werden darf (Lizenzeinräumung). Dies gilt nicht, soweit eine gesetzliche Schranke eingreift. Die Schranken des Datenbankherstellerrechts sind in § 87c UrhG gesetzlich geregelt.

Mit der seit dem 07.06.2021 geltenden Fassung des § 87c Abs. 1 UrhG finden sich dort neu die Nr. 4 “Text und Data Mining gemäß § 44b” und Nr. 5 “Text und Data Mining für Zwecke der wissenschaftlichen Forschung gemäß § 60d”.

In § 44 Abs. 1 UrhG kann man aufgrund des Verweises durch § 87c Abs. 1 Nr. 4 UrhG dorthin nachlesen, was Text und Data Mining nach Auffassung des Gesetzgebers (Definition) ist: Die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken bzw. Datenbanken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen. Mit anderen Worten: Wird ausschließlich zu diesen Zwecken eine Datenbank genutzt, so darf der Inhaber der Datenbank hierfür keine Lizenzgebühren verlangen! Diese Erlaubnis zum Text und Data Mining gestattet auch die Anfertigung einer vorübergehenden Kopie. Die Kopie ist also nach Abschluss des Text und Data Minings wieder zu löschen. Wie bemerkenswert diese Regelung ist, wird deutlich, wenn man sich

vergegenwärtigt, dass außerhalb dieser Schranke des Datenbankherstellerrechts Kopien einer Datenbank nicht ohne Erlaubnis (Lizenz) der begünstigten Person erfolgen dürfen, vgl. § 87b Abs. 1 UrhG.

Bei der Lektüre dieses Wortlautes wird klar: Unter die Gewinnung von “Informationen über Muster” kann man genau das verstehen, was die Grundlage für (generative) Künstliche Intelligenz bildet – neuronale Netze, die auf Basis bestehender Informationen angelernt und Muster abgeleitet bzw. gebildet haben.

Es wird deutlich, dass diese Schranke rückblickend – im Jahr 2021 war deren Bedeutung vermutlich den wenigsten klar – durch die von Schutzrechten begünstigten Personen inzwischen als wesentliche Beeinträchtigung ihrer Vermögensinteressen angesehen wird(Hahn 2023). Diese können – ihre Informationsgüter im Bereich der Text und Data Mining-Schranke nicht im gewohnten Umfang durch Lizenzen monetarisieren[5, S. 979].

Es ist vor dem Hintergrund der aktuellen Bedeutung von Large Language Modellen wenig überraschend, dass sich in der Rechtswissenschaft ein Streit darüber entwickelt hat, ob Schranke des Text und Data Minings tatsächlich auch für das Anlernen von generativen KI-Modellen gilt oder ob hierfür tatsächlich Lizenzgebühren zu entrichten wären, weil die Schranke diesen Fall nicht abdeckt. Dementsprechend wurden von der Contentindustrie auch bereits erste (Gefälligkeits-)Gutachten[1] veröffentlicht[So der Vorwurf durch 2]. Demgegenüber hält ein erstes Gerichtsurteil[4] auch die Fälle des Anlernens Künstlicher Intelligenz einschließlich generativer KI-Modelle (als Obiter Dictum) für gedeckt. Das Urteil ist auch in der rechtswissenschaftlichen Diskussion auf Zustimmung gestoßen[2, 3, 4 S. 170, 8, S. 907, 6, S. 795, ablehnend jetzt 7, 114 f.]. Im Wesentlichen stehen sich in der Rechtswissenschaft die Verteidiger der hergebrachten Rechte der Contentindustrie bzw. geschützten Personen und die am Wortlaut und dem gesetzgeberischen Willen orientierten Auslegungen gegenüber.

Einschränkungen des Text und Data Minings

Wer die Schranke des Text und Data Minings für sich in Anspruch nehmen möchte, sollte dabei jedoch die gesetzlichen Rahmenbedingungen beachten. Es ist vorgesehen, dass die vom Schutzrecht begünstige Person der Auswertung ihrer Datenbank zu Zwecken des Text und Data Minings zu widersprechen kann, § 44b Abs. 3 UrhG i. V. m. § 87 Abs. 1 Nr. 4 UrhG. Dabei ist ein Nutzungsvorbehalt bei online zugänglichen Werken nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt. Dies könnte beispielsweise in der robot.txt oder dem Text and Data Mining Reservation Protocol (TDMRep)(World Wide Web Consortium (W3C) 2024) der Webseite erfolgen. Das LG Hamburg (2024) ist allerdings auch der Meinung, dass ein in Textform hinterlegter Widerspruch in den allgemeinen Geschäftsbedingungen auf der Webseite diese Voraussetzungen erfüllt[4, Rn. 61, zweifelnd insoweit 2].

Wichtig ist außerdem, zu beachten, dass nach erfolgreichem Text und Data Mining die eventuell vorhandenen Vervielfältigungen der Datenbank gelöscht werden müssen, § 44b Abs. 2 Satz 2 UrhG i. V. m. § 87 Abs. 1 Nr. 4 UrhG.

Hier wird auch deutlich, wo die Grenzen des Text und Data Minings verlaufen. Nach der Durchführung dürfen nur Muster, Trends und Korrelationen zurückbleiben. Alles, was einer Kopie einer Datenbank ähnelt oder daran erinnert, ist kein zulässiges Text und Data Mining, sondern die Anfertigung eines lizenzpflichtigen Vervielfältigungsstücks (Kopie). Die genauen Grenzen zwischen Kopie auf der einen Seite und Mustern, Trends und Korrelationen auf der anderen, wird sich erst nach und nach in der Rechtsprechung herausbilden. Möglicherweise

wird man hier auf Regeln zurückgreifen, die sich bereits im Zusammenhang mit der Abgrenzung einer Freien von einer unfreien Bearbeitung in der Rechtsprechung etabliert hat, zurückgreifen, vgl. § 23 Abs. 1 UrhG.

OpenStreetMap-Datenbank und Open Database License

Der Open Database License gewährt für die Nutzung der OpenStreetMap-Daten alle Rechte, welche für eine sinnvolle Nutzung erforderlich sind. Die Lizenzbedingungen gewähren also in der Regel das Recht zur Vervielfältigung, Verbreitung, öffentlichen Zugänglichmachung und zur Bearbeitung an jede Person, die das möchte, vgl. Ziff. 3.1 ODbL 1.0.

Open Database License und gesetzliche Schranken

Das oben vorgestellte gesetzliche Recht zum Text und Data Mining, § 87c Abs. 1 Nr. 4 i. V. m. § 44b Abs. 1 UrhG geht vertraglichen Lizenzbedingungen vor. Es ist quasi eine gesetzlich gewährte Lizenz. Wie oben bereits erläutert kann die vom Schutzrecht begünstigte Person einer Nutzung zu Zwecken des Text und Data Minings widersprechen.

Es stellt sich daher die Frage, wie sich die Open Database License zu der neu in der Europäischen Union geschaffenen Schranken des Text und Data Minings verhält.

Schrankenregelungen zu gewährten Ausschließlichkeitsrechten kennt jede Rechtsordnung in der einen oder anderen Form. In Deutschland dienen solche Schranken unter anderem der Sicherstellung der Sozialbindung des Eigentums, vgl. Art. 14 Abs. 2 GG. Dementsprechend wurde dieser Fall auch in den Lizenzbedingungen der ODbL berücksichtigt.

Die Bestimmungen hierzu finden sich in Ziff. 6 ODbL. Als Grundaussage lässt es sich wie folgt zusammenfassen: Die von Gesetzgeber und der Rechtsprechung bestimmten Grenzen des Datenbankherstellerrechts bleiben von den Lizenzbestimmungen unberührt.

In Ziff. 6.1 a ODbL und Ziff. 6.2 sind die klassischen Schranken der Datenbankrichtlinie ausdrücklich erwähnt. Das Text und Data Mining befindet sich natürlich nicht in dieser Aufzählung. Es wurde erst nach der Veröffentlichung der ODbL geschaffen.

Damit kann als Zwischenergebnis festgehalten werden, dass die Schranke des Text und Data Minings durch die ODbL unangetastet bleibt. Dies führt jedoch zu einer weiteren Frage: Dass Gesetz sieht vor, dass die vom Schutzrecht begünstigte Person dem Text und Data Mining widersprechen kann, § 44b Abs. 3 UrhG i. V. m. § 87 Abs. 1 Nr. 4 UrhG. Wäre es also möglich, dass die OpenStreetMap Foundation (OSMF) als Lizenzgeberin dem Text und Data Mining widerspricht?

Open Database License und Widerspruchsmöglichkeit gegen Text und Data Mining

Die juristische Herausforderung an dieser Frage besteht darin, dass die ODbL gesetzliche Schranken einerseits unangetastet lässt. Als Besonderheit beim Text und Data Mining sieht der Gesetzgeber hier aber eine Widerspruchsmöglichkeit vor. Die begünstigte Person hat also vom Gesetzgeber das Recht bekommen, durch einseitige Erklärung die gesetzliche Schranke zu modifizieren. Aus der grundsätzlichen Zulässigkeit des Text und Data Minings kann die begünstigte Person ein Verbot des Text und Data Minings für die konkrete Datenbank herbeiführen.

Meiner Einschätzung nach wäre ein solcher Widerspruch jedoch mit den von der ODbL gewährten Grundfreiheiten inkompatibel. Es wäre juristisch nicht schlüssig, dass die Lizenz einerseits das kopieren und bearbeiten der Datenbank gestattet, andererseits ein Text und Data Mining untersagen würde. Alle interessierten könnten Text und Data Mining unter Berufung auf das Bearbeitungsrecht durchführen.

Wichtigster Unterschied hierbei ist, dass Ergebnisse des Text und Data Minings nicht der share alike Klausel, wie sie sich in Ziff. 4.6 ODbL findet, unterfallen, Analysen auf Basis des Bearbeitungsrechts hingegen schon. Diese müssten also wiederum unter den Bedingungen der Ziff. 4.6 ODbL bereitgestellt werden.

Die gesetzliche Schranke des Text und Data Mining führt nach der hier vertretenen Auffassung für die gewonnenen Muster, Trends und Korrelationen nicht auf Basis der ODbL lizenziert werden müssen. Diesen Fall kennt die ODbL auch für sogenannte “Produced Work” (Produkt). Auch diese fallen nicht unter die Bedingungen der Ziff. 4.6 ODbL, sondern müssen lediglich einen Hinweis auf die Datenquelle entsprechend Ziff. 4.3 ODbL enthalten. Im Gegensatz hierzu ist für die Erkenntnisse aus dem Text und Data Mining die Angabe der Herkunft der Erkenntnisse nicht erforderlich. Dies ergibt sich aus der Tatsache, dass die gesetzliche Schranke des Text und Data Mining eine solche Anforderung nicht aufstellt.

Im Ergebnis steht damit fest: Die OSMF kann einem Text und Data Mining nicht widersprechen. Text und Data Mining in der OpenStreetMap-Datenbank ist zulässig. Eine Quellenangabe ist rechtlich nicht erforderlich, kann sich aber gegebenenfalls unabhängig davon unter dem Aspekt guter wissenschaftlicher Praxis ergeben.

Chancen für OpenStreetMap

Nach der hier vertretenen Auffassung ist die OpenStreetMap Foundation nicht berechtigt, dem Text und Data Mining zu widersprechen. Es bleibt damit bei der gesetzlichen Grundregel, dass Text und Data Mining auf der OpenStreetMap-Datenbank zulässig sind.
In der Vergangenheit war es für OpenStreetMap eine große juristische Herausforderung den eigenen Datenbestand mit anderen geografischen Informationssammlungen zu vergleichen. Grund hierfür waren Lizenzinkompatibilitäten. Immer Stand die Gefahr, durch die Datenanalyse in unerlaubter weise eine abgeleitete Datenbank herzustellen. Dieser Aspekt war als (unsichtbarer) Elefant regelmäßig mit im Zimmer. Mit der gesetzlichen Schranke des Text und Data Mining dürfte sich die Rechtslage ein Stück weit entspannt haben. Immer wenn es darum geht, Muster, Trends und Korrelationen zwischen verschiedenen geografischen Daten zu analysieren ist dies im Rahmen dieser gesetzlichen Schranke möglich. Es muss nicht zwingen um die Gewinnung von Mustern für generative KI Modelle gehen. Wichtig dabei ist zu beachten, dass alle für die Analyse herangezogenen Datensätze dem Text und Data Mining zugänglich sein müssen. Datensätze bei denen die berechtigte Person dem Text und Data Mining widersprochen hat, sind einer Nutzung nicht zugänglich. Außerdem dürfen die gewonnenen Muster, Trends und Korrelationen nicht als (bearbeitete) Kopie der Originaldatensätze erscheinen. Kopien sind nur zum Durchführen der Analyse zulässig und müssen danach sofort wieder gelöscht werden. Nur die gewonnenen Erkenntnisse (Muster, Trends und Korrelationen) sind einer Weiternutzung ohne Einschränkungen zugänglich.

Kontakt zum Autor:

Falk Zscheile
c/o FOSSGIS e.V.
Bundesallee 23
10717 Berlin
falk.zscheile@fossgis.de

Literatur

[1] Tim W. Dornis und Sebastian Stober. Urheberrecht und Training generativer KI-Modelle. Technologische und juristische Grundlagen. Nomos Verlag, 2024.

[2] Thomas Hoeren. „Anmerkung zu LG Hamburg, Urteil vom 27.9.2024, 310 O 227/23“. In: MMR (2024), S. 978.

[3] Franz Hofmann. „Retten Schranken Geschäftsmodelle generativer KI-Systeme?“ In: ZUM (2024), S. 166–174.

[4] LG Hamburg. Urteil vom 27. Sep. 2024. Az. 310 O 227/23, ECLI:DE:LGHH:2024:0927.310O227.23.00.

[5] Christoph Matras und Patrick Grosmann. „Anmerkung zu LG Hamburg, Urteil vom 27.9.2024, 310 O 227/23“. In: MMR (2024).

[6] Benjamin Raue. „Die Freistellung von Datenanalysen durch die neuen Text und Data Mining-Schranken (§§ URHG § 44b, URHG § 60d UrhG)“. In: ZUM (2021), S. 793–802.

[7] Haimo Schack. „Auslesen von Webseiten zu KI-Trainingszwecken als Urheberrechtsverletzung de lege lata et ferenda“. In: NJW (2024), S. 113–117.

[8] Haimo Schack. „Schutzgegenstand, „Ausnahmen oder Beschränkungen“des Urhe-

berrechts“. In: GRUR (2021), S. 904–909.

[9] World Wide Web Consortium (W3C). TDM Reservation Protocol (TDMRep). 10. Mai 2024. url: https://www.w3.org/community/reports/tdmrep/CG-FINAL-tdmrep-20240510/ (besucht am 08. 02. 2025).

[10] Falk Zscheile. Gemeinfreiheit visueller Werke. Chemnitz, 12. März 2023. url: https://media.ccc.de/v/clt23-257-gemeinfreiheit-visueller-werke (besucht am 24. 06. 2024).

State of GeoServer Cloud

Daniel Koch, Nils Bühner



GeoServer ist ein vielseitiger Kartenserver, der in Microservice-Architekturen an seine Grenzen stößt. GeoServer Cloud als eigenes Projekt "on top" transformiert den GeoServer in skalierbare Einzelkomponenten für Container-Umgebungen wie Kubernetes.

Der GeoServer ist ein etablierter und extrem vielseitiger Kartenserver, der alle gängigen OGC-Standards sowie unzählige Datenformate unterstützt. Im Kontext von modernen Microservice-Architekturen, die eine Skalierung in Cloud-Umgebungen ermöglichen, gerät der GeoServer als monolithische Software jedoch an seine technische Grenzen. Vor diesem Hintergrund ist das GeoServer Cloud Projekt entstanden, welches

die bestehenden GeoServer-Implementierungen durch Erweiterung und Anpassung in individuell einsetzbare Einzelkomponenten transformiert. Diese können dann im Rahmen einer Microservice-Architektur in Container-basierten Verwaltungsumgebungen wie beispielsweise Kubernetes eingesetzt und beliebig skaliert werden.

Der Vortrag beleuchtet im Allgemeinen die technischen Konzepte sowie entscheidenden Unterschiede zum "Vanilla GeoServer" (wie z.B. das Datenverzeichnis/Katalog), diskutiert Vor- und Nachteile dieser Entwicklungen, stellt die wichtigsten Neuerungen aus dem Projekt vor und liefert somit wertvolle Entscheidungshilfen für die Wahl einer geeigneten GeoServer-Variante.

Neues von actinia - mit Knative weiter wolkenwärts

Markus Neteler



"Hallo, mein Name ist actinia. Ich bin eine REST API für GRASS GIS, die die Verwaltung und Visualisierung von Projects, Mapsets und Geodaten sowie die Ausführung von GRASS GIS Modulen ermöglicht und in einer Cloud-Umgebung installiert werden kann. Den aktuellen Trends folgend wird gezeigt, wie ich im ESA Projekt KNeo angepasst werde, um im Open Source Projekt Knative, einer Erweiterung für Kubernetes, installiert werden zu können. Um mehr Details zu erfahren, kommen Sie vorbei!"

"Hallo, mein Name ist actinia. Ich wurde schon häufiger auf der FOSSGIS vorgestellt, also vielleicht bin ich bereits bekannt. Falls nicht - ich bin eine REST API für GRASS GIS, die die Verwaltung und Visualisierung von Locations/Projects, Mapsets und Geodaten sowie die Ausführung der vielen GRASS GIS Module und Addons ermöglicht. Andere Tools wie GDAL werden ebenfalls unterstützt.

Ich kann in einer Cloud-Umgebung installiert werden und helfe dabei, große Mengen an Geoinformationen aufzubereiten, zu analysieren und bereitzustellen. Da ich ursprünglich auch für den Einsatz in einer Cloud-Umgebung entwickelt wurde, ist es naheliegend, mich nach den aktuellen Trends zu richten

und ständig weiter zu entwickeln. Das geschieht derzeit im Rahmen des ESA-geförderten Projektes KNeo. Hier werde ich so angepasst, dass ich mit Knative installiert werden kann. Knative ist ein von der Cloud Native Computing Foundation (CNCF) verwaltetes Open Source Community Projekt und eine Erweiterung der Container-Orchestrierungsplattform Kubernetes. Es ermöglicht, serverlose Workloads auf Kubernetes-Clustern auszuführen und Tools und Dienstprogramme bereitzustellen, die die Erstellung, Bereitstellung und das Management von containerisierten Anwendungen innerhalb von Kubernetes einfacher und „Kubernetes-nativer“ machen. Im Gegensatz zu früheren Serverless-Frameworks wurde Knative für die Bereitstellung aller modernen App-Workloads entwickelt - von monolithischen Apps bis hin zu Microservices und kleinen Funktionen.

In diesem Vortrag werden zunächst die Grundlagen von mir und Knative behandelt und anschließend gezeigt, wie ich für die Nutzung mit Knative angepasst wurde. Schlagwörter wie "cloud-native", "serverless", "event-triggered" und viele mehr werden nach diesem Vortrag keine Fremdwörter mehr sein. Um mehr Details zu erfahren, kommen Sie vorbei!"

Stand des GRASS GIS Projekts: Neuigkeiten und Perspektiven

Markus Neteler



Wir geben einen Überblick über die neuesten Entwicklungen im GRASS GIS Projekt mit Fokus auf die Version 8.4.0: Der Begriff "Location" ist nun einfach "Projekt", die Support-Vektor-Machine Methode zur Bildklassifikation ist in GRASS-core verfügbar und optionale JSON Ausgabe neu in verschiedenen Tools. Verbesserungen in der grafischen Benutzeroberfläche und neue Werkzeuge für Jupyter Notebooks machen GRASS GIS zugänglicher und skalierbarer für Anwendungen von der Wissenschaft bis zur Industrie.

Im Vortrag geben wir einen Überblick über die Fortschritte des GRASS GIS Projekts und die Highlights der Version 8.4.0, die mit über 520 Verbesserungen und Korrekturen veröffentlicht wurde. Eine bemerkenswerte Änderung ist die Ersetzung des traditionellen (aber oft verwirrenden) Begriffs "Location" durch "Project" in der graphischen Benutzeroberfläche, bei der Kommandozeile und in der gesamten Python API, wobei die Abwärtskompatibilität gewährleistet bleibt. Dies wird die Datenhierarchie für die Benutzer:innen intuitiver machen. Die Version führt mehrere neue Werkzeuge und Verbesserungen ein, darunter eine GRASS-Kernfunktion zur Bildklassifikation mit Hilfe von Support Vector Machines (i.svm). Im Bereich der Datenverarbeitung und -ausgabe unterstützt GRASS GIS 8.4.0 nun das JSON Ausgabeformat für verschiedene Werkzeuge und führt eine neue Klasse für Jupyter Notebooks zur Animation von Vektor- oder Rasterzeitreihen ein. Zusätzlich ermöglicht die Integration von ipyleaflet eine einfachere Kartenvisualisierung in Jupyter Notebooks. Zu den Verbesserungen der Benutzeroberfläche gehören ein verbessertes Single-Window-GUI mit abdockbarer Kartenansicht und ein neues History-Browser-Panel, das die Navigation durch die Workflow-Historie vereinfacht. Ein neuer GRASS GIS Programming Style Guide erleichtert die Orientierung während der Entwicklung.

Wir stellen die Roadmap für zukünftige GRASS GIS Versionen vor, die einen beschleunigten Releasezyklus und neue Werkzeuge beinhaltet. Insgesamt verbessern diese Neuerungen die Fähigkeiten von GRASS GIS in den Bereichen Geodatenanalyse, Datenvisualisierung und Benutzerinteraktion erheblich und demonstrieren die kontinuierliche Weiterentwicklung des Projekts, um den Bedürfnissen sowohl von Anfänger:innen als auch von erfahrenen Profis in der Geo-Community gerecht zu werden.

Fehlende Bürgersteige

Dr. Roland Olbricht

Die Fähigkeit zum Fußgängerrouting gilt als ein nahezu Alleinstellungsmerkmal von OpenStreetMap. Doch tatsächlich haben über 50% aller Hauptstraßen in Deutschland keine Information zu Bürgersteigen irgendeiner Art.

Zeit zu Mappen! Ich habe in meiner Heimatstadt Wuppertal angefangen und setze die Arbeit in den Nachbarstädten fort. Welche Werkzeuge und Tagging-Stile ich versucht und welche Werkzeuge und Tagging-Style ich nachher genutzt habe, ist Gegenstand dieses Vortrages.

Die Fähigkeit zum Fußgängerrouting gilt als ein nahezu Alleinstellungsmerkmal von OpenStreetMap. Diskussionen wie um Bürgersteig-Generierung und Flächenrouting vermitteln, dass die Datenversorgung bereits sehr gut sei. Und tatsächlich sind auch viele Routing-Engines und Apps aktiv, die die OpenStreetMap-Daten tatsächlich nutzen. Mein persönliches Highlight dafür ist die Zusammenarbeit mit der Dortmunder Community anlässlich des Kirchentags 2019 gewesen.

Doch tatsächlich haben über 50% aller Hauptstraßen in Deutschland keine Information zu Bürgersteigen (weder sidewalk- noch foot-Tag noch einen parallelen Fuß-Way in den Daten).

Zeit, das Problem ganz praktisch durch Mappen anzugehen. Ich habe in meiner Heimatstadt Wuppertal angefangen und werde mich dann durch die Nachbarstädte arbeiten. Welche Werkzeuge und Tagging-Stile ich vorab in Betracht gezogen habe und welche Werkzeuge und Tagging-Style ich nachher genutzt haben, ist Gegenstand dieses Vortrages.

Ableitung von Rasterkarten aus Vector Tiles für basemap.de

Sebastian Ratjens

Für die Produktion der Rasterkarten von basemap.de wird die Open-Source-Software "VT Raster Converter" entwickelt, mit der Vector-Tiles-Karten in Rasterbilder konvertiert werden können. Der Vortrag gibt einen Überblick über die Software und deren Nutzung in Verbindung mit dem MapProxy. Es werden die Erfahrungen und Herausforderungen in der Produktion der basemap.de Rasterkarten mittels Vektor-Raster-Konvertierung erläutert.

Der Kartendienst basemap.de stellt amtliche Karten u. a. in Form von Vector Tiles bereit. Da nicht alle Kartenanwendungen in der Lage sind Vector Tiles einzubinden, werden auch Rasterkarten-Dienste von basemap.de angeboten. Um die Pflege der Style-Konfigurationen zu vereinfachen und das Kartenbild der Vektor- und Rasterkarten zu vereinheitlichen, werden die Rasterbilder aus den Vector Tiles erzeugt. Dafür erfolgt eine Konvertierung mit der auf MapLibre Native basierenden Software "VT Raster Converter", die für die Produktion der basemap.de Rasterkarten entwickelt wurde. Für die Speicherung der konvertierten Rasterbilder wird der MapProxy verwendet.

Der Vortrag gibt einen Überblick über die Software "VT Raster Converter" und deren Nutzung in Verbindung mit dem MapProxy. Es werden die Erfahrungen und Herausforderungen in der Produktion der basemap.de Rasterkarten mittels Vektor-Raster-Konvertierung erläutert.

Die Software "VT Raster Converter" ist unter der Open-Source-Lizenz MIT veröffentlicht (https://github.com/Smart-Mapping/vt-raster-converter). Entwickelt werden die Dienste und Anwendungen von basemap.de in Zusammenarbeit der Vermessungsverwaltungen der Länder (AdV) und dem Bundesamt für Kartographie und Geodäsie (BKG).

QGIS und Blender für beeindruckende 3D-Darstellungen

Gordon Schlolaut

Blender ist eine freie Software für professionelle Visual Effects und Animationen. In Kombination mit QGIS können diese auch dafür verwendet werden, um Geo-Daten in 3D effektvoll darzustellen und zu animieren. Obwohl Blender sehr komplex ist, können sehr schnell und mit wenigen Schritten erste hervorragende Ergebnisse erzielt werden. Es wird ein Einstieg in den Workflow gegeben und die einzelnen relevanten Komponenten von Blender an mehreren Beispielen erklärt.

Der Elefant kann's auch allein: Graph-Erstellung aus OSM in der PostGIS-Datenbank

Matthias Daues

Vom osm-dump zum voll vernetzten Graphen: Mit osmium, osm2pgsql und einigen simplen Datenbank-Prozeduren gelingt die Umwandlung von rein geographischen Informationen in logische Datenstrukturen.

Im GitHub findest Du alles, was Du zum selber machen brauchst.

Matthias Daues (datenschoenheit) öffnet den Postgresql Werkzeugkasten und zeigt, wie man in vier Prozesschritten aus einem OSM-Dump sowohl ein logisches (also: Graph-Datenmodell) als auch ein geometrisches (also: Geo-Datenmodell) Netzwerk erzeugen und ohne Umwege für Analysen und Berechnungen zur Verfügung stellen kann:

Datenbeschaffung: In der shell oder mit Python

Vorprozessierung und DB-Beladung: Grobe Filterung mit Osmium, Modellierung und Beladung mit osm2pgsql

Graph-Erzeugung: PostGIS-Funktionen, SQL, und Python in der DB

Anwendung: Routing und Graph-Suche

Zum Vortrag gibt es ein Git-Repository mit allen erforderlichen Dockerfiles, Skripten und docker-compose Anweisungen, zum direkt Nachmachen oder Weiterentwickeln.

radinfra.de – gute Daten und Kampagnen zur Radinfrastruktur für ganz Deutschland

Tobias Jordans

OpenStreetMap bietet die einzigen flächendeckend verfügbaren und standardisierten Radinfrastrukturdaten in Deutschland. Um diese Daten für die Öffentlichkeit sowie für die planende Verwaltung zugänglich zu machen und systematisch zu verbessern, wurde radinfra.de entwickelt.

Die neu veröffentlichte radinfra.de bietet aufbereitete Daten, Statistiken und Kampagnen zur Radinfrastruktur für ganz Deutschland. Dafür werden täglich Daten aus OpenStreetMap ausgewertet und aufbereitet.

Daten zur Radinfrastruktur liegen in Deutschland sehr fragmentiert vor. Amtliche Daten sind nicht standardisiert, veraltet, ungenau oder gar nicht vorhanden. OSM bietet systematische Daten zur Radinfrastruktur, die fortlaufend aktuell gehalten werden können. Bisher wird dieses Potenzial nur ansatzweise genutzt, denn die Daten in OSM sind nicht einfach zugänglich und bedürfen weiterer Pflege und Detaillierung.

Unter radinfra.de ist jetzt ein neuer Anlaufpunkt für alle entstanden, die sich für offene Radinfrastrukturdaten interessieren und mithelfen wollen, diese zu verbessern.

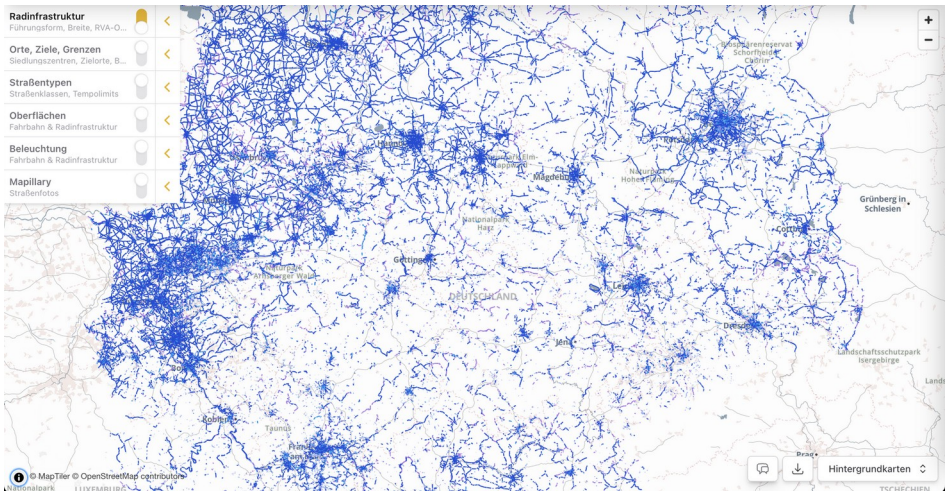
Einheitliche Daten für ganz Deutschland:

Dank OpenStreetMap können für ganz Deutschland einheitliche Daten produziert werden, die sowohl beschilderte als auch unbeschilderte Radinfrastruktur abbilden. Dabei werden die Rohdaten aus OpenStreetMap aufbereitet, vereinheitlicht und nach Führungsform klassifiziert, um eine einfache Nutzung für Planungszwecke zu ermöglichen.

Täglich aktuelle Daten:

Eine Stärke von OpenStreetMap ist ihr schneller Aktualisierungszyklus. Änderungen sind schon am nächsten Tag in den Daten zu sehen.

Kampagnen zur Datenpflege:



Neben dem einfachen Zugang zu Radinfrastruktur-Daten ist die Datenpflege ein zweites Ziel des Projekts. Dafür werden aus den Daten verschiedene Kampagnen zur Aktualität, Vollständigkeit und Richtigkeit der Daten abgeleitet. Ziel ist es, die Daten immer weiter zu verbessern und Attribute wie Breiten und Oberflächen systematisch zu erfassen.

Das Projekt ist Open Source und eine Kooperation der OSM-Verkehrswende-Community mit der FixMyCity GmbH. Die Prozessierung der Daten basiert auf dem Open Source-Projekt Radverkehrsatlas, das FixMyCity ebenfalls betreibt.

PMTiles - das cloud-native Format für Kacheln

Pirmin Kalberer

PMTiles ist ein neues Format zur effizienten Speicherung und einfachen Publikation von Vektor- und Rasterkacheln.

PMTiles ist ein cloud-natives Format, welches die Veröffentlichung von Tiles ohne spezielle Serverapplikation ermöglicht. Es wird nur ein HTTP-Server benötigt, um Vektor- oder Rasterkacheln auszuliefern.

Ein kompletter OpenStreetMap Vektor-Tiles Datensatz mit den Zoomstufen 0-15 kann in einem ca. 150 GB grossen PMTiles-Archiv gespeichert werden. Inzwischen wir die direkte Generierung von PMTiles von mehreren Tools unterstützt.

Dieser Vortrag zeigt die Vor- und Nachteile von PMTiles, wie man sie erzeugt und wie sie mit verschiedenen Karten-Viewern verwendet werden können.

Airborne-Laserscanning in Deutschland: Verfügbarkeiten, Herausforderungen, Potenziale

Jens Wiesehahn

Airborne-Laserscanning (ALS)-Daten werden in Deutschland großflächig durch die Bundesländer erhoben. Trotz einer Entwicklung hin zu mehr Open Access bestehen weiterhin Hürden aufgrund uneinheitlicher Datenqualität, Verfügbarkeit und Zugangsregelungen. Der Vortrag gibt einen Überblick über die Verfügbarkeit von ALS-Daten in Deutschland, zeigt bestehende Probleme bei der Datennutzung auf und diskutiert Strategien für eine verbesserte Bereitstellung.

Airborne-Laserscanning (ALS)-Daten sind für zahlreiche Anwendungsbereiche in Forschung und Entwicklung von großem Wert, da sie detaillierte Informationen zur dreidimensionalen Struktur von Landschaften liefern. Aus diesen Daten lassen sich hochpräzise Gelände- und Oberflächenmodelle erstellen, die beispielsweise zur Berechnung von Wasserabflüssen oder zur Schätzung der Holzbiomasse genutzt werden. Die Punktwolken liefern weitere Informationen, etwa wie die vertikale Bestandesstruktur zwischen Baumkronen und Boden aufbaut ist.

In Deutschland werden flächendeckende ALS-Kampagnen von den einzelnen Bundesländern und neuerdings auch vom Bund durchgeführt, jedoch ist die großflächige Nutzung dieser Daten nach wie vor mit Hürden verbunden. Während der Zugang in der Vergangenheit stark eingeschränkt war, zeigt sich zwar zunehmend ein Wandel hin zu mehr Open Access, die föderale Struktur führt allerdings weiterhin zu heterogenen Datenqualitäten, Verfügbarkeiten und Zugangsregelungen.

Die Präsentation führt in die Welt der Lidar-Daten ein: Wie werden sie erfasst, welche Standards existieren, und mit welchen Tools lassen sich die Daten verarbeiten? Der Schwerpunkt liegt auf der aktuellen Verfügbarkeit von ALS-Daten in Deutschland. Dazu wurde eine umfassende Recherche in Geoportalen, Social Media, persönlichem Mailverkehr etc. durchgeführt. Die Verfügbarkeit und Zugänglichkeit in den einzelnen Bundesländern wird dargestellt und im internationalen Kontext betrachtet. Es wird aufgezeigt, welche spezifischen Daten vorhanden sind, wie sie bereitgestellt werden und unter welchen Lizenzen sie stehen. Möglichkeiten und Herausforderungen bei der Arbeit mit ALS-Daten werden aus der Perspektive der Nutzenden beleuchtet, und es werden offene Fragen sowie potenzielle Strategien für eine verbesserte Datenbereitstellung diskutiert.

OpenSource sicher entwickeln und betreiben – Prozesse, Anforderung und Tools im Fokus

Florian Micklich

Open Source ist mehr als nur Software – sicher wird sie jedoch erst durch gezielte Entwicklungs- und Betriebsprozesse. Der Vortrag zeigt, wie Open Source-Projekte sicher gestaltet und betrieben werden können. Mit „Shifting Left“ und Automatisierung werden Sicherheitslücken frühzeitig erkannt, während Standards wie BSI-Grundschutz und der Cyber Resilience Act (CRA) wichtige Rahmenbedingungen setzen. Praxisnahe Einblicke in SAST, DAST und Supply-Chain-Security runden das Thema ab.

Open Source Software ist mehr als nur reine Software, doch das Thema Cybersecurity erfordert gezielte Maßnahmen in der Entwicklung und im Betrieb – wie bei jeder anderen Software auch. In diesem Praxisbericht zeige ich anhand konkreter Beispiele, wie man Open Source-Projekte sicher entwickeln und nachhaltig betreiben kann. Der Bericht beleuchtet bewährte Methoden und Open-Source-Tools, die dazu beitragen, Open Source-Projekte robuster und vertrauenswürdiger zu machen.

Ein zentraler Fokus liegt auf dem „Shifting Left“-Ansatz, der Sicherheitsprüfungen frühzeitig in die Entwicklungsphase integriert. Mithilfe von Automatisierungen wie statischer Code-Analyse (SAST), dynamischen Sicherheitstests (DAST), Dependency Scanning und Supply-Chain-Security können Sicherheitsrisiken bereits während der Entwicklung reduziert werden.

Der BSI-Grundschutz wird dabei mit seinen wichtigsten Bausteinen für die Umsetzung und den Betrieb von Open Source-Projekten vorgestellt. Jeder Baustein liefert praxisnahe Anforderungen und Empfehlungen, die dabei helfen, typische Sicherheitslücken systematisch zu schließen und so die Betriebssicherheit langfristig zu gewährleisten.

Abgerundet wird der Praxisbericht durch den in drei Jahren in Kraft tretenden Cyber Resilience Act (CRA) und die Herausforderungen, die auf Open Source-Projekte zukommen. Dieser neue Rechtsrahmen fordert eine intensive Vorbereitung und Zusammenarbeit, um die zukünftigen Anforderungen zu erfüllen.

Zum Abschluss lade ich zum Austausch und zur Diskussion ein: Gemeinsam erörtern wir, wie sich Open Source-Projekte sicher umsetzen und betreiben lassen, um den Anforderungen der Zukunft gerecht zu werden.

PTNA: Qualitätssicherung für ÖPNV-Linien in OpenStreetMap

Toni Erdmann

Abstract

„PTNA - Public Transport Network Analysis“ [1] ermöglicht eine SOLL-IST-Analyse für ÖPNV-Linien in OpenStreetMap. Soll-Daten bestehen aus einer CSV-Liste von ÖPNV-Linien, jeweils mindestens aus der Liniennummer und dem Fahrzeugtyp. Diese Informationen werden in einer IST-Analyse mit den OSM-Daten verglichen. Zusätzlich erfolgt eine Fehleranalyse der einzelnen ÖPNV-Relationen in OSM.

PTNA sollte auf der FOSSGIS 2020 in Freiburg vorgestellt werden. Wegen Corona gibt es lediglich einen Beitrag im Tagungsband [2]. Dieser Beitrag stellt die wichtigsten Erweiterungen an PTNA seit März 2020 vor: eine Analyse von GTFS-Daten, sowie einen Vergleich von OSM-Daten mit GTFS-Daten. „GTFS“ steht für: "General Transit Feed Specification", „GTFS is a community-driven open standard for rider-facing transit information.“ [3][4][5].

Anwendung von PTNA

PTNA wird seit Februar 2017 entwickelt um einen **SOLL-IST-Vergleich** im Rahmen von **OSMs ÖPNV-Linien** zu realisieren. Die folgenden 4 Anwendungsfälle sind derzeit möglich:

1.) IST-Analyse = Bestandsaufnahme

- ✔ Definition einer Area in dem die ÖPNV-Linen gesucht werden sollen
- ✔ Einfache Auflistung der gefundenen *route_master* und *route* Relationen nach Typ des Fahrzeuges und Linien-Nummer bzw. -Bezeichnung
- ✔ Dedizierte Fehleranalyse der Relationen basierend auf der „Public Transport Version 2“ („PTv2“) -Spezification [7]

2.) IST-Analyse = Bestandsaufnahme für einen oder mehrere ‚network‘-Werte

- ✚ Filtern nach ‚*network*‘-Werten, z.B.:
  - „`Münchner Verkehrs- und Tarifverbund`“ und „`MVV`“

3.) SOLL-IST-Analyse anhand einer Liste real existierender Linien

- ✚ Erstellung und Wartung einer SOLL-Liste (CSV-Format) im OSM-Wiki
- ✚ Lesen der CSV-Liste vom OSM-Wiki, z.B. ein Eintrag:
  - `210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH`
- ✚ Fehlermeldung: Linien in der CSV-Liste sind nicht in OSM gemapped
- ✚ Auflistung nicht in der CSV-Liste stehender aber in OSM existierender Linien

- Einfache Strukturierung des Report-Layout (Abschnitte, Text, …)
- Erweiterte Fehleranalyse von *route_master* und *route* Relationen
  - Passen *tags* des *route_masters* zu denen der zugehörigen *routen*

4.) SOLL-IST-Analyse anhand von Vorgaben aus GTFS-*feed*

- Erstellung der CSV-Liste im OSM-Wiki durch Export der GTFS-*routen*
- Wartung der CSV-Liste im OSM-Wiki mit Hilfe der GTFS-Daten
- Erweiterung des CSV-Eintrags um GTFS-spezifische Elemente
  - *feed*“ = Name des Datensatzes
  - „*route_id*“ = ID der Linie im GTFS-Datensatz
  - „*release_date*“ = optionales Element: die Version des Datensatzes
    - `210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH;DE-BY-MVV;mvv-19-210-1;`
- Analyse, Prüfung und Verlinkung der GTFS-Informationen im PTNA-Report
  - GTFS-Links auf die Analyse von GTFS-*route* bzw. GTFS-*trip*
  - Vergleichs-Icons zum Start eines Vergleichs von
    - GTFS-*route* mit OSM-*route_master*
    - GTFS-*trip* mit OSM-*route*

Über die Einstiegsseite von PTNA [1] gelangt man via Klick auf eine Flagge bei „**Take me to the results …**“ auf eine Übersicht der verfügbaren Reports für das entsprechende Land. Die wichtigsten Spalten in der Tabelle:

- Die **1. Spalte**, „Name“ führt zum **Report** der Analysen
- Die **3.-letzte Spalte**, „Letzte Änderungen“ führt zu einem Report bei der zusätzlich die **Veränderungen seit der letzten Analyse** markiert sind
- Die **letzte Spalte**, „Linien“ führt zu den **SOLL-Daten**, der CSV-Liste im OSM-Wiki

Im Folgenden wird im Wesentlichen der zuletzt genannte, 4.) Anwendungsfall beschrieben, bei dem ein Vergleich von GTFS mit OSM möglich ist. Wo keine GTFS-Daten zur Verfügung stehen kommt Anwendungsfall 3.) zum Einsatz.

Die Vorbereitung der GTFS-Daten in PTNA wird erläutert. Die Begriffe aus GTFS und OSM werden anschließend kurz gegenübergestellt.

GTFS-Daten

Import der GTFS-Daten

GTFS-Daten müssen bezüglich ihrer Lizenz mit OSM kompatibel sein. CC0 [16] ist ideal, für CC-BY-4.0 [17] gilt das zum Beispiel nur, wenn die Form der „Namensnennung“ konkreter genannt wird. Für einige GTFS-Daten liegen Angaben seitens der Eigentümer vor, die die Namensnennung im Verzeichnis der Beitragenden im OSM-Wiki [6] als ausreichend bestätigen.

PTNA prüft GTFS-Quellen regelmäßig auf neue Versionen der Daten. Es wird in der Regel nur die erste im Monat verfügbare Version importiert. Auf Anfrage kann ein Import angestoßen werden.

Aufbereitung der GTFS-Daten

Beim Import der GTFS-Daten wird aus den diversen CSV-Dateien des Datensatzes eine sqlite3-Datenbank aufgebaut. Ein zweiter, wichtiger Schritt ist die Reduktion der Datenmenge durch Aggregierung, aus OSM-Sicht, redundanter Informationen. Das ist im Wesentlichen die Zusammenfassung identischer Trips mit der selben Reihenfolge von Haltestellen aber unterschiedlichen Abfahrtzeiten zu einem repräsentativen Trip.

Analyse der GTFS-Daten

Um die Verwendbarkeit der GTFS-Daten für OSM zu bestimmen, werden von PTNA diverse Analysen durchgeführt:

1.) Ist ein Trip eine Teilroute eines anderen Trips bezüglich der Haltestellen

2.) Besteht ein Trip nur aus 2 Haltestellen (Ausnahme: Seilbahnen, Fähren, ...)

3.) Haben die erste und zweite Haltestelle den selben Namen, die selbe *stop_id*

4.) Haben die vorletzte und letzte Haltestelle den selben Namen, die selbe *stop_id*

5.) Unterscheiden sich zwei Trips nur durch die *shape_id* der Fahrstrecke

6.) Unterscheiden sich zwei Trips nur durch die *stop_id* der Haltestellen

7.) Beträgt die Fahrzeit eines Trips 0 Sekunden

Die Ergebnisse dieser Prüfungen werden als „PTNA Kommentar“ zum jeweiligen GTFS-*trip* ausgegeben. Diese Kommentare sollen als Hilfe beim Mappen dienen, wenn entschieden werden soll, ob ein GTFS-*trip* auch in den OSM-Daten gemapped werden soll. Gerade Punkt 6.) ist häufig bei GTFS-Daten für Züge zu sehen, die in Bahnhöfen mal das eine und mal ein anderes Gleis nutzen: *stop_name* identisch, *stop_id* unterschiedlich.

Normalisierung von Namen im GTFS-Datensatz

In GTFS-Datensätzen sind die Namen von Haltestellen häufig abgekürzt. In OSM hingegen gilt die Regel: Namen werden ausgeschrieben. Dieser Unterschied kann spaäter beim Vergleich von GTFS- mit OSM-Daten zu Fehlermeldungen führen. PTNA führt daher für den deutschsprachigen Raum eine Normalisierung/Expansion von Namen für die GTFS-Daten durch, z.B.:

- „Friedenstr.“
  - de_DE und de_AT: „Friedenstraße“
  - de_CH: „Friedenstrasse“
- „W.-Heisenberg-W.“
  - „Werner-Heisenberg-Weg“
- „Bruck Abzw.“
  - de_DE und de_CH: „Bruck Abzweig“
  - de_AT: „Bruck Abzweigung“

Die Eigentümer der GTFS-Datensätze sind beim Abkürzen sehr kreativ: PTNA kann bezüglich der Normalisierung der Namen nie vollständig sein.

Darstellung der GTFS-Daten auf der PTNA Webseite

Über die Einstiegsseite von PTNA [1] gelangt man via Klick auf eine Flagge hinter „**Take me to the GTFS …**“ auf eine Übersicht der **verfügbaren GTFS-Daten** für das entsprechende Land. Die wichtigste Spalte in der Tabelle:

- Die **1. Spalte**, „Name“ führt zur Übersicht der GTFS-*routen* des GTFS-*feeds*

Auf der folgenden Übersichtsseite werden die **GTFS-*routen* des GTFS-*feeds*** aufgelistet, sortiert nach Fahrzeugtyp und Linien-Nummer bzw. -Bezeichnung. Die wichtigste Spalte in der Tabelle:

- Die **1. Spalte**, „Linie“ führt zur Übersicht der GTFS-*trips* der GTFS-*route*

Auf der nun folgende Übersichtsseite werden die **GTFS-*trips* der GTFS-*route*** im Abschnitt „Existierenden Linienvarianten“ aufgelistet, sortiert nach den Namen der Haltestellen (1., letzte, 2., 3., …, letzte). Diese Art der Sortierung hält Linienvarianten zusammen, die identische Start- und End-Haltestellen haben, dazwischen aber Unterschiede aufweisen. Die wichtigste Spalte in der Tabelle:

- Die 2**. Spalte**, „Trip-ID“ führt zur Analyse des GTFS-*trips*

Auf der finalen Übersichtsseite erfolgt die **Analyse des GTFS-*trips*** mit Karte, Taggingvorschlägen für OSM, Haltestellentabelle, Verkehrszeiten und, wenn vorhanden, dem Fahrweg (*shape*).

Begriffe: GTFS versus OSM

GTFS und OSM nutzen unterschiedliche Begriffe für Objekte des ÖPNV. Tabelle 1 zeigt einen Überblick über die von PTNA verwendeten Elemente und deren Entsprechungen. Für OSM gilt dabei die so genannte „Public Transport Version 2“ („PTv2“) [7].

GTFS Begriff	OSM Begriff	Anmerkung
feed	*network:guid*	Viele GTFS-Datensätze werden von den Verkehrsverbünden herausgegeben. PTNA vergibt generische *feed*-Namen [8], die in OSM als *network:guid* verwendet werden können: - *network* = Münchner Verkehrs-und Tarifverbund - *network:short* = MVV - *network:guid* = DE-BY-MVV
route	*route_master*	Fasst alle Fahrten einer ÖPNV-Linie zusammen.
route_id		Eindeutige ID einer *route.*
route_short_name	*ref*	Das OSM tag *ref* einer *route_master*/*route* Relation.
route_color	*colour*	Die Hintergrundfarbe der Linienbezeichnung.
route_text_color	*colour:text*	Die Textfarbe der Linienbezeichnung.
trip	*route*	Ein GTFS-*trip* entspricht einer einzelnen Fahrt zu einer bestimmten Uhrzeit, definiert durch die angefahrenen Stops und optionaler Fahrstrecke (*shape)*. Eine OSM-*route* beschreibt alle Fahrten, die die selben Stops anfahren plus deren Fahrstrecke.
trip_id		Eindeutige ID eines *trips.*

GTFS Begriff	OSM Begriff (PTv2)	Anmerkung
shape		In GTFS: die eigentliche Fahrstrecke (optional). In OSM ist das die Kette der *highway*/*railway*/... member einer *route* Relation (ohne ‚*role*‘).
shape_id		Eindeutige ID eines *shapes.*
stop	*public_transport=platform*	Ein *stop* entspricht der neben dem Fahrweg liegenden *platform*.
stop_id	*ref:IFOPT, ...*	Eindeutige ID eines *stops*, nicht immer identisch mit *ref:IFOPT*, ...
stop_lat	*lat*	Breitengrad der Haltestelle.
stop_lon	*lon*	Längengrad der Haltestelle.
stop_name	*name*	Name einer Haltestelle.
	public_transport=stop_position	Den Begriff „Halteposition des Fahrzeugs“ gibt es in GTFS nicht.
agency_name	*operator*	Bei manchen GTFS-Daten kann der Name der *agency* als OSM-*operator* einer *route_master*/*route* Relation genutzt werden.

Tabelle 1: Vergleich GTFS vs. OSM Begriffe

Für den Vergleich von GTFS- und OSM-Daten können die aufbereiteten und analysierten GTFS-Daten auf verschiedene Arten genutzt werden.

Vergleich von GTFS-Daten mit OSM-Daten

Nutzung von GTFS-Daten in der CSV-Liste im OSM-Wiki

Die Definition des GTFS-*feed* und der GTFS-*route_id* (**SOLL-Daten**) erfolgt in der **CSV-Liste im OSM-Wiki**.

- Dieser Schritt sollte immer zuerst erfolgen
- Änderungen erfolgen lediglich an Daten im OSM-Wiki, nicht an OSM-Daten
- Die erstmalige Erstellung kann durch PTNA unterstützt werden
  - Export der GTFS-*routen* als CSV-Liste
- Wartung hält sich in Grenzen,
  - wenn die GTFS-*route_id* stabil [5] bleibt und weil GTFS-*trip_ids* nicht enthalten sind
  - Lediglich neue und eingestellte Linien müssen dann eingepflegt werden
- Im PTNA-Report erscheint jeweils ein **GTFS**-Link und ein **Vergleichs**-Icon
- Beispiel für CSV-Daten im OSM-Wiki für den Bus 210 im Münchner Verkehrs- und Tarifverbund:
  - `210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH;`**`DE-BY-MVV;mvv-19-210-1`**

Nutzung von GTFS-Daten in OSM route_master / route Relationen

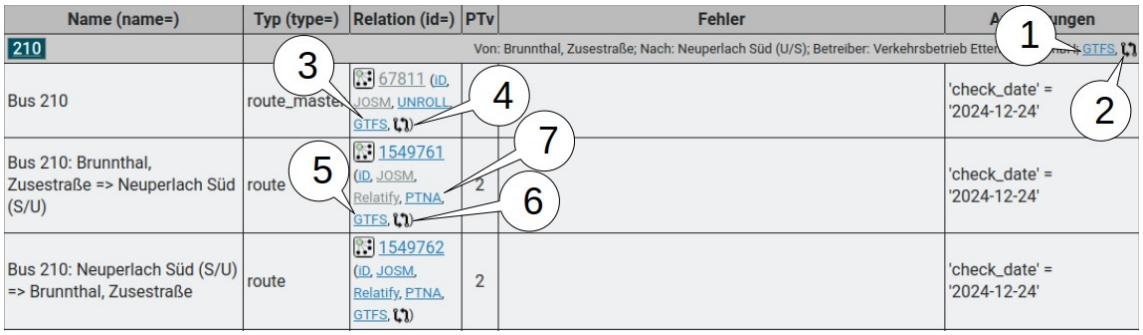
Die Definition des GTFS-*feed*, der GTFS-*route_id* bzw. GTFS-*trip_id* erfolgt im OSM-Datensatz der *route_master*- bzw. *route*-Relation als **IST-Daten** [9].

- Änderungen erfolgen an OSM-Daten
- Aufwändig in der Erstellung
- Gegebenenfalls aufwändig in der Wartung
  - Das hängt von der Stabilität der GTFS-*id*-Daten über mehrere Versionen hinweg ab (*route_id*, *trip_id*, *shape_id*, *stop_id*) [5]
- Erstellung und Wartung kann durch den vorangegangenen Abschnitt (GTFS-Daten in CSV-Liste im OSM-Wiki) und den dadurch möglichen Vergleich erleichtert werden
- Im PTNA-Report erscheint für jede so getaggte Relation ein **GTFS**-Link und ein **Vergleichs**-Icon
- Beispiel für die tags in einer OSM-*route*-Relation des Busses 210:
  - `gtfs:feed=DE-BY-MVV`
  - `gtfs:route_id=mvv-19-210-1`
  - `gtfs:trip_id:sample=mvv-19-210-1-1-1-H-0-MoTuWeThFrSaSu-119-344`

Bedeutung von Links im PTNA-Report

Abbildung 1 zeigt einen Auszug aus einem PTNA-Report. Der Auszug korrespondiert mit dem folgenden Eintrag in den CSV-Daten im OSM-Wiki:

```
210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH;DE-BY-MVV;mvv-19-210-1
```



Name (name=)	Typ (type=)	Relation (id=)	PTv	Fehler	Anmerkungen
210				Von: Brunnthal, Zusestraße; Nach: Neuperlach Süd (U/S); Betreiber: Verkehrsbetrieb Etten	GTFS,
Bus 210	route_master	67811 (ID, JOSM, UNROLL, GTFS,)			'check_date' = '2024-12-24'
Bus 210: Brunnthal, Zusestraße => Neuperlach Süd (S/U)	route	1549761 (ID, JOSM, Relatify, PTNA, GTFS,)	2		'check_date' = '2024-12-24'
Bus 210: Neuperlach Süd (S/U) => Brunnthal, Zusestraße	route	1549762 (ID, JOSM, Relatify, PTNA, GTFS,)	2		'check_date' = '2024-12-24'

Abbildung 1: Auszug aus PTNA-Report für den Bus 210

Für die Analyse und den Vergleich von GTFS- und OSM-Daten finden sich diverse Links:

① Ein **GTFS**-Link in der Kopfzeile außen rechts führt zu einer **Analyse** der GTFS-*route* mit einem Überblick über die zugehörigen GTFS-*trips*. Die genutzten GTFS-Daten werden der **CSV-Liste im OSM-Wiki** (= SOLL-Daten) entnommen (Abbildung 2).

Variante	Trip-ID	Anzahl Fahrten	Gültigkeit		Haltestellen			PTNA Info
			Von	Bis	Start	Über	Ende	Kommentar
1	mvv-19-210-1-1-1-H-0-MoTuWeThFrSaSu-119-344	7037 (7325)	2025-01-13	2025-03-30	Brunnthal, Zusestraße	Taufk., W.-Messerschmitt-Str. => Taufkirchen, Lilienthalstr. => Ottobrunn, Einsteinstraße => Ottobrunn, Röntgenstraße => Ottobrunn, Robert-Koch-Str. => Ottobrunn, Ottostraße => Ottobrunn, Hubertusstraße => Ottobrunn, Ortsmitte => Ottobrunn, Jahnstraße => Ottobrunn, Bahnhofstraße => Ottobrunn, Alte Landstraße => Neubiberg, W.-Heisenberg-W.	Neuperlach Süd	
2	mvv-19-210-1-1-4-R-0-Su-1-1	6896	2025-01-13	2025-03-30	Neuperlach Süd	Neubiberg, W.-Heisenberg-W. => Ottobrunn, Alte Landstraße => Ottobrunn, Bahnhofstraße => Ottobrunn, Jahnstraße => Ottobrunn, Ortsmitte => Ottobrunn, Hubertusstraße => Ottobrunn, Ottostraße => Ottobrunn, Finkenstraße => Ottobrunn, Robert-Koch-Str. => Ottobrunn, Röntgenstraße => Ottobrunn, Einsteinstraße => Taufkirchen, Lilienthalstr. => Taufk., W.-Messerschmitt-Str.	Brunnthal, Zusestraße	
3	mvv-19-210-1-1-7-R-1-MoTuWeThFr#2-225-268	624	2025-01-13	2025-03-28	Neuperlach Süd	Neubiberg, W.-Heisenberg-W. => Ottobrunn, Alte Landstraße => Ottobrunn, Bahnhofstraße => Ottobrunn, Jahnstraße	Ottobrunn, Ortsmitte	
4	mvv-19-210-1-1-2-H-1-MoTuWeThFr#2-219-447	288	2025-01-13	2025-03-28	Ottobrunn, Jahnstraße	Ottobrunn, Bahnhofstraße => Ottobrunn, Alte Landstraße => Neubiberg, W.-Heisenberg-W.	Neuperlach Süd	Fahrt ist Teilroute von: mvv-19-210-1-1-1-H-0-MoTuWeThFrSaSu-119-344

Abbildung 2: Übersicht über die *trips* einer GTFS-*route*

② Ein **Vergleichs**-Icon in der Kopfzeile außen rechts führt zu einem Vergleich der GTFS-*route* mit dem OSM-*route_master*. Die genutzten GTFS-Daten werden der **CSV-Liste im OSM-Wiki** (= SOLL-Daten) entnommen (Abbildung 7 im folgen Abschnitt: „Vergleich von GTFS-*route*- mit OSM-*route_master*-Daten“).

③ Ein **GTFS**-Link in der dritten Spalte einer OSM-*route_maste*r-Relation führt zu einer **Analyse** der GTFS-*route* mit einem Überblick über die zugehörigen GTFS-*trips*. Die genutzten GTFS-Daten werden den ***gtfs:**** **Angaben der *route_master*-Relation** (= IST-Daten) entnommen (siehe Abbildung 2 oben).

④ Ein **Vergleichs**-Icon in der dritten Spalte einer OSM-*route_maste*r-Relation führt zu einem Vergleich der GTFS-*route* mit dem OSM-*route_master*. Die genutzten GTFS-Daten werden den ***gtfs:**** **Angaben der *route_master*-Relation** (= IST-Daten) entnommen (siehe Abbildung 7 im folgenden Abschnitt: „Vergleich von GTFS-*route*- mit OSM-*route_master*-Daten“).

⑤ Ein **GTFS**-Link in der dritten Spalte einer OSM-*route*-Relation führt zu einer **Analyse** des GTFS-*trips*, mit Karte und Details (scrollen) zum Trip. Die genutzten GTFS-Daten werden den ***gtfs:**** **Angaben der *route*-Relation** (= IST-Daten) entnommen (Abbildung 4).

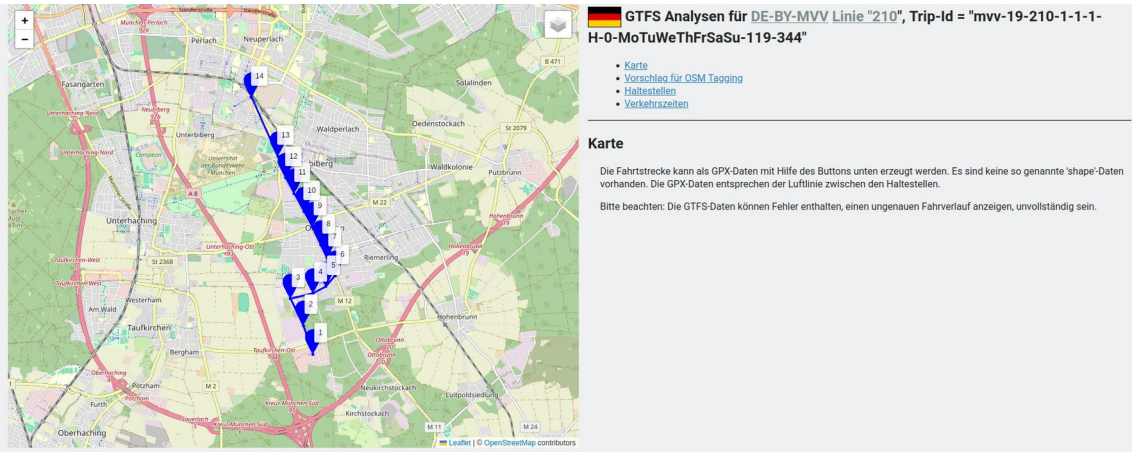


Abbildung 4: Analyse eines GTFS-*trips* durch PTNA

⑥ Ein **Vergleichs**-Icon in der dritten Spalte einer OSM-*route*-Relation führt zu einem Gegenüberstellung von GTFS-*trip* und OSM-*route*. Die genutzten GTFS-Daten werden den ***gtfs:**** **Angaben der *route*-Relation** (= IST-Daten) entnommen (siehe Abbildungen 8 und 9 im folgenden Abschnitt: „Vergleich von GTFS-*trip*- mit OSM-*route*-Daten“).

⑦ Ein **PTNA**-Link in der dritten Spalte führt zu einer **Analyse** der OSM-*route*-Relation, mit Karte und Details der einzelnen Elemente der Relation (Abbildung 3).

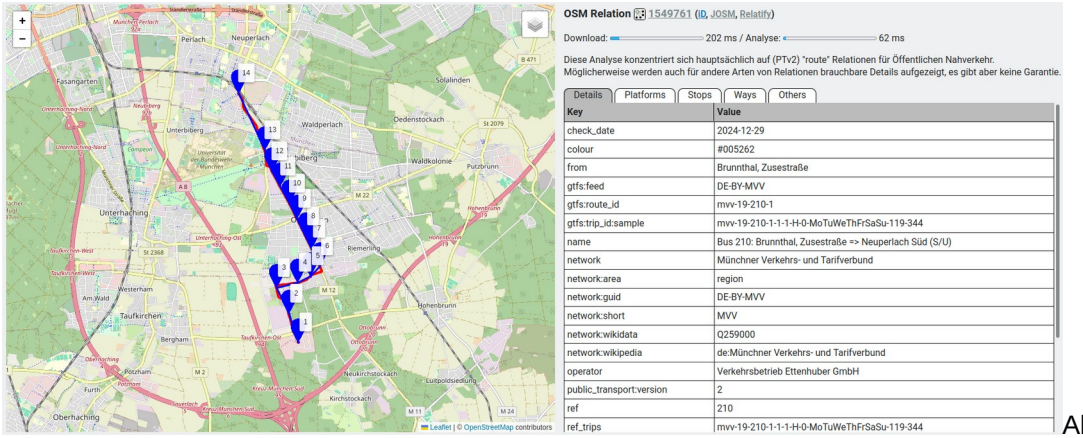


Abbildung 3: Analyse einer OSM-*route*-Relation durch PTNA

Abbildung 4 zeigt eine OSM-*route*-Relation, bei der die Reihenfolge der Haltestellen nicht korrekt erscheint: Haltestellen (PTv2-*platform*) 3 und 4 sind vertauscht.

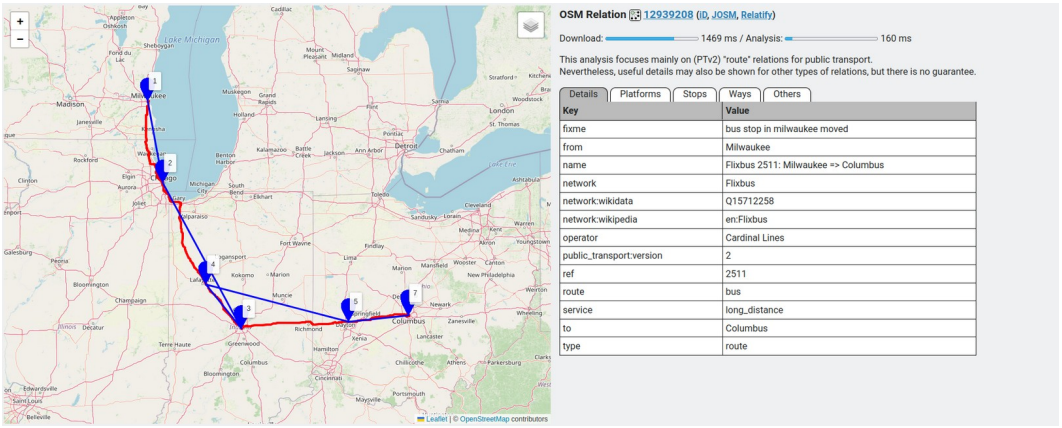


Abbildung 4: Analyse einer fehlerhaften OSM-*route*-Relation durch PTNA

Visueller Vergleich

Hierbei erfolgt der Vergleich eines GTFS-*trip* und einer OSM-*route* auf getrennten Karten

- Keinerlei Änderungen an bestehenden OSM-Daten notwendig
- Mühsam in der Anwendung.

Die Links zu **GTFS** (Link 5) und **PTNA** (Link 7) im PTNA-Report führen für GTFS-*trip* und OSM-*route* zu identisch aufgebauten Karten auf den Webseiten.

Ohne GTFS-Link im PTNA-Report muss die Analyse über die Länderspezifische GTFS-Übersicht, den GTFS-*feed*, die GTFS-*route* und den passenden GTFS-*trip* gesucht werden.

Ein visueller Vergleich kann mittels zweier Browserfenster erfolgen (Abbildung 5 und 6 zeigen Ausschnitte der Fenster).

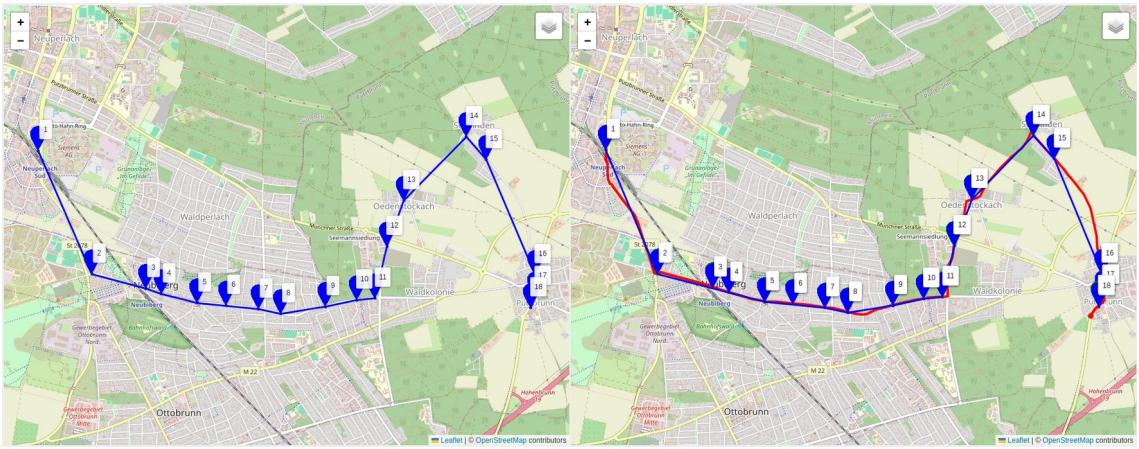
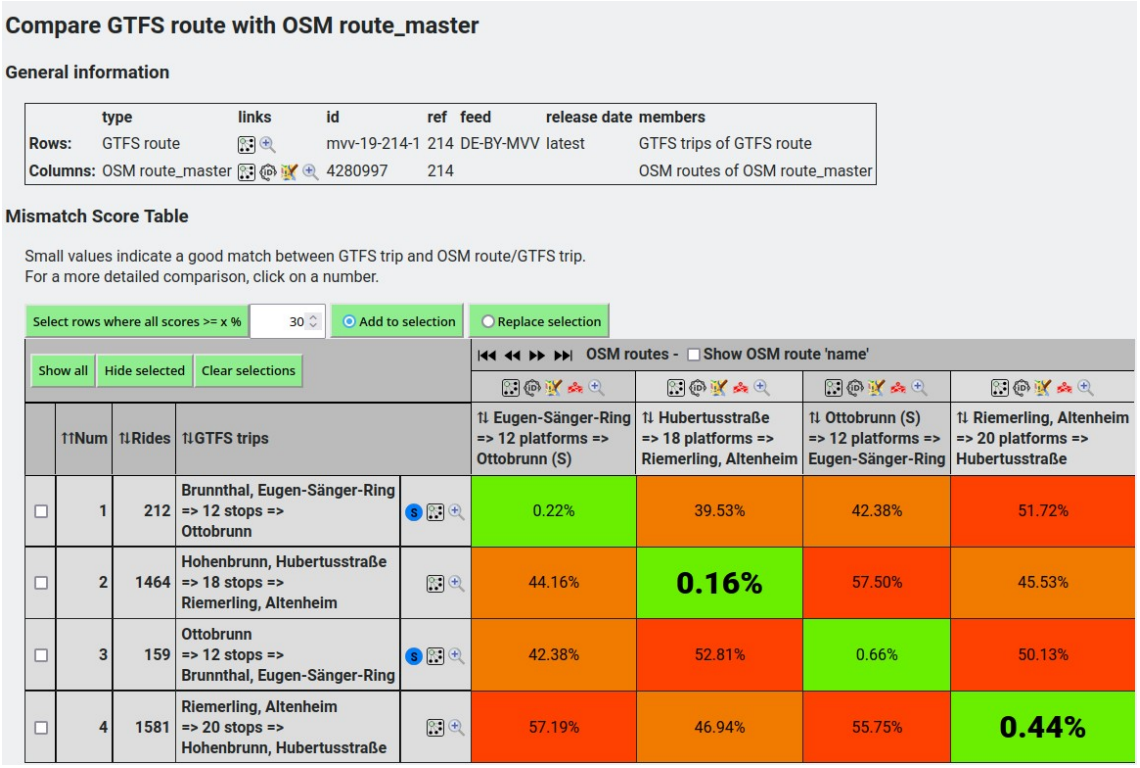


Abbildung 5: GTFS-*trip* auf Karte (ohne *shape*) Abbildung 6: OSM-*route* auf Karte (mit *shape*)

Vergleich von GTFS-route- mit OSM-route_master-Daten

Der Link bei dem **Vergleichs-Icon** für eine **OSM-*route_master***-Relation führt zu dem Vergleich einer GTFS-*route* versus OSM-*route_master*. Der Vergleich erfolgt in Tabellenform mit so genannten **score**-Werten für jede Kombination GTFS-*trip*/OSM-*route*: je kleiner der **score**-Wert, desto besser die Übereinstimmung der beiden Datensätze. Ist ein **score**-Wert in fetter Schrift abgebildet, so stimmen *gtfs:feed* und *gtfs:trip_id:sample* in der OSM-*route*-Relation mit der GTFS-*trip_id* überein (Abbildung 7).



Compare GTFS route with OSM route_master

General information

	type	links	id	ref	feed	release date	members
Rows:	GTFS route		mvv-19-214-1	214	DE-BY-MVV	latest	GTFS trips of GTFS route
Columns:	OSM route_master		4280997	214			OSM routes of OSM route_master

Mismatch Score Table

Small values indicate a good match between GTFS trip and OSM route/GTFS trip.
For a more detailed comparison, click on a number.

Select rows where all scores >= x % 30 | Add to selection | Replace selection

Show all | Hide selected | Clear selections

OSM routes - Show OSM route 'name'

	↑↑Num	↑↓Rides	↑↓GTFS trips		↑↓ Eugen-Sänger-Ring => 12 platforms => Ottobrunn (S)	↑↓ Hubertusstraße => 18 platforms => Riemerling, Altenheim	↑↓ Ottobrunn (S) => 12 platforms => Eugen-Sänger-Ring	↑↓ Riemerling, Altenheim => 20 platforms => Hubertusstraße
☐	1	212	Brunnthal, Eugen-Sänger-Ring => 12 stops => Ottobrunn		0.22%	39.53%	42.38%	51.72%
☐	2	1464	Hohenbrunn, Hubertusstraße => 18 stops => Riemerling, Altenheim		44.16%	**0.16%**	57.50%	45.53%
☐	3	159	Ottobrunn => 12 stops => Brunnthal, Eugen-Sänger-Ring		42.38%	52.81%	0.66%	50.13%
☐	4	1581	Riemerling, Altenheim => 20 stops => Hohenbrunn, Hubertusstraße		57.19%	46.94%	55.75%	**0.44%**

Abbildung 7: Vergleich von GTFS-*route* mit OSM-*route_master* mit **score**-Tabelle

Ein Klick auf einen **score**-Wert führt zum Vergleich eines GTFS-*trips* mit einer OSM-*route* (nächster Abschnitt).

Vergleich von GTFS-trip- mit OSM-route-Daten

Der Link bei dem **Vergleichs-Icon** für eine **OSM-*route***-Relation führt zu dem Vergleich GTFS-*trip* versus OSM-*route*. Der Vergleich erfolgt auf einer Karte und in Tabellenform mit den **score**-Werten für einzelne Metriken: je kleiner der **score**-Wert, desto besser die Übereinstimmung.

Die Metriken basieren auf Eigenschaften von Haltestellen: Anzahl, Position, Entfernung, Namen, ID, … Die Shapes spielen keine Rolle, sie sind häufig nicht enthalten, mitunter fehlerhaft und schwer zu vergleichen (häufig: „false-positives“).

Beim Scrollen nach unten wird ein Vergleich der Haltestellen in Tabellenform sichtbar.

Abbildungen 8 und 9 zeigen einen **positiven Fall,** wo die Übereinstimmung sehr gut ist. Lediglich die Positionen der Haltestelle 12 stimmt nicht ganz überein: OSM ist hier richtig.

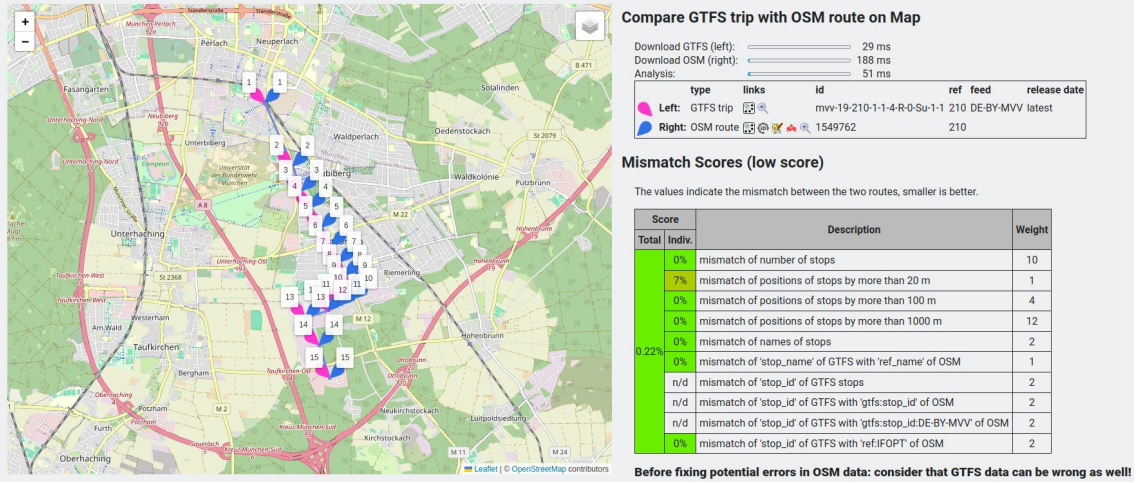


Abbildung 8: Vergleich von GTFS-*trip* mit OSM-*route* auf Karte und **score**-Tabelle: positiv

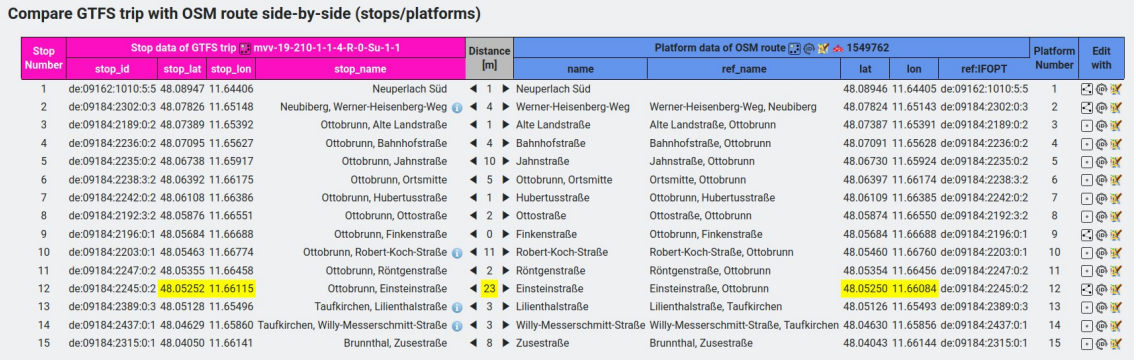


Abbildung 9: Vergleich von GTFS-*trip* mit OSM-*route* mit Haltestellen-Tabelle: positiv

Abbildungen 10 und 11 zeigen einen **negativen Fall,** wo die Übereinstimmung schlecht ist. Der **score** ist 22.06%. Der GTFS-*trip* stellt bezüglich Stops 6 und 7 die Werktagsvariante bis ca. 20:30 dar, die OSM-*route* stellt bezüglich Haltestelle 6 die Abend- und Sonntagsvariante dar (fährt nicht durch das Gewerbegebiet).

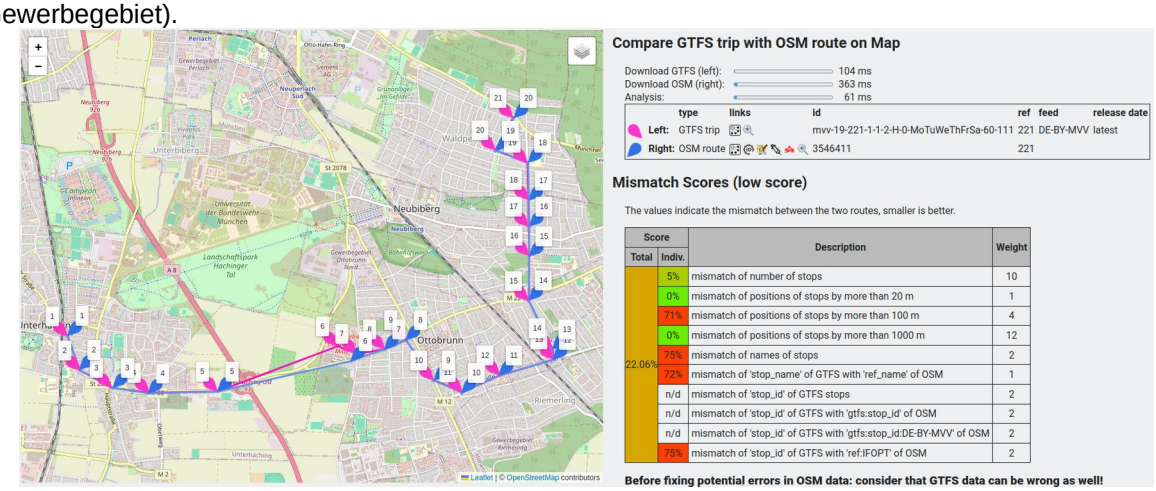


Abbildung 10: Vergleich GTFS-*trip* (werktags) mit OSM-*route* (abends, sonntags): Karte/Scores

Compare GTFS trip with OSM route side-by-side (stops/platforms)

Stop Number	Stop data of GTFS trip mvv-19-221-1-1-2-H-0-MoTuWeThFrSa-60-111				Distance [m]	Platform data of OSM route 3546411				Platform Number	Edit with
	stop_id	stop_lat	stop_lon	stop_name		name	ref_name	lat	lon	ref:IFOPT		
1	de:09184:2310:1:1	48.06520	11.61318	Unterhaching	3	Unterhaching (S)		48.06521	11.61314	de:09184:2310:1:1	1	
2	de:09184:2366:0:1	48.06204	11.61489	Unterhaching, Jahnstraße	11	Jahnstraße		48.06209	11.61477	de:09184:2366:0:1	2	
3	de:09184:2368:0:1	48.06044	11.61926	Unterhaching, Ottobrunner Straße	11	Ottobrunner Straße	Unterhaching, Ottobrunner Straße	48.06040	11.61940	de:09184:2368:0:1	3	
4	de:09184:2378:0:1	48.05995	11.62448	Unterhaching, Friedhof	16	Friedhof	Unterhaching, Friedhof	48.05992	11.62427	de:09184:2378:0:1	4	
5	de:09184:2365:0:1	48.06008	11.63392	Unterhaching, Kirchlandweg	0	Kirchlandweg	Unterhaching, Kirchlandweg	48.06008	11.63392	de:09184:2365:0:1	5	
6	de:09184:2206:0:1	48.06428	11.65058	Ottobrunn, Siemensstraße	204	Daimlerstraße	Daimlerstraße, Ottobrunn	48.06290	11.65239	de:09184:2380:2:3	6	
7	de:09184:2380:1:1	48.06355	11.65324	Ottobrunn, Daimlerstraße	285	Lindenstraße	Lindenstraße, Ottobrunn	48.06396	11.65703	de:09184:2239:0:1	7	
8	de:09184:2239:0:1	48.06399	11.65709	Ottobrunn, Lindenstraße	237	Ottobrunn, Ortsmitte	Ortsmitte, Ottobrunn	48.06484	11.66001	de:09184:2238:2:3	8	
9	de:09184:2238:2:3	48.06484	11.66005	Ottobrunn, Ortsmitte	504	Hubertusstraße	Ottobrunn, Hubertusstraße	48.06109	11.66385	de:09184:2242:0:2	9	
10	de:09184:2242:0:2	48.06108	11.66386	Ottobrunn, Hubertusstraße	313	Friedrich-Ebert-Platz	Friedrich-Ebert-Platz, Ottobrunn	48.05993	11.66770	de:09184:2281:1:1	10	
11	de:09184:2281:1:1	48.06001	11.66786	Ottobrunn, Friedrich-Ebert-Platz	409	Hans-Kandler-Weg	Hans-Kandler-Weg, Ottobrunn	48.06155	11.67285	de:09184:2190:0:1	11	
12	de:09184:2190:0:1	48.06158	11.67303	Ottobrunn, Hans-Kandler-Weg	566	Ottobrunn (S)	Ottobrunn	48.06297	11.68037	de:09184:2210:1:1	12	
13	de:09184:2210:1:1	48.06305	11.68055	Ottobrunn	106	Karl-Stieler-Straße	Karl-Stieler-Straße, Ottobrunn	48.06398	11.68030	de:09184:2188:1:1	13	
14	de:09184:2188:1:1	48.06409	11.68024	Ottobrunn, Karl-Stieler-Straße	545	Am Brunneck/Uhlandstraße	Am Brunneck/Uhlandstraße, Ottobrunn	48.06849	11.67699	de:09184:2187:0:1	14	
15	de:09184:2187:0:1	48.06847	11.67701	Ottobrunn, Am Brunneck/Uhlandstraße	456	Neubiberg, Gymnasium	Gymnasium, Neubiberg	48.07257	11.67713	de:09184:2278:0:1	15	
16	de:09184:2278:0:1	48.07261	11.67709	Neubiberg, Gymnasium	300	Walkürenstraße	Walkürenstraße, Neubiberg	48.07531	11.67708	de:09184:2275:0:1	16	
17	de:09184:2275:0:1	48.07534	11.67705	Neubiberg, Walkürenstraße	266	Brunhildenstraße	Brunhildenstraße, Neubiberg	48.07773	11.67704	de:09184:2290:0:2	17	
18	de:09184:2290:0:2	48.07781	11.67701	Neubiberg, Brunhildenstraße	382	Salzmannstraße	Salzmannstraße, München	48.08125	11.67690	de:09162:968:3:SAL 1	18	
19	de:09162:968:3:SAL 1	48.08126	11.67688	Salzmannstraße	350	Waldperlach	Waldperlach, München	48.08230	11.67243	de:09162:967:1:1	19	
20	de:09162:967:1:1	48.08237	11.67239	Waldperlach	383	Waldheimplatz	Waldheimplatz, München	48.08536	11.67496	de:09162:966:5:6	20	
21	de:09162:966:5:6	48.08538	11.67486	Waldheimplatz	8							

Abbildung 11: Vergleich GTFS-*trip* (werktags) mit OSM-*route* (abends, sonntags): Tabelle

Icons und ihre Bedeutung

Auf den Seiten des PTNA-Reports und der GTFS-/OSM-Vergleiche sind diverse Icons zu sehen. Diese führen entweder zu weiteren Informationen oder laden relevante Daten in OSM-Editoren, wo eine Bearbeitung erfolgen kann.


Das Vergleichs-Icon: ein Klick führt zum Vergleich von GTFS- mit OSM-Daten. Auf der Übersichtsseite für GTFS-*feeds* führt ein Klick zum Vergleich zweier GTFS-*feed*-Versionen.


Lade die OSM Daten des Objektes (*relation*, *way* oder *node*) in den **JOSM**-Editor.


Mit Hilfe des **JOSM**-Editors werden Daten (key, value) in die OSM-*route_master- bzw. route*-Relation injiziert. Mouseover zeigt an, welche Daten das sind.


Lade die OSM-Daten (*relation*, *way* oder *node*) in den **iD**-Editor.


Lade die OSM-*route*-Relation in den **Relatify**-Editor.


Der angezeigte Name (GTFS-*stop_name*) wurde „Normalisiert“.
Mouseover zeigt den Originalnamen.


Zeige eine Kurzinformationen der Objektdaten in einem Pop-Up-Fenster.


Der angezeigte GTFS-*trip* stellt eine Teilroute (sub-route) mindestens eines anderen GTFS-*trips* dar. Mouseover zeigt an, welcher das ist.


Der angezeigte GTFS-*trip* hat verdächtige Merkmale.
Mouseover zeigt an, welche das sind


Der angezeigte GTFS-*trip* hat nur zwei Haltestellen.


Der angezeigte GTFS-*trip* ist bezüglich *stop_name* identisch mit mindestens einem anderen GTFS-*trip* unterscheidet sich aber in *stop_id* bzw. *shape_id*. Mouseover zeigt an, welcher das ist.


Das Objekt ist ein Punkt (OSM: *node*)
Ein Klick auf das Icon bzw. die ID zeigt die Objektdaten auf https://osm.org an.


Das Objekt ist eine Strecke (OSM: *way*)
Ein Klick auf das Icon bzw. die ID zeigt die Objektdaten auf https://osm.org an.


Das Objekt ist eine Sammlung (GTFS: *route*, *trip*; OSM: *relation*)
GTFS : ein Klick führt zur Übersicht von *route* bzw. *trip*.
OSM : ein Klick auf das Icon bzw. die ID zeigt die Objektdaten auf https://osm.org an.

⇅ ⇈ ⇊ Diese Zeichen im Tabellenkopf einer Spalte erlauben eine Sortierung der Tabelle nach dieser Spalte: nicht sortiert, aufsteigend, absteigend.

Sonstiges

Auf die Angabe von Beispielen mit IDs von realen GTFS- oder OSM-Daten wird hier bewusst verzichtet. Diese können bis zum Zeitpunkt der Konferenz schon wieder veraltet sein.

Als maximale akzeptable Distanz zwischen GTFS-*stop* und zugehöriger OSM-*platform* wurden 20 Meter gewählt. Das entspricht etwa der Länge eines Gelenkbusses (18 m) und erlaubt dennoch das Erkennen von falschen Bussteigen in Busbahnhöfen. Beim Vergleich lassen sich gleichwohl alle „Distanz-Werte“ und „Gewichtungen“ GTFS-*feed*-spezifisch konfigurieren.

Über das Vergleichs-Icon in der letzten Spalte der Tabelle einer länderspezifischen GTFS-Übersicht (Beispiel: DE [10]) lässt sich ein Vergleich GTFS versus GTFS anstoßen, z.B.: was hat sich zwischen zwei GTFS-Versionen geändert?

Im PTNA-Report gibt es am Anfang unter „*Datum der Daten*“ einen Button „**Start einer neuen Analyse anfordern**“. Ein Cron-Job schaut alle 15 Minuten nach solchen Anforderung und startet diese. Man muss nicht auf den nächtlichen Start (zwischen 02:00-05:00 Ortszeit der zu analysierenden Region) des PTNA-Reports warten.

Die Protokolle (log) und Ergebnisse aller automatisierten Aktivitäten von PTNA sind über die „**Statistik**“-Seite [11] (vor allem für den Autor selbst) erreichbar.

Ausblick / Ideen:

- Weiterführung der Umstellung des OSM-Daten-Downloads von Overpass API [18] zu Planet-Dumps/Updates/Filter/Extracts [19][20][21]
  - 109 Analysen sind umgestellt, 244 Konfigurationen müssen noch angepasst werden
- Ausgabe der Ergebnisse (score-Werte) des Vergleichs von GTFS- und OSM-Daten im PTNA-Report
  - Basierend auf den CSV-Daten im OSM-Wiki
  - Basierend auf den *gtfs:** tags in den OSM-Relationen
  - Portieren von > 1000 Lines-of-Code von JavaScript nach Perl
  - Problem: nur wenige score-Werte sind tatsächlich 0.00%, wie behandelt man die anderen?
- Einbau einer „sidebar“ über die Fehlermeldungen (Kategorien) im PTNA-Report ein- und ausgeblendet werden können
  - Analog zu KeepRight [12] und PTSA [13]
- Erweitern von Statistiken
  - Entwicklung der Anzahl Fehler einer Kategorie über einen Zeitraum
- Re-Design des sqlite3-Datenbankschemas (PTNA-Tabellen) für GTFS-Daten
  - Man lernt nie aus
- Bei der PTNA-Analyse (Link 7) einer OSM-*route*-Relation könnte man eine Prüfung auf Lücken und falsche Sortierung bei den Ways einbauen
  - Analog zum Relations-Editor in JOSM
  - Portieren von 600-1000 Lines-of-Code in Perl nach JavaScript

Anhang

Kontakt zum Autor:

Antonius „Toni“ Erdmann
Friedenstraße 21
D-85521 Ottobrunn
osm-ToniE@web.de
Community: „PTNA: news for Public Transport Network Analysis“ [14]
PTNA-Source-Code [15]

Literatur / Links:

[1] „PTNA - Public Transport Network Analysis“, *https://ptna.openstreetmap.de*

[2] „PTNA: Qualitätssicherung für ÖPNV in OpenStreetMap“, Toni Erdmann, FOSSGIS 2020 Tagungsband, Seite 109, *https://www.fossgis-konferenz.de/2020/data/FOSSGIS_Tagungsband_2020.pdf*

[3] „GTFS“, "General Transit Feed Specification", *https://gtfs.org/*

[4] „GTFS reference“, *https://gtfs.org/documentation/schedule/reference/*

[5] „GTFS best practices“, *https://gtfs.org/documentation/schedule/schedule-best-practices/*

[6] „OSM contributors“, *https://wiki.openstreetmap.org/wiki/Contributors*

[7] „Proposal:Public Transport“, approved version “PTv2”, *https://wiki.openstreetmap.org/w/index.php?title=Proposed_features/Public_Transport&oldid=625726*

[8] „Codes for identifying the principal subdivisions (e.g., provinces or states) of all countries coded in ISO 3166-1“, ISO, *https://en.wikipedia.org/wiki/ISO_3166-2*

[9] „GTFS“, Spezifikation der *gtfs:** tags im OSM-Wiki, *https://wiki.openstreetmap.org/wiki/GTFS*

[10] Übersicht über die von PTNA analysierten GTFS-*feeds* in Deutschland, *https://ptna.openstreetmap.de/gtfs/DE/index.php*

[11] Statistiken und Logging von PTNA, *https://ptna.openstreetmap.de/en/statistics.php*

[12] „KeepRight“, *https://keepright.at/report_map.php?lang=de*

[13] „PTSA - Public Transport Stop Analysis“, *https://gauss.whz.de/ptsa/#3/50.72/12.50*

[14] „PTNA: news for Public Transport Network Analysis“, OpenStreetMap Community Forum, *https://c.osm.org/t/ptna-news-for-public-transport-network-analysis/8383/*

[15] PTNA-Source-Code, *https://github.com/osm-ToniE*

[16] Lizenz CC0, *https://creativecommons.org/publicdomain/zero/1.0/*

[17] Lizenz CC-BY-4.0, *https://creativecommons.org/licenses/by/4.0/*

[18] Overpass-API, *https://wiki.openstreetmap.org/wiki/Overpass_API*

[19] Planet-Dumps, *https://planet.openstreetmap.org/pbf/planet-latest.osm.pbf*

[20] pyosmium-up-to-date, *https://github.com/osmcode/pyosmium*

[21] Osmium-Filter, Osmium-Extract, *https://github.com/osmcode/osmium-tool*

RoutingPlus goes Masterportal: Unser Weg zu einer modernen Benutzeroberfläche

Benjamin Würzler

Steigt mit ein in unseren Lightning Talk und hört von unseren Erkenntnissen aus dem agilen Entwicklungsprozess einer neuen graphischen Benutzeroberfläche für den Routingdienst RoutingPlus des Bundesamtes für Kartographie und Geodäsie (BKG). Dabei erfahrt ihr die Gründe, warum wir uns für die Open Source Lösung Masterportal entschieden haben, und welche Vorteile sich dadurch nicht nur für uns, sondern auch die Open Source Community ergeben.

Geplante Entwicklungsroute:

Das bewährte, aber in die Jahre gekommene RoutingPlus Demoportal ablösen und auf Grundlage des Masterportal 3.0 durch eine neue Anwendung ersetzen.

Berechnete Entwicklungsfahrzeit:

Fertigstellung innerhalb von knapp 2 Jahren bis 2025.

Wer möchte mitkommen?

Basierend auf dem leistungsstarken openrouteservice (ORS) des Heidelberg Institute for Geoinformation Technology (HeiGIT) ist RoutingPlus seit mehreren Jahren als etablierten Routingdienst des Bundes anerkannt. Für die Bundesverwaltung stellt dieser Dienst eine wichtige Infrastruktur dar, die über eine flexible Dienste-API in eigene Anwendungen und Prozesse integriert werden kann.

seit einigen Jahren wächst jedoch der Bedarf die RoutingPlus Dienste auch über eine graphische Benutzeroberfläche verwenden zu können. Vor allem für weniger technikaffine Nutzerinnen und Nutzer wurde dafür 2017 das RoutingPlus Demoportal veröffentlicht. Dieses dient zudem als Demonstrator, welcher die umfangreichen Funktionalitäten sichtbar und zugänglich macht. Gleichzeitig hat das Masterportal in den letzten Jahren mit einem Interface-Modul eine weitere Möglichkeit geschaffen, die Routingdienste des BKG anzubieten und den Zugang für alle Anwender zu erleichtern.

Aufgrund von veränderten Anforderungen und Sicherheitsaspekten sowie steigenden Erwartungen an die Leistungsfähigkeit und Flexibilität des Softwarebetriebs haben wir Anfang 2023 den Startschuss für die Ablösung des bisherigen RoutingPlus Demoportals gegeben. Diese Entscheidung markierte den Beginn eines umfassenden Projekts, das ein völlig neues, zukunftsfähiges Portal hervorbringen soll. Welche Fortschritte wir seither gemacht haben, welche Herausforderungen uns begegneten und welche „Abkürzungen“ oder unerwarteten „Umwege“ wir genommen haben, erfahrt ihr in unserem Beitrag auf der FOSSGIS 2025.

Steigt deshalb mit ein in unseren Lightning Talk und hört von unseren Erkenntnissen aus dem agilen Entwicklungsprozess einer neuen graphischen Benutzeroberfläche für den Routingdienst RoutingPlus des Bundesamtes für Kartographie und Geodäsie (BKG). Dabei erfahrt ihr die Gründe, warum wir uns für die Open Source Lösung Masterportal entschieden haben, und welche Vorteile sich dadurch nicht nur für uns, sondern auch die Open Source Community ergeben. Als Early-Adopter der Masterportal Version 3.0 teilen wir die Erfahrungen, die wir während der Entwicklung gesammelt haben, und die dabei aufgetretenen Herausforderungen, sowohl technisch als auch organisatorischer Natur.

Zudem geben wir Einblicke, wie wir durch agile Prozesse und enge Zusammenarbeit die Erfassung und Umsetzung von User-Anforderungen vorantreiben, um das neue RoutingPlus Portal optimal an die Bedürfnisse zukünftiger Nutzerinnen und Nutzer anzupassen. Natürlich führen wir auch den aktuellen Entwicklungsstand der neuen Anwendung vor und zeigen darüber hinaus, wie das Projekt langfristig angelegt ist, um den neuen Herausforderungen und amtlichen Anforderungen gerecht zu werden.

Römische Grabdenkmäler im 3D-Webviewer

Johannes Frank[1], Homayoon Afsharpoor[1]

[1]Hochschule Mainz - University of Applied Sciences

Zusammenfassung: Wie kann aus hochgenauen Geodaten wie 3D-Punktwolken und 3D-Objekten ein Museumsviewer entwickelt werden, der auch von fachfremden Personen genutzt werden kann? Hierfür wurden die Open-Source-Webframeworks Potree und 3D HOP erweitert, um einen immersiven und intuitiven Museumsrundgang zu ermöglichen. Die Ergebnisse des digitalen Museums werden in einem Kurzvortrag präsentiert.

Schlüsselwörter: 3D-Webviewer, Römische Grabdenkmäler, Potree, Museumsviewer

Die Ausstellungsstücke:

Museen fungieren als Hüter zeitlicher Relikte und gewähren faszinierende Einblicke in vergangene Epochen. Das Rheinische Landesmuseum in Trier beherbergt unter anderem die beeindruckende Gräberstraße, eine umfassende Sammlung von Grabmonumenten, Denkmälern und Grabsteinen aus der Epoche der Römer (siehe Abb. 1). Die dort ausgestellten Objekte sind Überreste der Gräberfelder der Römer. Diese erstreckten sich traditionell außerhalb der Stadtmauern römischer Städte, wie Augusta Treverorum (heute Trier), entlang der Hauptverkehrsstraßen. Insbesondere in den vorderen als "vornehmer" geltenden Reihen, zeugen aufwendige Grabmonumente von der gesellschaftlichen Stellung des Verstorbenen und seiner Familie. Diese Gräber bieten eine reiche Quelle von Einblicken in den Alltag der Römer, indem sie zahlreiche Details aus ihrem täglichen Leben offenbaren. Ein herausragendes Beispiel eines der größten zusammenhängend ausgegrabenen und rekonstruierten Gräberfelder ist das der Neumagener Funde. Die Originalfunde sind in der Gräberstraße aufgebaut. Diese Spolien stammen aus den Gräberfeldern von Trier und wurden während der Befestigungsarbeiten in Neumagen dorthin transportiert und in die Mauern integriert.

Die Datengrundlage:

Die digitale Rekonstruktion des Museumsraums des Rheinischen Landesmuseums Trier basiert auf einer umfangreichen Datenerfassung. Durch diese vorangegangenen Messkampagnen ist eine präzise Modellierung der Ausstellungsräume ermöglicht worden. Mittels photogrammetrischer Verfahren mit einer Nikon Z7 II Systemkamera wurden detaillierte texturierte 3D-Modelle einiger Ausstellungsstücke erstellt. Zusätzlich trug eine Drohne (DJI Mavic 3 Enterprise) zur Datensammlung bei. Außerdem wurde mit dem hochpräzisen TLS RTC 360 ein dreidimensionales Abbild des Bereichs in der Gräberstraße er-

stellt. Durch die Einrichtung eines Festpunktfeldes und die Bestimmung lokaler Koordinaten der Festpunkte sind alle generierten Daten – von den texturierten 3D-Modellen bis zu den Punktwolken – auf ein gemeinsames lokales Koordinatensystem abgestimmt.

Die Umsetzung:

Für die Visualisierung dieser Daten wird eine Kombination der Webanwendungen Potree [1] für die großvolumigen Punktwolken und 3D HOP [2] für die Darstellung der einzelnen texturierten 3D-Objekte verwendet. Eine zentrale Herausforderung besteht darin, eine nahtlose und benutzerorientierte Schnittstelle zwischen diesen beiden Anwendungen herzustellen, die ein intuitives Erleben der virtuellen Museumswelt gewährleistet.



Abb. 2: Interaktion mit der Annotation zu dem 3D-Objekt „Grabturm“. Hierbei kann der Objektviewer geöffnet werden.

Es wurde ein Prozessablauf entwickelt, um die Punktwolken der Gräberstraße miteinander zu verknüpfen und in einem Octree-Format [3] speichereffizient vorzuhalten. Dadurch ist eine flüssige Visualisierung der Punktwolke in einem Webbrowser mit dem Open-Source-Webframework Potree ermöglicht worden.

Auch die speicherintensiven 3D-Objekte im PLY-Format wurden in das hierarchische Nexus-Format [4] konvertiert. Das Potree-Framework wurde erweitert, um auch 3D-Objekte in diesem Format gemeinsam mit der Punktwolke anzeigen zu lassen (siehe Abb. 2). Darüber hinaus wurden zu jedem Objekt die Metadaten (die Informationen der Ausstellungstafeln) dynamisch mit dem 3D-Objekt verknüpft. Somit war eine einfache Nutzerinteraktion mit einem Ausstellungsstück durch die sogenannten Annotationen möglich. Die dynamische Erstellung der Annotationen wurde so implementiert, dass bei wechselnden Ausstellungsstücken oder Räumlichkeiten, keine Änderungen im Quellcode des Museumsviewers vorgenommen werden müssen. Dadurch kann die Wiederverwendbarkeit und Erweiterbarkeit des Museumsviewers gewährleistet werden.

Die Bewegung ist durch die Maus und die Tastatur frei in den Ausstellungsräumen möglich. Hierbei wurde auf die Gamification zurückgegriffen und die typischen WASD-Tasten zur Bewegung implementiert. Als Unterstützung und Orientierungshilfe wurde eine Mini-Map (siehe Abb. 3), sowie ein Rundflug in den Museumsviewer integriert.

Neben dem Museumsviewer wurde der Objektviewer entwickelt, mit dem

einzelne Ausstellungsstücke als 3D-Objekte visualisiert werden können (siehe Abb. 4). Die Webseite bietet auch die Möglichkeit zugehörige Metadaten zu den Objekten anzuzeigen und detaillierte Untersuchungen an dem Objekt vorzunehmen.

Der Objektviewer eines jeden Ausstellungsstück kann entweder durch die Interaktion mit einer Annotation oder über eine Übersichtstabelle im Museumsviewer in einem neuen Browser Tab geöffnet werden. Dadurch bleibt der Fortschritt des digitalen Museumsrundganges erhalten und zusätzlich können sich die kleinen Details auf den Objekten ohne die Punktwolke betrachten lassen. Auch hier können sich alle zu einem Ausstellungsstück gespeicherten Metadaten, wie Inschriften mit Übersetzungen, Beschreibungen und Materialangaben anzeigen lassen. Dadurch kann der Museumsviewer nicht nur als Visualisierungstool von Ausstellungsstücken genutzt werden, sondern auch als digitale Alternative zur Wissensvermittlung eingesetzt werden. Durch die dynamische Nutzung der Frameworks lassen sich einfach neue Punktwolken, 3D-Objekte und Metadaten addieren, welche automatisch in den Museumsviewer übernommen werden.

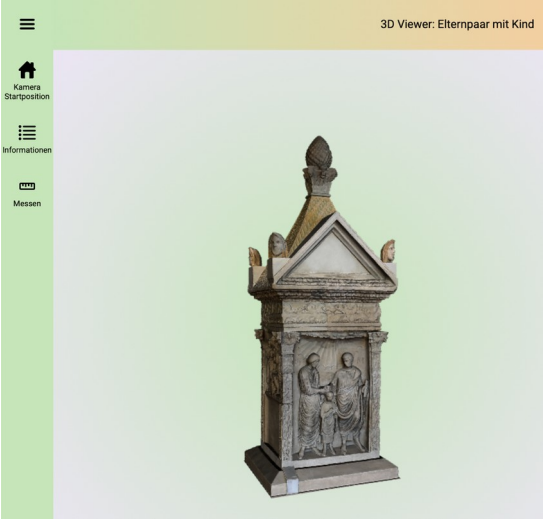


Abb. 4: Der dynamisch erzeugte Objektviewer für das Ausstellungsstück „Elternpaar mit Kind“ umgesetzt durch das 3D HOP-Framework.

Die Nutzung:

Der Museumsviewer basiert auf einer Messkampagne des i3 Mainz [5]. Da die Datenschutzbestimmungen der Geodaten nicht ausreichend geklärt sind, wird der Museumsviewer nur prototypisch auf den eigenen Servern der Hochschule Mainz gehostet.

Kontakt zum Autor:

Herr Johannes Frank
Hochschule Mainz
Lucy-Hillebrand-Straße 2
55128 Mainz
Johannes-frank@hotmail.com

Literatur

[1] Potree github-Repository: https://github.com/potree/potree (zuletzt aufgerufen: 30.01.2025)

[2] 3D HOP-Framework: https://3dhop.net (zuletzt aufgerufen: 30.01.2025)

[3] Potree Converter: https://github.com/potree/PotreeConverter (zuletzt aufgerufen: 30.01.2025)

[4] Nexus Format: http://vcg.isti.cnr.it/vcgtools/nexus// (zuletzt aufgerufen: 30.01.2025)

[5] i3 Mainz Forschungsprojekt: https://i3mainz.hs-mainz.de/projekte/grabdenkmaeler-des-trevererge-bietes/ (zuletzt aufgerufen: 30.01.2025)

fiboa - Standardisierte Feldgrenzen

Matthias Mohr

Das Projekt Field Boundaries for Agriculture (fiboa) zielt darauf ab, Feldgrenzen / Feldblöcke in einem einheitlichen Format auf globaler Ebene offen zugänglich zu machen. Dies wird ermöglicht durch eine offene Spezifikation für Feldgrenzen. Dazu sind quelloffene Programme und frei evrfügbare Daten verfügbar, mit denen man schnell und einfach loslegen kann. Diese Präsentation gibt eine kurze EInführung in das Projekt, mit der Hoffnung auf einen regen Austausch im Anschluss.



Der Kern von fiboa ist eine Spezifikation für die Darstellung von Feldbegrenzungsdaten in GeoJSON und GeoParquet in einer standardisierten Form, mit mehreren optionalen Erweiterungen, die zusätzliche Attribute spezifizieren. Das Kerndatenschema von fiboa ist recht einfach - es besteht aus einer Reihe von Definitionen für Attributnamen und -werte. Die Anzahl der Attribute im Kern ist von vornherein recht gering. Die Idee ist, dass die meisten der „interessanten“ Daten über das Feld in den „Erweiterungen“ zu finden sind.

Mehr über die Technologien hinter der fiboa-Spezifikation erfahren, den vollständigen Text der Spezifikation lesen und verfügbare offene Datensätze und Erweiterungen unter dem folgenden Link erkunden: https://fiboa.org

Absicherung von Diensten mit Keycloak

Astrid Emde

Nicht jeder OWS Dienst soll offen verfügbar sein. Nicht jeder Serversoftware über die Dienste konfiguriert werden können, stellt eine Möglichkeit der Absicherung zur Verfügung. Wie kann eine einfache und einheitliche Lösung für den geschützten Zugriff aussehen?

Hier kommt Keycloak ins Spiel.

Keycloak ist ein Open Source Projekt und unterstützt Single Sign-On mit Identitäts- und Zugriffsmanagement für Anwendungen und Dienste. Dabei unterstützt Keycloak verschiedene Protokolle wie z.B OpenID oder OAuth. Es wird eine Benutzerverwaltung sowie der Aufbau von Rollen unterstützt. Die Möglichkeiten sind sehr umfangreich.

Anhand eines einfachen Beispiels wird die Absicherung demonstriert.

Datenbankschema mit Mermaid visualisieren

Jakob Miksch

Vorstellung von mermerd und Mermaid zu visuellen Darstellung von (Geo)-Datenbankstrukturen. Zusätzlich wird gezeigt wie man diesen Prozess automatisiert.

Einer Geodateninfrastruktur liegt üblicherweise ein durchdachtes Datenbankschema zugrunde. Um dies allen Beteiligten verständlich zu machen bieten sich ER-Diagramme (Entitiy-Relationsship) an. Mittels mermerd ist es möglich die Tabellenstruktur einer Datenbank (z.B. Postgres) zu extrahieren und mittels der Aufzeichnungssprache Mermaid visuell darzustellen. Dieser Vortrag beschreibt wie dieser Prozess mittels Continuous Integration automatisiert werden kann, damit jederzeit das eine visuelle Darstellung des Datenbank Schemas für alle Team-Mitglieder zu sehen ist.

Koordinatenreferenzsysteme für D-A-CH

Javier Jimenez Shaw

Die Eigenschaften und Unterschiede der alten und neuen Koordinatenreferenzsysteme aus der D-A-CH-Region werden mithilfe von spatialreference.org erklärt.

In Deutschland, Österreich und der Schweiz gibt es verschiedene Koordinatenreferenzsysteme (CRS). Einige sind alt wie “DHDN / 3-degree Gauss-Kruger zone 3” oder “LV03”, einige neue wie “ETRS89/ DREF91/2016 / UTM zone 32N (N-zE)” oder “LV95”. Die Eigenschaften der CRS und ihre Unterschiede werden erklärt.

Wir werden sehen, wie Koordinatenreferenzsysteme mit spatialreference.org (von PROJ) gefunden werden können. Außerdem werden die Unterschiede zwischen Pulkovo, DHDN, CH1903+, MGI und ETRS89 datums betrachtet.

Des Weiteren werden die verschiedenen vertikalen Koordinatenreferenzsysteme der drei Länder wie “DHHN2016”, “LN02” oder “LHN95” beleuchtet.

Lebewohl Web Merkator

Pirmin Kalberer

Die meisten heutigen Webkarten verwenden die Web-Merkator-Projektion.

Das Hauptproblem der Web-Merkator-Projektion ist die starke Verzerrung weit vom Äquator entfernter Flächen.

2018 veröffentlichten Bojan Šavrič, Tom Patterson und Bernhard Jenny ihre Arbeit zur

Equal-Earth-Kartenprojektion, einer flächentreuen

Projektion für Weltkarten.

Dieser Vortrag zeigt, wie man die Equal-Earth-Kartenprojektion für Webkartographie mit verschiedenen Arten von Datenquellen nutzen kann.

Eine wachsende Sammlung von Informationen zur Verwendung von Equal Earth ist verfügbar unter equal.bbox.earth.