TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

ARCHIVING DIGITAL MAPS WITH GEOPACKAGE AND VECTOR-TILE DISSEMINATION

Yunnan Chen

Supervisors: Dr. Markus Jobst (BEV)
Prof. Dr. Georg Gartner (TUW)
Dr.-Ing. Christian Murphy (TUM)

Outline

1. Introduction
2. Literature Review and Theoretical Background
3. Methodology and Case Study
4. Results
5. Discussion and Future Work
6. Conclusions
7. References

1.Introduction

- Background and Motivation
- Research Goal and Objectives
- Research Questions

1.Introduction

1.1 Background and Motivation

“Currently, there are no easy answers as how best to guarantee that digital data will last into the future without active management to preserve the data.” - Bleakly, D. R., 2002

- Limitations of current geospatial data format in archiving digital maps
- Time consuming operation for archiving digital maps
- Digital map distribution Issues

1.Introduction

1.2 Research Goal and Objectives

- Goal:

  An alternative method for long-term archiving and distribution of digital maps

- Objective:

1. To determine and analyze a suitable data format that has the ability to store geospatial data and its relevant information such as metadata and style.

2. To establish a suitable method for archiving, publishing, and distributing digital maps with customized styling.

3. To determine a suitable programming and to develop a script that minimizes the human effort during the procedure.

1.Introduction

1.3 Research Questions

- Is it possible to embed relevant metadata for future usage in digital maps using GeoPackage?
- Is it possible to embed styling from GeoPackage when distributing digital maps that use vector-tile format?
- Is it possible to create a Python script that partially automate the procedures of embedding relevant information into GeoPackage?

2. Literature Review and Theoretical Background

- Principle of Digital Maps Preservation and Archiving
- Status Quo and Challenges
- Evaluation of Geospatial Data Format for Archiving Purpose

2. Literature Review and Theoretical Background

2.1 Principle of Digital Maps Preservation and Archiving

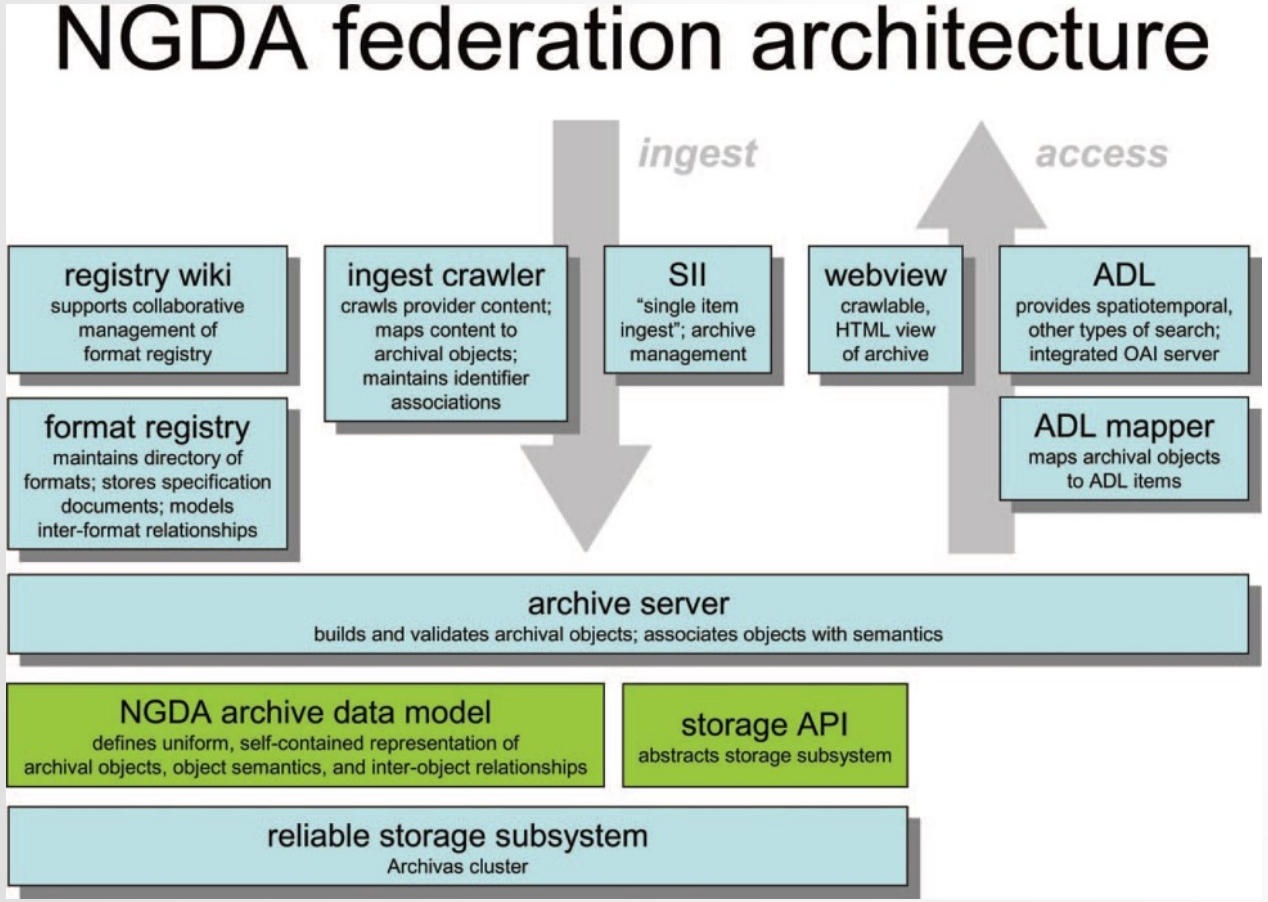
Principles for the Preservation of Geospatial Data

1. Archiving begins at the point of data creation.
2. Data should be preserved in a way that non-specialists can understand.
3. Information objects should be independently understandable.
4. A graphical representation of the data should be preserved with the data itself.
5. Avoid restrictive assumptions about future use of the data.

Five principles for the preservation of geospatial data (Clark, 2016)

2. Literature Review and Theoretical Background

2.2 Status Quo and Challenges



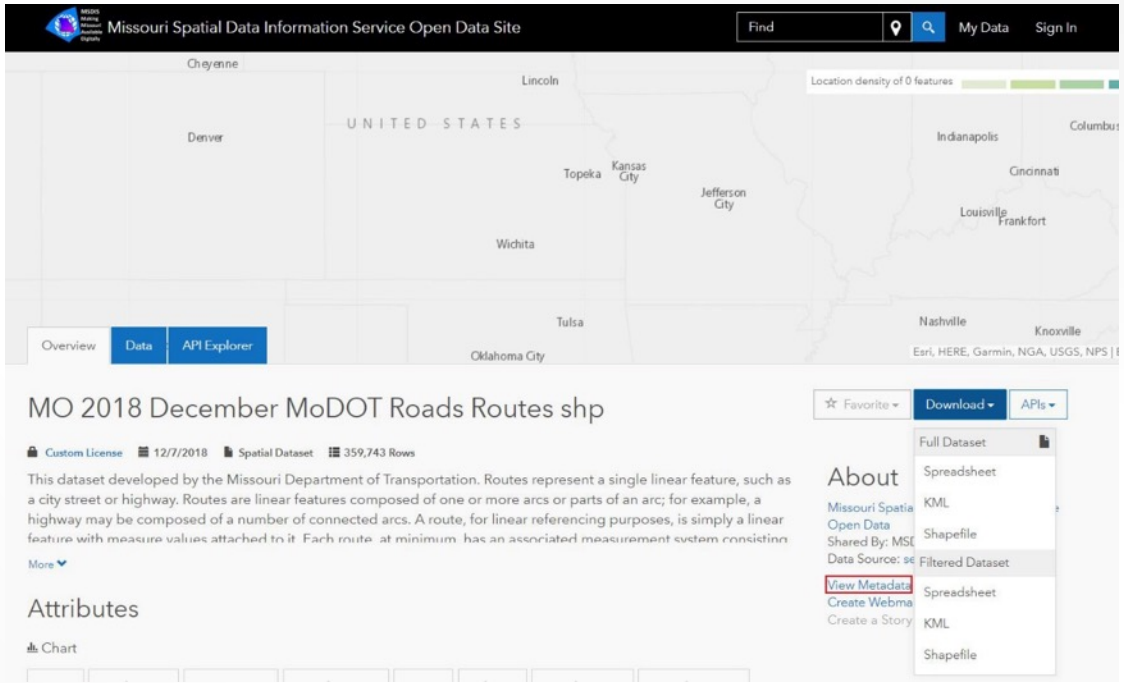
The NGDA architecture at the UCSB (Erwin & Sweetkind-Singer, 2009)

- Modern cartography through digital approach requires a more complicated framework to ensure the processes of reproduction, sharing, and publishing. (Jobst, M & Gartner, G, 2011)
- Backup and storage of digital maps; The importance of storing data representation. (Lauriault et al., 2011)
- Proprietary nature of geospatial data format raises a technical risk for long-term archiving. (Clark, 2016)
- Preserve the application schemas and its relationships with the relating geospatial dataset. (Shaon et al., 2011)
- Decentralized Production; Complexity of Data; Managing Versions of the Same data. (Locher & Thermens, 2012)

2. Literature Review and Theoretical Background

2.2 Status Quo and Challenges

GeoMAPP Partners (North Carolina Center for Geographic Information and Analysis & North Carolina Department of Cultural Resources, 2011)



Support of Metadata on Geoportal of MSDIS website

2. Literature Review and Theoretical Background

2.3 Evaluation of Geospatial Data Format for Archiving Purpose

2.3.1 Vector Data Format – Shapefile

- Most commonly used vector data format for archiving and distributing geospatial data
- Non-topological binary format
- Multi-File format
- 4GB size limit
- Poor support for attribute data type
- Metadata in separated data file
- No support for symbology and feature labels
- Data compression issue with data dissemination

2. Literature Review and Theoretical Background

2.3 Evaluation of Geospatial Data Format for Archiving Purpose

2.3.2 Raster Data Format – GeoTIFF

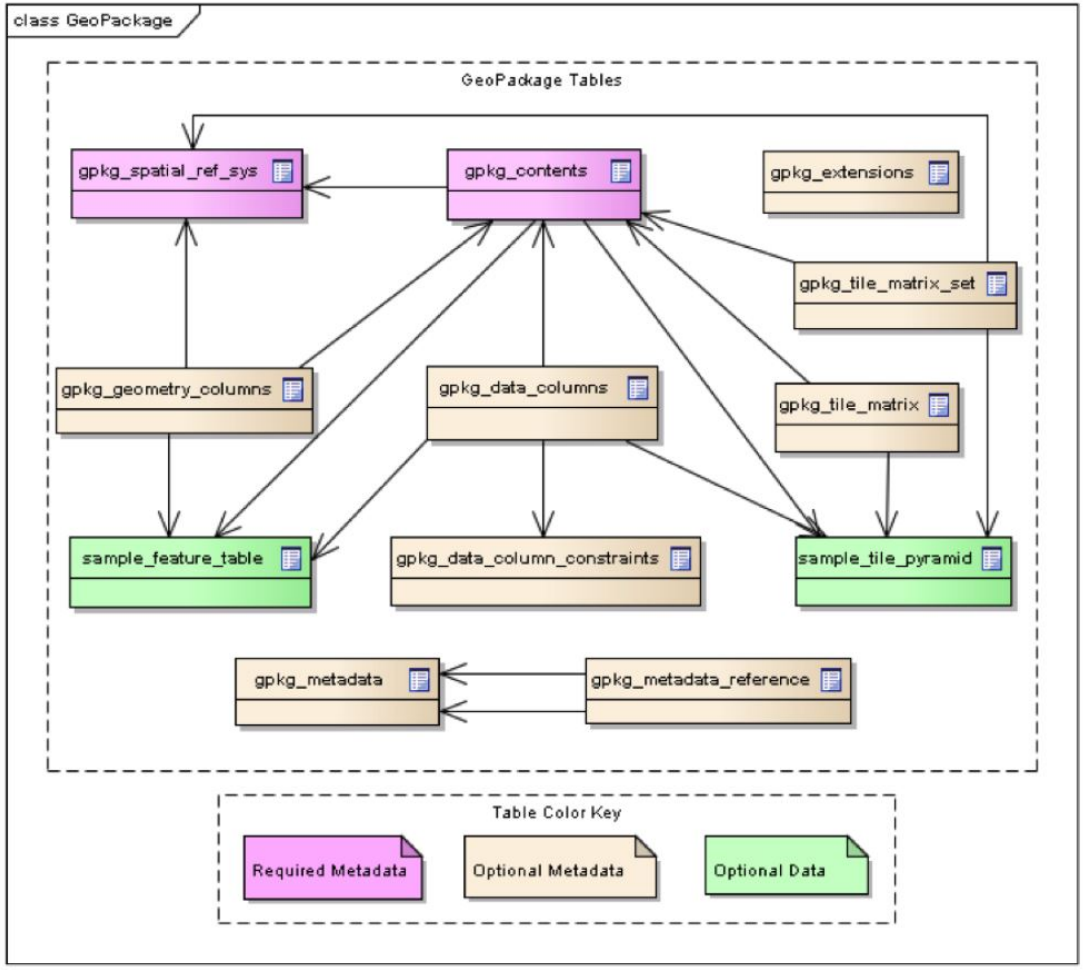
- Most commonly used raster data format for archiving and distributing geospatial data
- 32bit binary format
- Multi-File format
- 4GB size limit
- Metadata in separated data file
- No support for symbology
- Data compression and size issue with data dissemination

2. Literature Review and Theoretical Background

2.3 Evaluation of Geospatial Data Format for Archiving Purpose

2.3.2 Alternative Data Format – GeoPackage

- *"open, standards-based, platform-independent, portable, self-describing, compact format for transferring geospatial information" – GeoPackage Encoding Standard(2018)*
- 64bit ASCII format
- SQLite database file – single file format
- Support for storing both vector and raster data
- 150TB storage size
- Metadata support by default with ISO 19115
- Style support through extension
- *"Interoperable file between SDI and the mobile system" - Bogossian et al.,(2014)*



3. Methodology and Case Study

3.1 Case Study Background and Scenario

3.2 Overview of Approach

3.2.1 Original Dataset Review

3.2.2 Geospatial Data Format Conversion

3.2.3 Extracting Metadata from PDF

3.2.4 Inserting Metadata

3.2.5 Editing Style in QGIS

3.2.6 Adding Style into GeoPackage

3.2.7 Uploading GeoPackage to Geoserver

3.2.8 Importing Style and Digital Maps Distribution

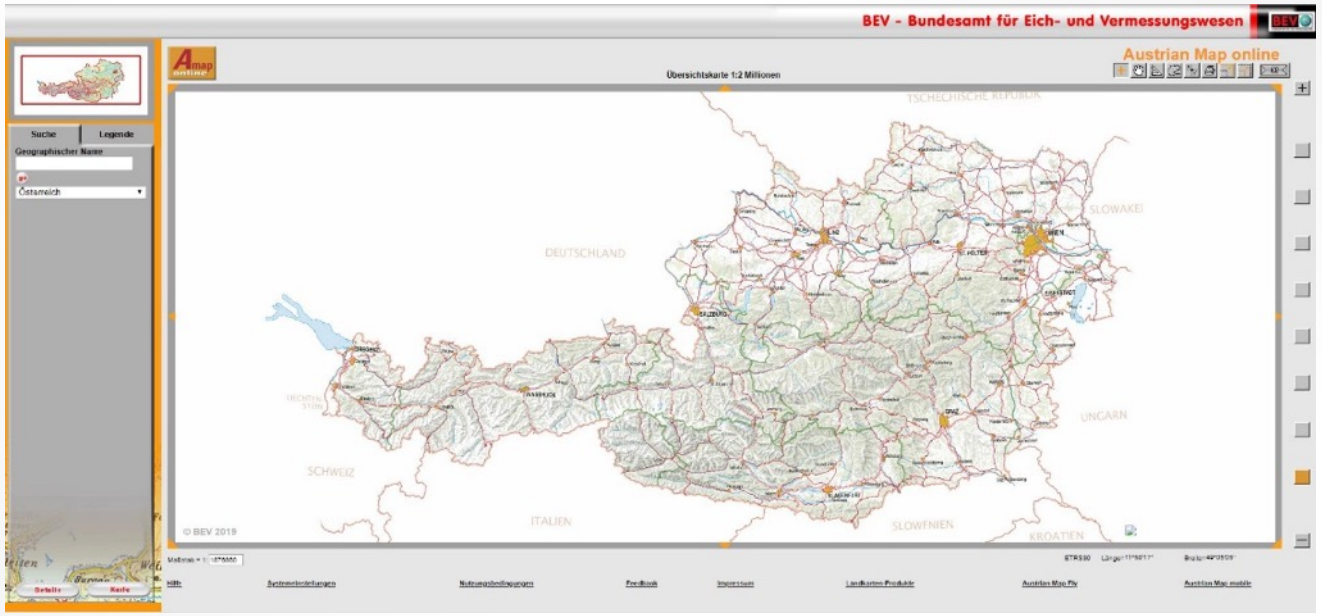
3. Methodology

3.1 Case Study Background and Scenario

Scenario 1
GIS technician: “I want to archive the national maps, but I only have the shapefile from BEV. Should I create the map from scratch?”

Scenario 2
Regular Austrian Citizen: “I want to have the city map with national style as a digital collection, do I have to buy it from BEV? ”

Scenario 3
Emergency Response Event such as Etna Volcano eruption in Sicily, Italy, “Is it possible to make and share the maps in a short time period”

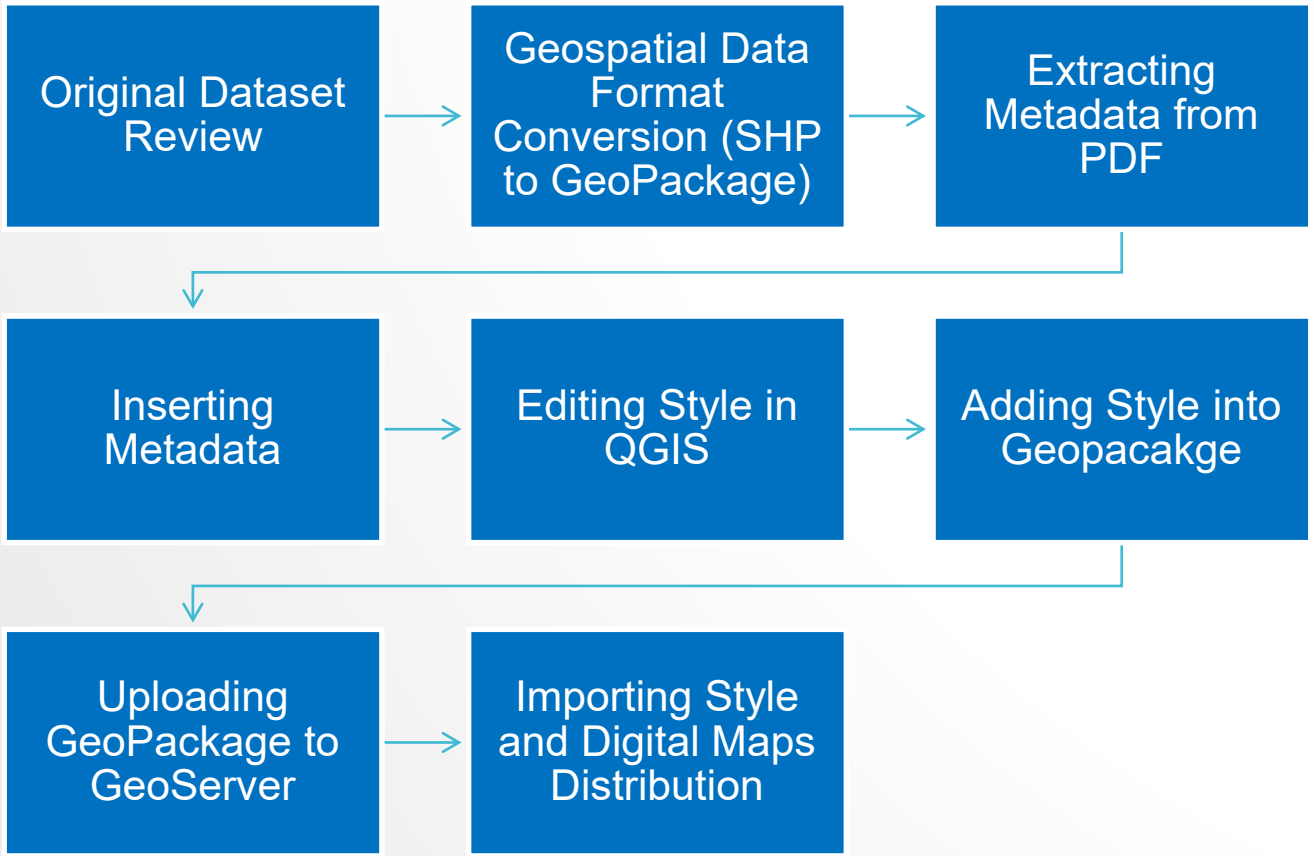


Austria Map Online (AMap Online) –A Web Map Viewer hosted by BEV

Case Study
The cartographic model (KM1000-V) and national map style from Austrian Mapping Agency(BEV)

3. Methodology

3.1 Overview of Approach



Workflow Overview

3. Methodology

3.2 Overview of Approach

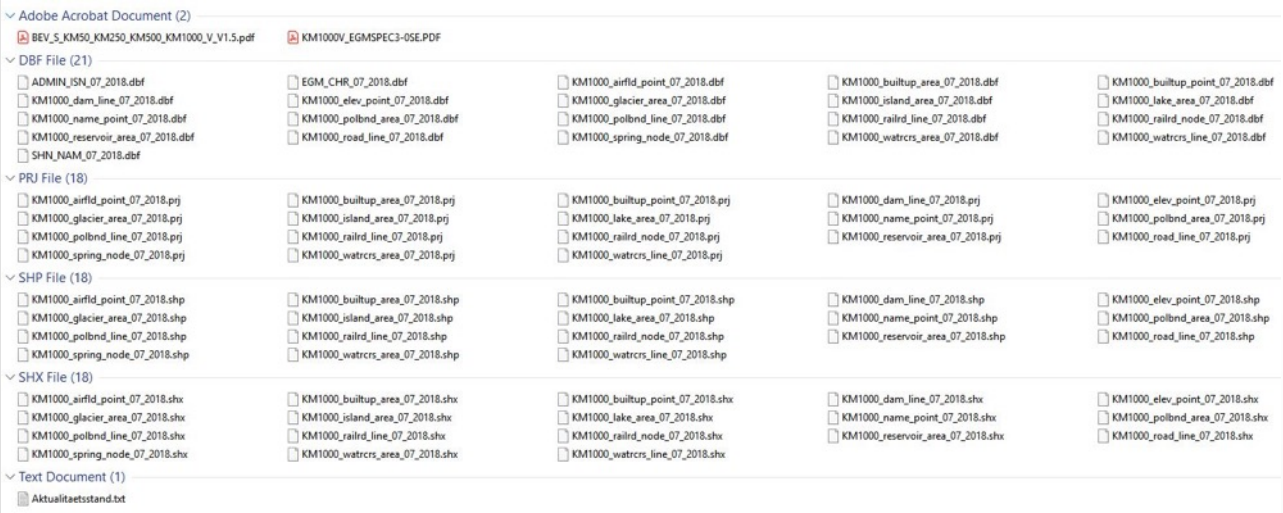
WORKFLOW	SOFTWARE/PROGRAMMING LANGUAGE
Original Dataset Review	PDF Reader
Geospatial Data Format Conversion	QGIS 3.4.1
Extracting Metadata From PDF	PDF Reader/Python 3 IDLE/SQLite
Inserting Metadata	Python 3 IDLE/SQLite
Editing Style in QGIS	QGIS 3.4.1/Color Picker
Adding Style into GeoPackage	QGIS 3.4.1
Uploading GeoPackage	GeoServer 2.15
Importing Style and Digital Maps Distribution	GeoServer 2.15

Table 1. Software and Programming Language Requirement

3. Methodology

3.2 Overview of Approach

3.2.1 Original Dataset Review



Original Dataset Overview

18 SHPs, 2 PDF files, 3 DBF files, 1 TXT file, 78 files in total

BEV - Bundesamt für Eich- und Vermessungswesen

2.4 Kartographisches Modell 1:1 Million – Vektor (KM1000-V)

2.4.1 General structure

2.4.1.1 About KM1000-V

Data of KM1000-V is the Austrian part of EuroGlobalMap (EGM) the pan-European vector dataset at small scale. EGM Database is intended to be used in map scale 1:1 000 000.
Detailed specifications are described in ***EGMspec3-0se.pdf***. This document is a summary of the most relevant specifications.

2.4.1.2 Data format and file table

Data of KM1000-V is stored in these files:

Shape Files	Description	Type
KM1000_AIRFLD_POINT.shp	Airport / Airfield	Point
KM1000_DAM_LINE.shp	Dam/ Weir	Line
KM1000_ELEV_POINT.shp	Height point	Point
KM1000_GLACIER_AREA.shp	Glacier	Area
KM1000_ISLAND_AREA.shp	Island	Area
KM1000_LAKE_AREA.shp	Lake	Area
KM1000_NAME_POINT.shp	Named location	Point
KM1000_POLBND_AREA.shp	Administrative area	Area
KM1000_POLBND_LINE.shp	Administrative boundary	Line
KM1000_RAILRD_NODE.shp	Railway station	Point (Node)
KM1000_RAILRD_LINE.shp	Railway	Line
KM1000_RESERVOIR_AREA.shp	Reservoir	Area
KM1000_ROAD_LINE.shp	Road	Line
KM1000_BUILTUP_AREA.shp	Built-up area	Area
KM1000_BUILTUP_POINT.shp	Built-up point	Point
KM1000_SPRING_NODE.shp	Spring/ Water hole (connected)	Point (Node)
KM1000_WATRCRS_AREA.shp	Watercourse	Area
KM1000_WATRCRS_LINE.shp	Watercourse	Line

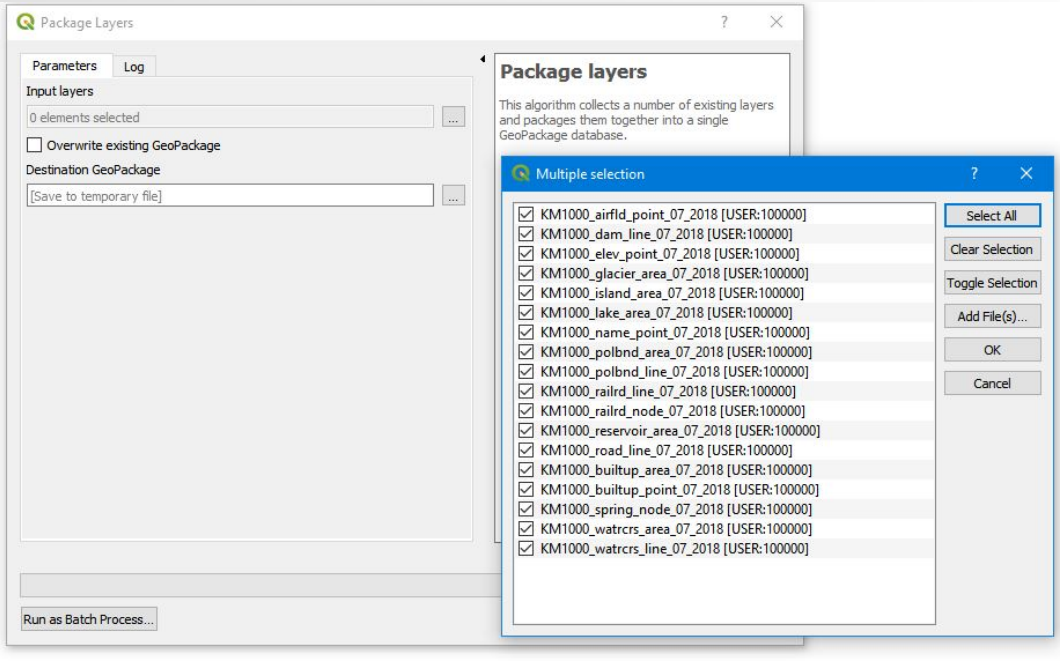
Info Tables	Description
ADMIN_ISN.dbf	This table includes the names of the administrative hierarchy levels
EGM_CHR.dbf	This table describes the national character sets used for each language.
SHN_NAM.dbf	The table includes the names of the units of all administrative levels.

Example of Metadata in BEV PDF Page 22

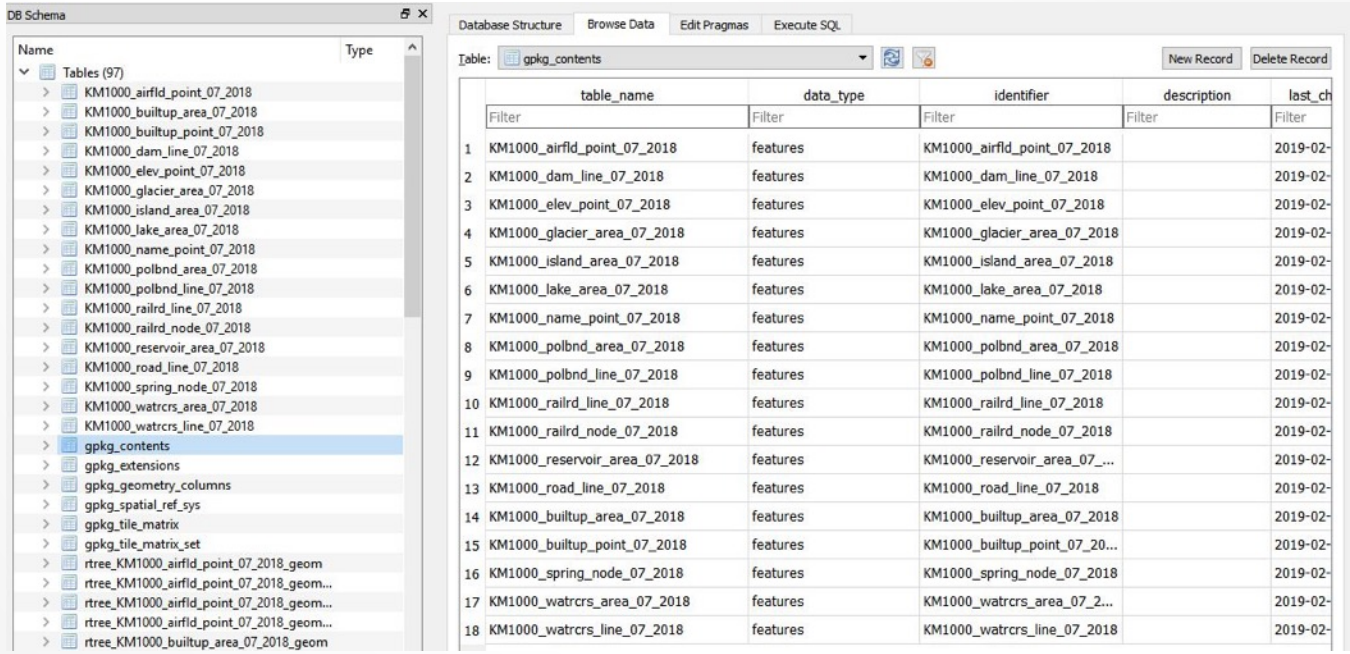
3. Methodology

3.2 Overview of Approach

3.2.2 Geospatial Data Format Conversion



Package Layer Tool Box in QIS



gpkg_contents table in output GeoPackage using DB Brower for SQLite

3. Methodology

3.2 Overview of Approach

3.2.3 Extracting Metadata from PDF

```python
##Convert PDF to Text
##Source Code: http://stanford.edu/~mgorkove/cgi-bin/rpython_tutorials/Using%20Python%20to%20Convert%20PDFs%20to%20Text%20Files.php
def convert(fname, pages=None):
    if not pages:
        pagenums = set()
    else:
        pagenums = set(pages)

    codec = 'utf-8'
    output = StringIO()
    manager = PDFResourceManager()
    converter = TextConverter(manager, output, laparams=LAParams())
    interpreter = PDFPageInterpreter(manager, converter)

    infile = open(fname, 'rb')
    for page in PDFPage.get_pages(infile, pagenums):
        interpreter.process_page(page)
    infile.close()
    converter.close()
    text = output.getvalue()
    output.close
    return text

output = convert(BEVPDF, pages=[21])
print (output)

file = open("PDFPage22.txt", "w", encoding='utf-8')
file.write(output)
file.close()
print ("FINISHED Converting pdf to txt")
```

```python
##Convert tables in PDF to CSV
tabula.convert_into(BEVPDF,"DataFormatFileTable.csv",encoding='utf-8',multiple_tables= True,output_format='data_format',pages="22")
tabula.convert_into(BEVPDF,"tablepage23.csv",encoding='utf-8',guess = False, multiple_tables= True,output_format='data_format',pages="23")
tabula.convert_into(BEVPDF,"tablepage24.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="24")
tabula.convert_into(BEVPDF,"tablepage25.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="25")
tabula.convert_into(BEVPDF,"tablepage2627.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="26,27")
tabula.convert_into(BEVPDF,"tablepage28.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="28")
tabula.convert_into(BEVPDF,"tablepage2930.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="29,30")
tabula.convert_into(BEVPDF,"tablepage31.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="31")
tabula.convert_into(BEVPDF,"tablepage32.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="32")
tabula.convert_into(BEVPDF,"tablepage33.csv",encoding='utf-8',guess = True, lattice = True, multiple_tables= True,output_format='data_format',pages="33")
print ("FINISHED EXTRACT Table from PDF")
```

3. Methodology

3.2 Overview of Approach

3.2.3 Extracting Metadata from PDF

2.4 Kartographisches Modell 1:1 Million – Vektor (KM1000-V)

2.4.1 General structure

2.4.1.1 About KM1000-V
Data of KM1000-V is the Austrian part of EuroGlobalMap (EGM) the pan-European vector dataset at small scale. EGM Database is intended to be used in map scale 1:1 000 000.
Detailed specifications are described in EGMspec3-0se.pdf. This document is a summary of the most relevant specifications.

2.4.1.2 Data format and file table

Data of KM1000-V is stored in these files:

Shape Files

KM1000_AIRFLD_POINT.shp

KM1000_DAM_LINE.shp

KM1000_ELEV_POINT.shp

KM1000_GLACIER_AREA.shp

KM1000_ISLAND_AREA.shp

KM1000_LAKE_AREA.shp

KM1000_NAME_POINT.shp

KM1000_POLBND_AREA.shp

KM1000_POLBND_LINE.shp
KM1000_RAILRD_NODE.shp

KM1000_RAILRD_LINE.shp

KM1000_RESERVOIR_AREA.shp

KM1000_ROAD_LINE.shp

tabletest

Shape Files	Description	Type
KM1000_AIRFLD_POINT.shp	Airport / Airfield	Point
KM1000_DAM_LINE.shp	Dam/ Weir	Line
KM1000_ELEV_POINT.shp	Height point	Point
KM1000_GLACIER_AREA.shp	Glacier	Area
KM1000_ISLAND_AREA.shp	Island	Area
KM1000_LAKE_AREA.shp	Lake	Area
KM1000_NAME_POINT.shp	Named location	Point
KM1000_POLBND_AREA.shp	Administrative area	Area
KM1000_POLBND_LINE.shp	Administrative boundary	Line
KM1000_RAILRD_NODE.shp	Railway station	Point (Node)
KM1000_RAILRD_LINE.shp	Railway	Line
KM1000_RESERVOIR_AREA.shp	Reservoir	Area
KM1000_ROAD_LINE.shp	Road	Line
KM1000_BUILTUP_AREA.shp	Built-up area	Area
KM1000_BUILTUP_POINT.shp	Built-up point	Point
KM1000_SPRING_NODE.shp	Spring/ Water hole (connected)	Point (Node)
KM1000_WATRCRS_AREA.shp	Watercourse	Area
KM1000_WATRCRS_LINE.shp	Watercourse	Line
Info Tables	Description	
ADMIN_ISN.dbf	This table includes the names of the administrative	
	hierarchy levels	
EGM_CHR.dbf	This table describes the national character sets used for	
	each language.	
SHN_NAM.dbf	The table includes the names of the units of all	
	administrative levels.	

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata

```python
GeoPackage = sg.PopupGetFile('Please Select the GeoPackage:')
if GeoPackage != "":
    conn = sqlite3.connect(GeoPackage)
    sg.Popup(GeoPackage, 'Database Connection Successful')
else:
    sg.Popup(GeoPackage, 'Database Connection Failed','Program Stopped')
    sys.exit("Error")

print ("Database Connection Successful")

cursor = conn.cursor()
```

Set up initial connection with Geopackage

```python
##Create gpkg_metadata table
cursor.execute('DROP TABLE IF EXISTS gpkg_metadata')
cursor.execute('''''CREATE TABLE gpkg_metadata (
  id INTEGER CONSTRAINT m_pk PRIMARY KEY ASC NOT NULL,
  md_scope TEXT NOT NULL DEFAULT 'dataset',
  md_standard_uri TEXT NOT NULL,
  mime_type TEXT NOT NULL DEFAULT 'text/xml',
  metadata TEXT NOT NULL DEFAULT ''
);''')
conn.commit()
##Create gpkg_metadata_reference table
cursor.execute('DROP TABLE IF EXISTS gpkg_metadata_reference')
cursor.execute('''''CREATE TABLE gpkg_metadata_reference (
  reference_scope TEXT NOT NULL,
  table_name TEXT,
  column_name TEXT,
  row_id_value INTEGER,
  timestamp DATETIME NOT NULL DEFAULT (strftime('%Y-%m-%dT%H:%M:%fZ','now')),
  md_file_id INTEGER NOT NULL,
  md_parent_id INTEGER,
  CONSTRAINT crmr_mfi_fk FOREIGN KEY (md_file_id) REFERENCES gpkg_metadata(id),
  CONSTRAINT crmr_mpi_fk FOREIGN KEY (md_parent_id) REFERENCES gpkg_metadata(id)
);''')
conn.commit()
```

Create metadata tables based on OGC GeoPackage Encoding Standard

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata

- Inserting general metadata information of KM-1000 to GeoPackage

```
1.  ##Path Location of Aktualitaetsstand
2.  LatestInfo = sg.PopupGetFile('Please Select Aktualitaetsstand.txt file in Unzipped KM-
    1000V Folder:')
3.  if LatestInfo !="":
4.      sg.Popup(LatestInfo, 'Import Successful')
5.  else:
6.      sg.Popup(LatestInfo, 'Import Failed','Program Stopped')
7.      sys.exit("Error")
8.
9.  lines = []
10.
11. file = open(LatestInfo, 'rt',encoding="latin-1")
12.
13. for line in file:
14.     lines.append(line)
15.
16. ReleaseDate = lines[3] + lines[5]
17. print (ReleaseDate)
18.
19.
20. ##Extract General Structure Information of KM-1000
21. ExtractGeneralStructure = "PDFPage22.txt"
22. file = open(ExtractGeneralStructure, 'r',encoding='utf-8')
23. content = file.read()
24. GeneralStructurePattern = r'2.4.1.1  About KM1000-V (.*\n.*\n.*\n.*\n.*) .*'
25. test = re.search(GeneralStructurePattern, content, re.MULTILINE)
26. GeneralStructure = str(test.group())
27. GeneralInformation = ReleaseDate + GeneralStructure
28. print (GeneralInformation)
29.
30. ##Append general metadata infotmation to gpkg
31. metadata = [1,'dataset','NA', 'text', GeneralInformation]
32. print (metadata)
33. cursor.execute('insert into gpkg_metadata values (?,?,?,?,?)', metadata)
34. conn.commit()
35. print ('''''FINISHED Extracting Text from Aktualiaetsstand and PDF Page 22
36. FINISHED Adding general metadata infotmation to gpkg''')
37.
38. file.close()
39. if os.path.exists("PDFPage22.txt"):
40.   os.remove("PDFPage22.txt")
41. else:
42.   print("The file does not exist")
```

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata

```python
##add description of layer to gpkg for KM1000_polbnd_area
FeatureClassName = "KM1000_polbnd_area"
Definition = "An area controlled by administrative authority."
EGM_Feature_Class = "PolbndA"
FeatureType = "Area"
PrimitiveType = "Face"
PortrayalCriteria = "Each administrative unit consists of one main area and occasionally of one main area with exclave(s). Exclaves bigger than 3 km² included. If a country has national administrative levels below a country level, then the lowest level in EU-countries is a level equivalent to NUTS3 level and in other countries the lowest level is comparable to this level."
attributeTable = "KM1000_polbnd_area_07_2018_description_of_attributes"
table = [1,FeatureClassName, Definition, EGM_Feature_Class, FeatureType, PrimitiveType, PortrayalCriteria]
cursor.execute('DROP TABLE IF EXISTS KM1000_polbnd_area_07_2018_description')
cursor.execute('''''CREATE TABLE KM1000_polbnd_area_07_2018_description (
        id INTEGER CONSTRAINT m_pk PRIMARY KEY ASC NOT NULL,
        featureClassName TEXT NOT NULL,
        definition TEXT NOT NULL,
        EGMFeatureClass TEXT NOT NULL,
        featureType TEXT NOT NULL,
        primitiveType TEXT NOT NULL,
        portrayalCriteria TEXT NOT NULL
        );''')
cursor.execute('INSERT INTO KM1000_polbnd_area_07_2018_description VALUES (?,?,?,?,?,?,?)', table)
conn.commit()
print ("FINISHED Inserting description for KM1000_polbnd_area")

##add description of attribute table to gpkg for KM1000_polbnd_area
skiprows = list(range(9)) + list(range(17,26))
DescriptionOfAttributes = pd.read_csv("tablepage2627.csv",skiprows = skiprows, usecols = [0,2,3],encoding='latin1')
cursor.execute('DROP TABLE IF EXISTS KM1000_polbnd_area_07_2018_description_of_attributes')
DescriptionOfAttributes.to_sql("KM1000_polbnd_area_07_2018_description_of_attributes",conn, if_exists='append',index=False)
conn.commit()
print ("FINISHED Inserting description of attributes for KM1000_polbnd_area")
```

KM1000_polbnd_area

Administrative area

Definition: An area controlled by administrative authority.
EGM - Feature class: PolbndA
Feature type: Area
Primitive type: Face
Portrayal criteria: Each administrative unit consists of one main area and occasionally of one main area with exclave(s). Exclaves bigger than 3 km² included. If a country has national administrative levels below a country level, then the lowest level in EU-countries is a level equivalent to NUTS3 level and in other countries the lowest level is comparable to this level.

Layer Description

Attribute	Definition	Value/Code or Example Value description
FCODE	FACC feature code	FA001 Administrative area
TAA	Type of the administrative area	0 Unknown 1 Mainland 3 Exclave or island 4 Condominium 7 Water only
SHN0	Id-code of country-level (ISO 3166 Nation Code + number of zeros, so that fields SHN0 – SHN4 have equal width).	FI000000 *(Example)* XXYY000 *(Example)* For in dispute areas between countries XX and YY
SHN1	ID Code of 1st order administrative unit.	FI600000 *(Example)* N_A Not applicable (if country has no more than the country level in EGM)

KM50, KM250, KM500, KM1000 - Vektor Version **1.5** Seite 26 von 33

Description of Attribute Table

BEV - Bundesamt für Eich- und Vermessungswesen

SHN2	ID Code of 2nd order administrative unit.	FI108000 *(Example)* N_A Not applicable (if country has no more than the 1st order national level in EGM)
SHN3	Id-code of the 3rd order administrative unit.	DE01005300000 (Example) N_A Not applicable (if country has no more than the 2nd order national level in EGM)
SHN4	Id-code of the 4th order administrative unit.	GB11QL0000 *(Example)* N_A Not applicable (if country has no more than the 3rd order national level in EGM)

Metadata Overview of KM1000_polbnd_area Layer in BEV PDF

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata

Table: KM1000_polbnd_area_07_2018_description

New Record | Delete Record | Mode: Text | Import | Export | Set as NULL

	id	eatureClassNam	definition	:GMFeatureClass	featureType	primitiveType	portrayalCriteria
	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	1	KM1000_polb...	An area controlled by ...	PolbndA	Area	Face	Each administrative unit consists of one main area and occasi...

Each administrative unit consists of one main area and occasionally of one main area with exclave(s). Exclaves bigger than 3 km² included. If a country has national administrative levels below a country level, then the lowest level in EU-countries is a level equivalent to NUTS3 level and in other countries the lowest level is comparable to this level.

Overview of KM1000_polbnd_area_07_2018_description

Table: KM1000_polbnd_area_07_2018_description_of_attributes

New Record | Delete Record | Mode: Text | Import

	Attribute	Definition	Value/Code or Example Value description
	Filter	Filter	Filter
1	FCODE	FACC feature code	FA001 Administrative area
2	TAA	Type of the administrative area	0 Unknown1 Mainland3 Exclave or island4 Condominium7 Water only
3	SHN0	Id-code of country-level (ISO 3166 NationC...	FI000000 (Example)XXYY000 (Example) For in dispute areas between countriesXX and YY
4	SHN1	ID Code of 1st order administrative unit.	FI600000 (Example)N_A Not applicable (if country has no more thanthe country level in EGM)
5	SHN2	ID Code of 2nd order administrative unit.	FI108000 (Example)N_A Not applicable (if country has no more than the 1st ordernational level in EGM)
6	SHN3	Id-code of the 3rd order administrative unit.	DE01005300000 (Example)N_A Not applicable (if country has no more than the 2ndordernational level in EGM)
7	SHN4	Id-code of the 4th order administrative unit.	GB11QL0000 (Example)N_A Not applicable (if country has no more than the 3rd ordernational level in EGM)

0 Unknown
1 Mainland
3 Exclave or island
4 Condominium
7 Water only

Overview of KM1000_polbnd_area_07_2018_description_of_attribute

DB Schema

Name | Ty

- KM1000_airfld_point_07_2018
- KM1000_airfld_point_07_2018_description
- KM1000_airfld_point_07_2018_description_of_attributes
- KM1000_builtup_area_07_2018
- KM1000_builtup_area_07_2018_description
- KM1000_builtup_area_07_2018_description_of_attributes
- KM1000_builtup_point_07_2018
- KM1000_builtup_point_07_2018_description
- KM1000_builtup_point_07_2018_description_of_attributes
- KM1000_dam_line_07_2018
- KM1000_dam_line_07_2018_description
- KM1000_dam_line_07_2018_description_of_attributes
- KM1000_elev_point_07_2018
- KM1000_elev_point_07_2018_description
- KM1000_elev_point_07_2018_description_of_attributes
- KM1000_glacier_area_07_2018
- KM1000_glacier_area_07_2018_description
- KM1000_glacier_area_07_2018_description_of_attributes
- KM1000_island_area_07_2018
- KM1000_island_area_07_2018_description
- KM1000_island_area_07_2018_description_of_attributes
- KM1000_lake_area_07_2018
- KM1000_lake_area_07_2018_description
- KM1000_lake_area_07_2018_description_of_attributes
- KM1000_name_point_07_2018
- KM1000_name_point_07_2018_description
- KM1000_name_point_07_2018_description_of_attributes
- KM1000_polbnd_area_07_2018
- KM1000_polbnd_area_07_2018_description
- KM1000_polbnd_area_07_2018_description_of_attributes
- KM1000_polbnd_line_07_2018
- KM1000_polbnd_line_07_2018_description
- KM1000_polbnd_line_07_2018_description_of_attributes
- KM1000_railrd_line_07_2018
- KM1000_railrd_line_07_2018_description
- KM1000_railrd_line_07_2018_description_of_attributes
- KM1000_railrd_node_07_2018
- KM1000_railrd_node_07_2018_description
- KM1000_railrd_node_07_2018_description_of_attributes
- KM1000_reservoir_area_07_2018
- KM1000_reservoir_area_07_2018_description
- KM1000_reservoir_area_07_2018_description_of_attributes
- KM1000_road_line_07_2018
- KM1000_road_line_07_2018_description
- KM1000_road_line_07_2018_description_of_attributes
- KM1000_spring_node_07_2018
- KM1000_spring_node_07_2018_description
- KM1000_spring_node_07_2018_description_of_attributes
- KM1000_watrcrs_area_07_2018

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata – INSPIRE Metadata Support

```
1.  ##insert KM1000_polbnd_line XML to gpkg
2.  ##polbnd_line_xml = "KM1000AU__d2b8d67f-737c-4d49-b220-ca0ef422197d.xml"
3.  polbnd_line_xml = sg.PopupGetFile('Please Select Downloaded XML File:')
4.  if LatestInfo !="":
5.      sg.Popup(polbnd_line_xml, 'Import Successful')
6.  else:
7.      sg.Popup(polbnd_line_xml, 'Import Failed','Program Stopped')
8.      sys.exit("Error")
9.
10. polbnd_line_xml_copy = "polbnd_line_xml_copy.xml"
11. copyfile(polbnd_line_xml, polbnd_line_xml_copy)
12.
13. xmlformatpreserved = xml.dom.minidom.parse(polbnd_line_xml_copy)
14. xmlformatpreserved = xmlformatpreserved.toprettyxml()
15. xmlformatpreserved = str(xmlformatpreserved)
16.
17.
18. metadata = [2,'feature','http://www.isotc211.org/2005/gmd', 'xml', xmlformatpreserved]

19. cursor.execute('insert into gpkg_metadata values (?,?,?,?,?)', metadata)
20. conn.commit()
21. print ("Finishing inserting xml to GeoPackage")
22.
23. timestamp = datetime.datetime.now()
24. metadata_reference = ['table','KM1000_POLBND_LINE_07_2018', '', '',
25.                       timestamp, 2, 1]
26. cursor.execute('insert into gpkg_metadata_reference values (?,?,?,?,?,?,?)', metadata_r
    eference)
27. conn.commit()
28.
29. def replace_in_config(old, new):
30.     with open(polbnd_line_xml_copy, 'r') as f:
31.         text = f.read()
32.
33.     with open(polbnd_line_xml_copy, 'w') as f:
34.         f.write(text.replace(old, new))
35.
36. replace_in_config('gco:', '')
37. replace_in_config('gmd:', '')
```

```
38.
39.
40. tree = ElementTree.parse(polbnd_line_xml_copy)
41. root = tree.getroot()
42.
43. extractxml = ""
44.
45. for child in root.iter():
46.     tagstring = str(child.tag)
47.     textstring = str(child.text)
48.     extractxml += tagstring + textstring
49.
50. extractxml = str(extractxml)
51.
52. metadata = [3,'feature','http://www.isotc211.org/2005/gmd', 'text', extractxml]
53. cursor.execute('insert into gpkg_metadata values (?,?,?,?,?)', metadata)
54. conn.commit()
55.
56. timestamp = datetime.datetime.now()
57. metadata_reference = ['table','KM1000_POLBND_LINE_07_2018', '', '',
58.                       timestamp, 3, 1]
59. cursor.execute('insert into gpkg_metadata_reference values (?,?,?,?,?,?,?)', metadata_r
    eference)
60. conn.commit()
61.
62. print ("Finishing inserting xml to GeoPackage")
63. if os.path.exists("polbnd_line_xml_copy.xml"):
64.     os.remove("polbnd_line_xml_copy.xml")
65. else:
66.     print("The file does not exist")
```

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata – INSPIRE Metadata Support


```
<?xml version="1.0" ?>
<gmd:MD_Metadata xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:gmd="http://www.isotc211.org/2005/gmd" xmlns:gml="http://www.opengis.net/gml" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.isotc211.org/2005/gmd http://schemas.opengis.net/iso/19139/20060504/gmd/gmd.xsd">

        <gmd:fileIdentifier>

                <gco:CharacterString>d2b8d67f-737c-4d49-b220-ca0ef422197d</gco:CharacterString>

        </gmd:fileIdentifier>

        <gmd:language>

                <gmd:LanguageCode codeList="http://www.loc.gov/standards/iso639-2/" codeListValue="ger"/>

        </gmd:language>
```


Type of data currently in cell: Text / Numeric
21261 char(s)

DB Schema

- HillShade50m
- KM1000_airfld_point_07_2018
- KM1000_airfld_point_07_2018_description
- KM1000_airfld_point_07_2018_description_of_attributes
- KM1000_builtup_area_07_2018
- KM1000_builtup_area_07_2018_description
- KM1000_builtup_area_07_2018_description_of_attributes
- KM1000_builtup_point_07_2018
- KM1000_builtup_point_07_2018_description
- KM1000_builtup_point_07_2018_description_of_attributes

Database Structure | Browse Data | Edit Pragmas | Execute SQL

Table: gpkg_metadata New Record Delete R

	id	md_scope	md_standard_ur	mime_type	metadata
	Filter	Filter	Filter	Filter	Filter
1	1	dataset	NA	text	Produkt: Kartographisches Modell - Vektor 1:1 000 000 (KM1000-V)Aktualitätsstand (Stichtag) der enthaltenen Daten: 07.2018
2	2	feature	http://www.is...	xml	<?xml version="1.0" ?><gmd:MD_Metadata xmlns:gco="http://www.isotc211.org/2005/gco" xmlns:gmd="http://www.isotc2...
3	3	feature	http://www.is...	text	MD_Metadata fileIdentifier

3. Methodology

3.2 Overview of Approach

3.2.4 Inserting Metadata – INSPIRE Metadata Support

Mode: Text

```
MD_Metadata
 fileIdentifier
    CharacterStringd2b8d67f-737c-4d49-b220-ca0ef422197dlanguage
    LanguageCodeNonehierarchyLevel
    MD_ScopeCodeNonecontact
    CI_ResponsibleParty
      organisationName
        CharacterStringBundesamt für Eich- und VermessungswesencontactInfo
        CI_Contact
          address
            CI_Address
              electronicMailAddress
                CharacterStringkundenservice@bev.gv.atrole
        CI_RoleCodeNonedateStamp
    DateTime2018-05-16T16:00:20metadataStandardName
    CharacterStringISO19115metadataStandardVersion
    CharacterString2003/Cor.1:2006referenceSystemInfo
    MD_ReferenceSystem
      referenceSystemIdentifier
        RS_Identifier
          code
            CharacterStringEPSG:3416codeSpace
            CharacterStringhttp://www.opengis.net/def/crs/EPSG/0/3416identificationInfo
    MD_DataIdentification
```

Type of data currently in cell: Text / Numeric
7858 char(s)

DB Schema

Name
- KM1000_watrcrs_line_07_2018
- KM1000_watrcrs_line_07_2018_description
- KM1000_watrcrs_line_07_2018_description_of_attributes
- KM500_R_WALD_07_2018
- gpkg_contents
- gpkg_data_column_constraints
- gpkg_data_columns
- gpkg_extensions
- gpkg_geometry_columns
- gpkg_metadata

Database Structure | Browse Data | Edit Pragmas | Execute SQL

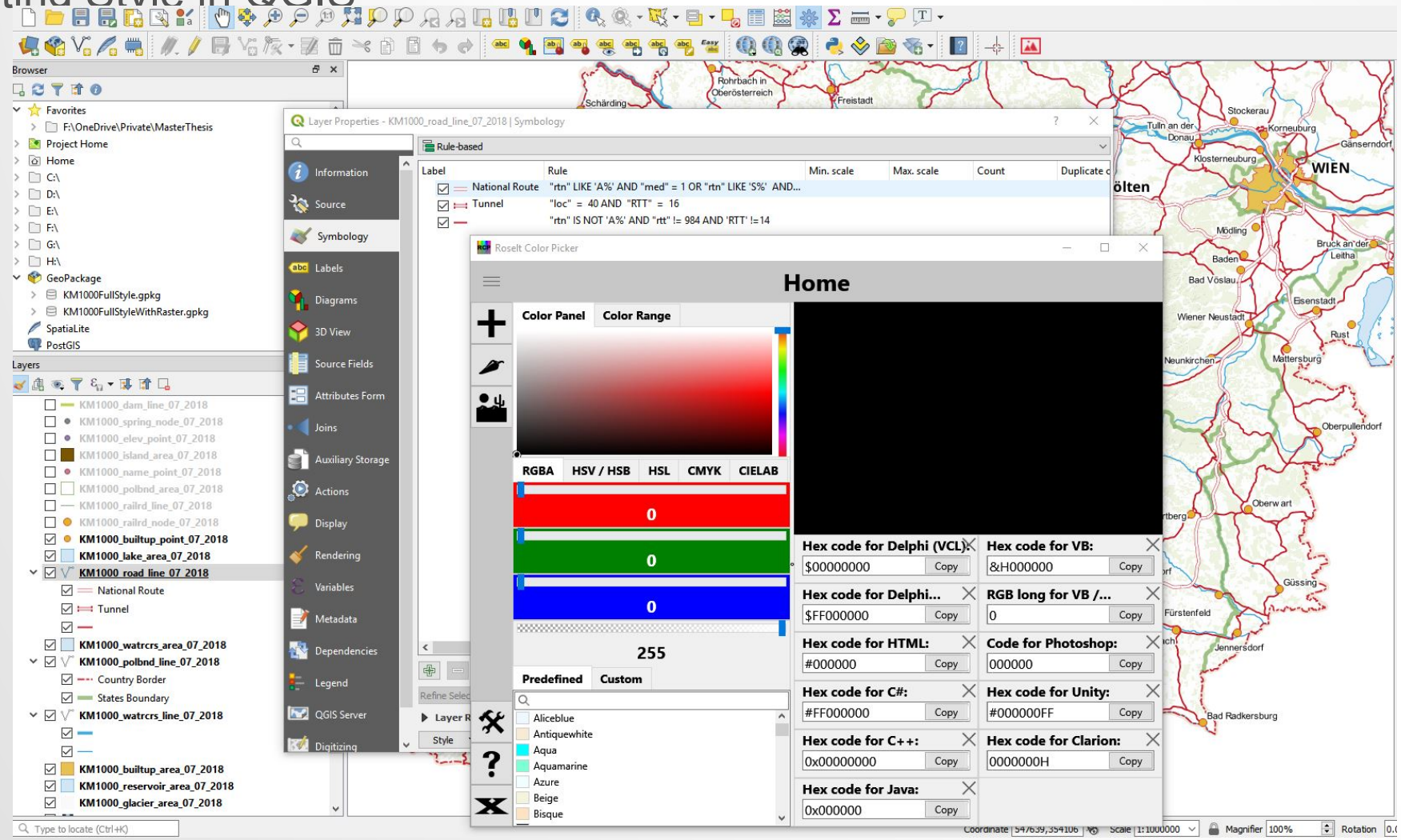
Table: gpkg_metadata

	id	md_scope	md_standard_uri	mime_type	metadata
	Filter	Filter	Filter	Filter	Filter
1	1	dataset	NA	text	Produkt: Kartographisches Modell - Vektor 1:1 000 000 (KM1000-V)Aktualitätsstand (Stichta...
2	2	feature	http://www.isotc211.org/2005/gmd	xml	<?xml version="1.0" ?><gmd:MD_Metadata xmlns:gco="http://www.isotc211.org/2005/gc...
3	3	feature	http://www.isotc211.org/2005/gmd	text	MD_Metadata fileIdentifier

3. Methodology

3.2 Overview of Approach

3.2.5 Editing Style in QGIS



3. Methodology

3.2 Overview of Approach

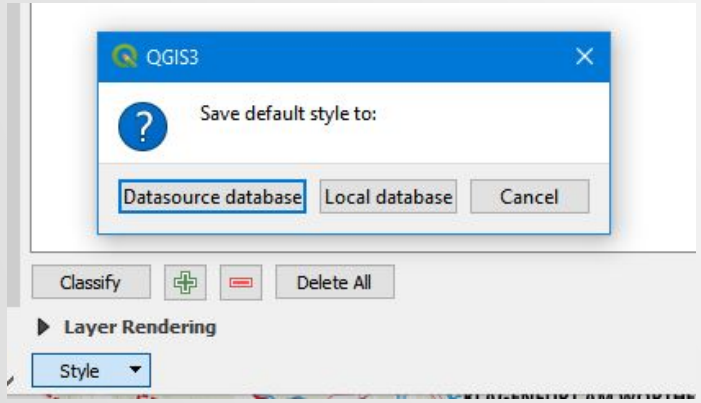
3.2.5 Editing Style in QGIS

Additional Cartographic Representations created to match the National Map Style

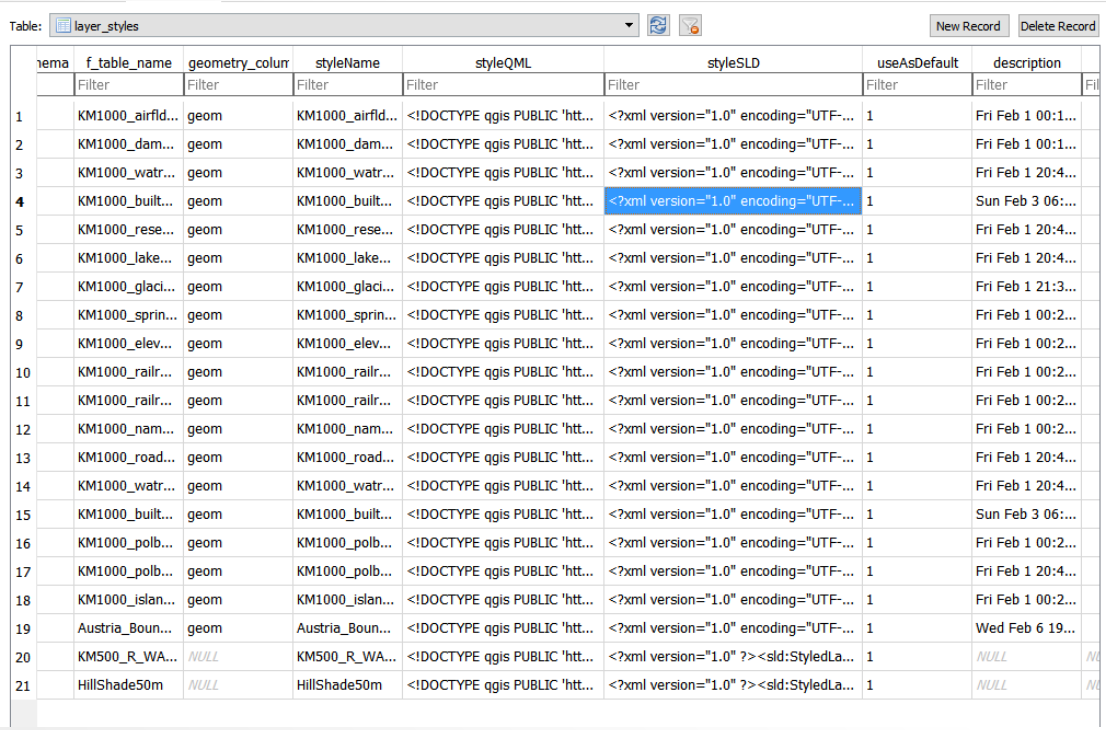
3. Methodology

3.2 Overview of Approach

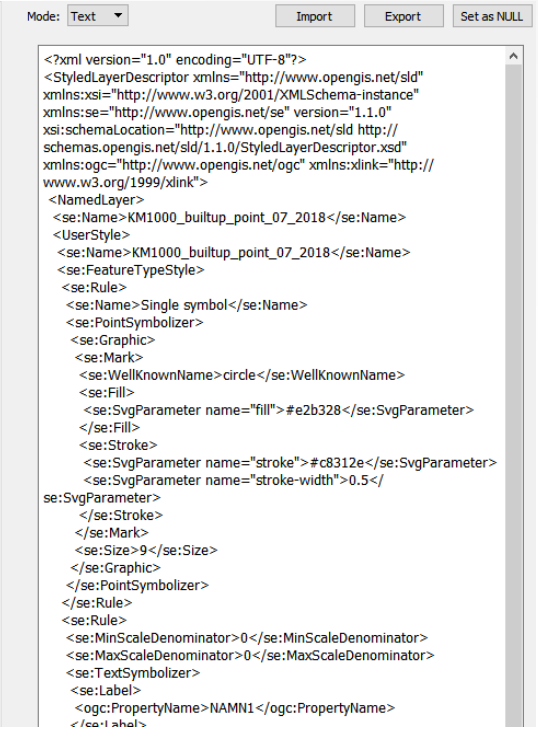
3.2.6 Adding Style into GeoPackage



Add Layer Style to GeoPackage in QGIS



	f_table_name	geometry_colum	styleName	styleQML	styleSLD	useAsDefault	description
1	KM1000_airfld...	geom	KM1000_airfld...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:1...
2	KM1000_dam...	geom	KM1000_dam...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:1...
3	KM1000_watr...	geom	KM1000_watr...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 20:4...
4	KM1000_built...	geom	KM1000_built...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Sun Feb 3 06:...
5	KM1000_rese...	geom	KM1000_rese...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 20:4...
6	KM1000_lake...	geom	KM1000_lake...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 20:4...
7	KM1000_glaci...	geom	KM1000_glaci...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 21:3...
8	KM1000_sprin...	geom	KM1000_sprin...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
9	KM1000_elev...	geom	KM1000_elev...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
10	KM1000_railr...	geom	KM1000_railr...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
11	KM1000_railr...	geom	KM1000_railr...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
12	KM1000_nam...	geom	KM1000_nam...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
13	KM1000_road...	geom	KM1000_road...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 20:4...
14	KM1000_watr...	geom	KM1000_watr...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 20:4...
15	KM1000_built...	geom	KM1000_built...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Sun Feb 3 06:...
16	KM1000_polb...	geom	KM1000_polb...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
17	KM1000_polb...	geom	KM1000_polb...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 20:4...
18	KM1000_islan...	geom	KM1000_islan...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Fri Feb 1 00:2...
19	Austria_Boun...	geom	Austria_Boun...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" encoding="UTF-...	1	Wed Feb 6 19...
20	KM500_R_WA...	NULL	KM500_R_WA...	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" ?><sld:StyledLa...	1	NULL
21	HillShade50m	NULL	HillShade50m	<!DOCTYPE qgis PUBLIC 'htt...	<?xml version="1.0" ?><sld:StyledLa...	1	NULL




```
<?xml version="1.0" encoding="UTF-8"?>
<StyledLayerDescriptor xmlns="http://www.opengis.net/sld"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:se="http://www.opengis.net/se" version="1.1.0"
xsi:schemaLocation="http://www.opengis.net/sld http://
schemas.opengis.net/sld/1.1.0/StyledLayerDescriptor.xsd"
xmlns:ogc="http://www.opengis.net/ogc" xmlns:xlink="http://
www.w3.org/1999/xlink">
 <NamedLayer>
  <se:Name>KM1000_builtup_point_07_2018</se:Name>
  <UserStyle>
   <se:Name>KM1000_builtup_point_07_2018</se:Name>
   <se:FeatureTypeStyle>
    <se:Rule>
     <se:Name>Single symbol</se:Name>
     <se:PointSymbolizer>
      <se:Graphic>
       <se:Mark>
        <se:WellKnownName>circle</se:WellKnownName>
        <se:Fill>
         <se:SvgParameter name="fill">#e2b328</se:SvgParameter>
        </se:Fill>
        <se:Stroke>
         <se:SvgParameter name="stroke">#c8312e</se:SvgParameter>
         <se:SvgParameter name="stroke-width">0.5</
se:SvgParameter>
        </se:Stroke>
       </se:Mark>
       <se:Size>9</se:Size>
      </se:Graphic>
     </se:PointSymbolizer>
    </se:Rule>
    <se:Rule>
     <se:MinScaleDenominator>0</se:MinScaleDenominator>
     <se:MaxScaleDenominator>0</se:MaxScaleDenominator>
     <se:TextSymbolizer>
      <se:Label>
       <ogc:PropertyName>NAMN1</ogc:PropertyName>
```

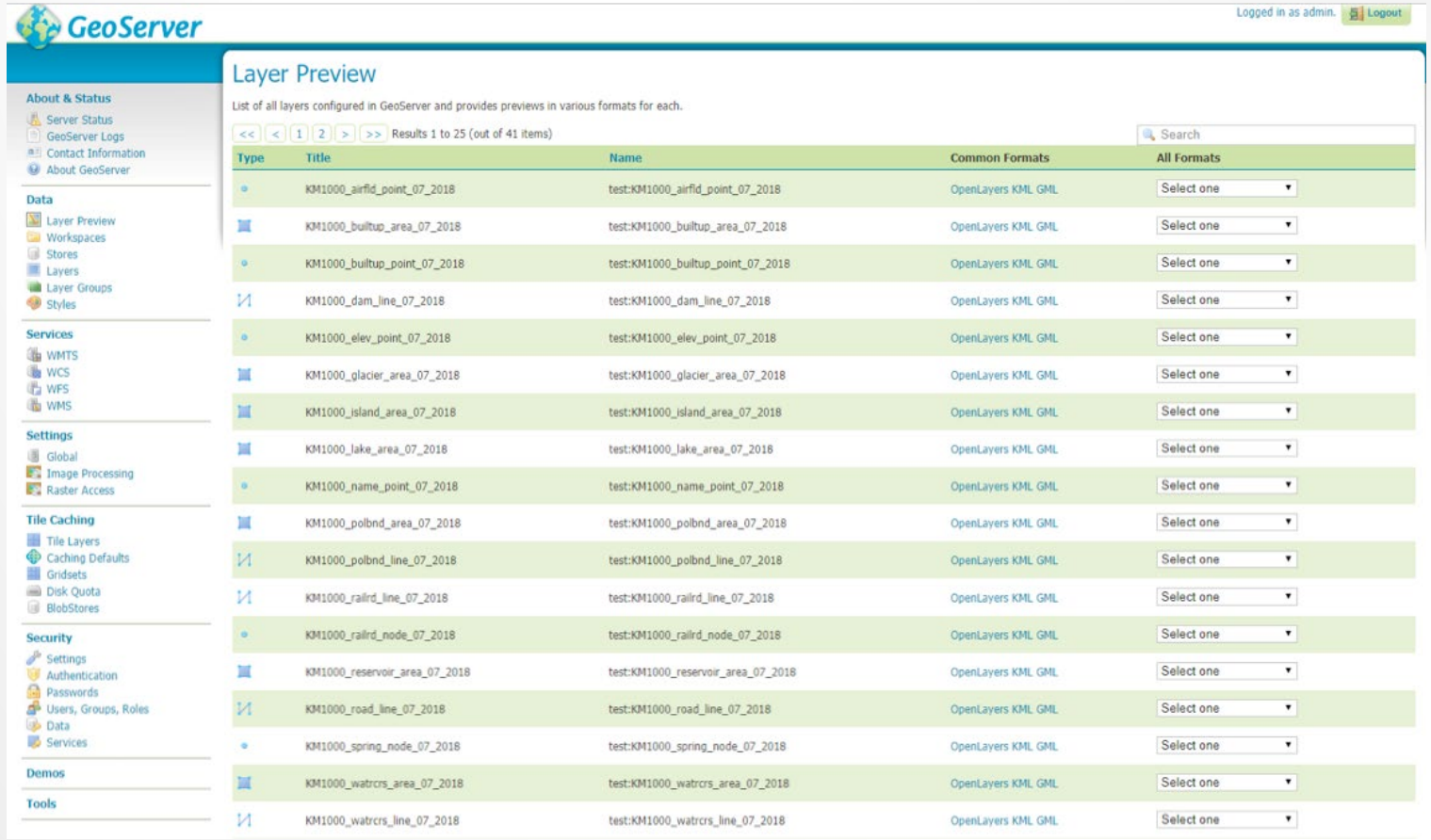
3. Methodology

3.2 Overview of Approach

3.2.7 Uploading GeoPackage to Geoserver

1. Austria_Boundary_Buffer_Effect
2. HillShade50m
3. KM500_R_WALD_07_2018
4. KM1000_glacier_area_07_2018
5. KM1000_reservoir_area_07_2018
6. KM1000_builtup_area_07_2018
7. KM1000_watrcrs_line_07_2018
8. KM1000_polbnd_line_07_2018
9. KM1000_watrcrs_area_07_2018
10. KM1000_road_line_07_2018
11. KM1000_lake_area_07_2018
12. KM1000_builtup_point_07_2018

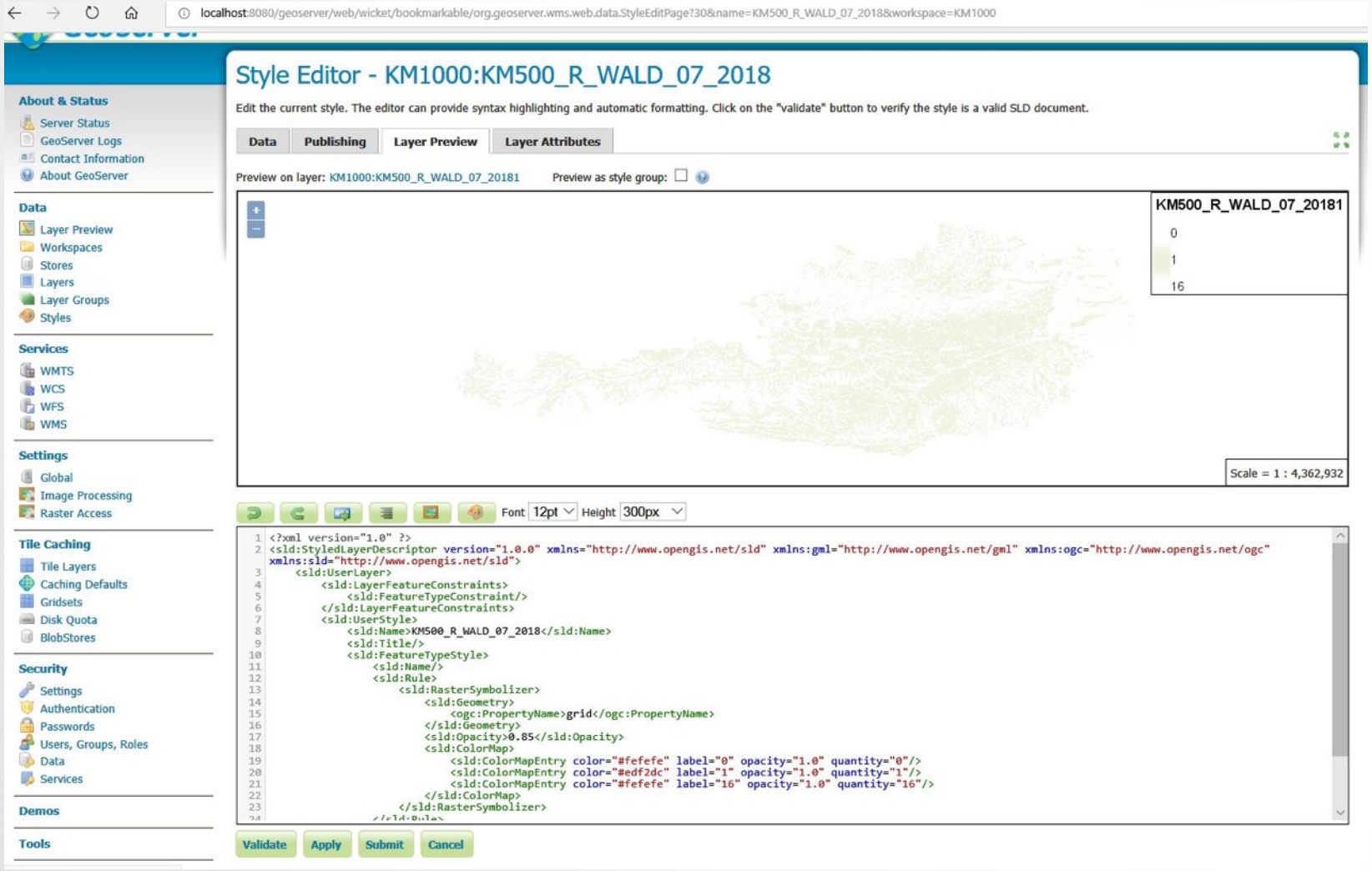
Layer Overlay Order



3. Methodology

3.2 Overview of Approach

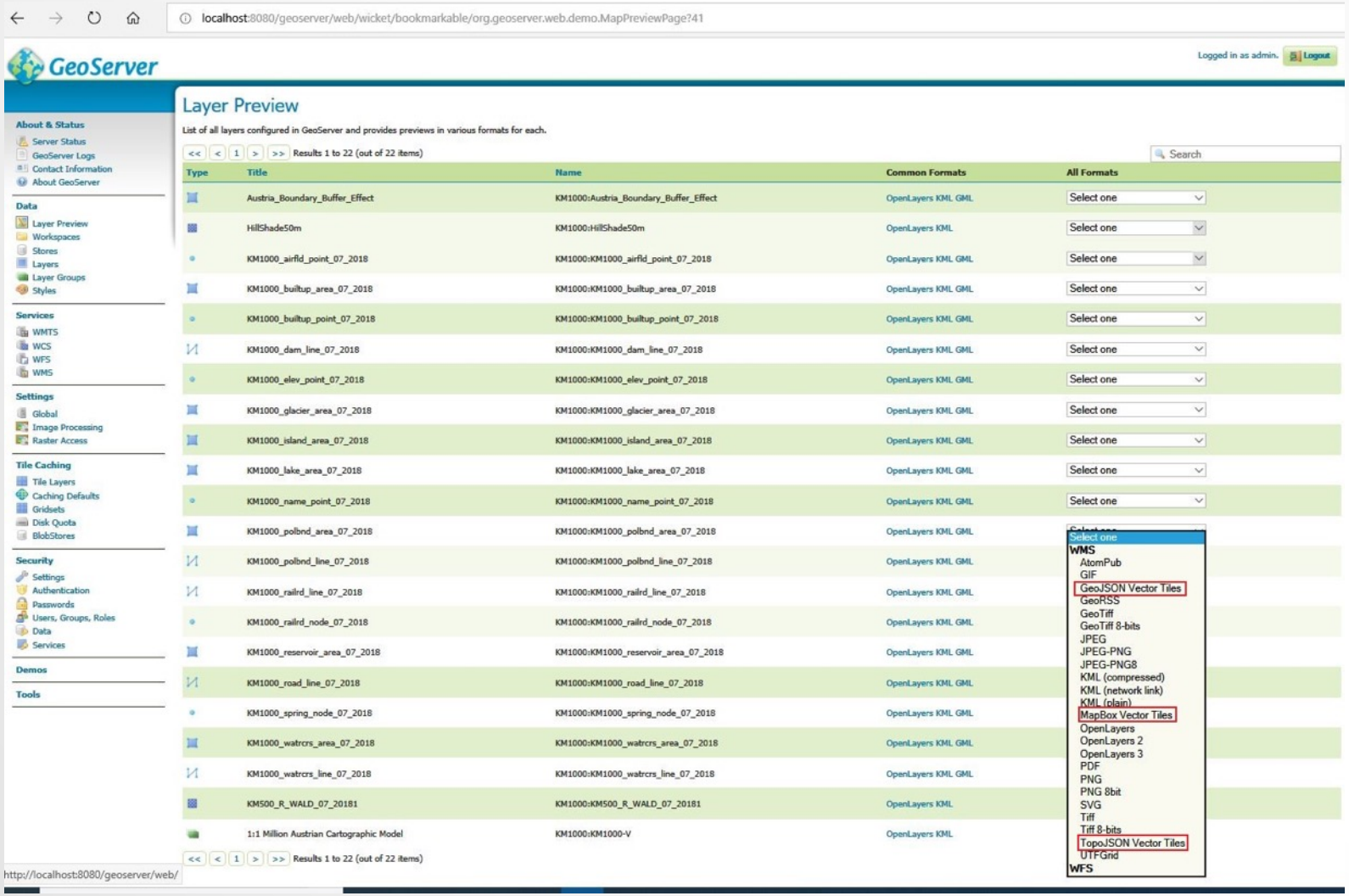
3.2.8 Importing Style and Digital Maps Distribution



3. Methodology

3.2 Overview of Approach

3.2.8 Importing Style and Digital Maps Distribution



4. Results

- Archiving Digital Maps with GeoPackage

Name	Size	Allocated	Files	Folders	% of Parent (Allocated)
486.6 MB F:\MasterThesis\FinalResult\Size	**486.5 MB**	**486.6 MB**	**15**	**1**	**100.0 %**
430.9 MB KM1000-V Raster	**430.9 MB**	**430.9 MB**	**12**	**0**	**88.6 %**
55.6 MB [3 Files]	55.6 MB	55.6 MB	3	0	11.4 %
36.4 MB KM1000FullStyleWithRaster.gpkg	36.4 MB	36.4 MB	1	0	65.5 %
15.2 MB KM1000FullStyle.gpkg	15.2 MB	15.2 MB	1	0	27.3 %
4.0 MB KM1000-V.zip	4.0 MB	4.0 MB	1	0	7.2 %

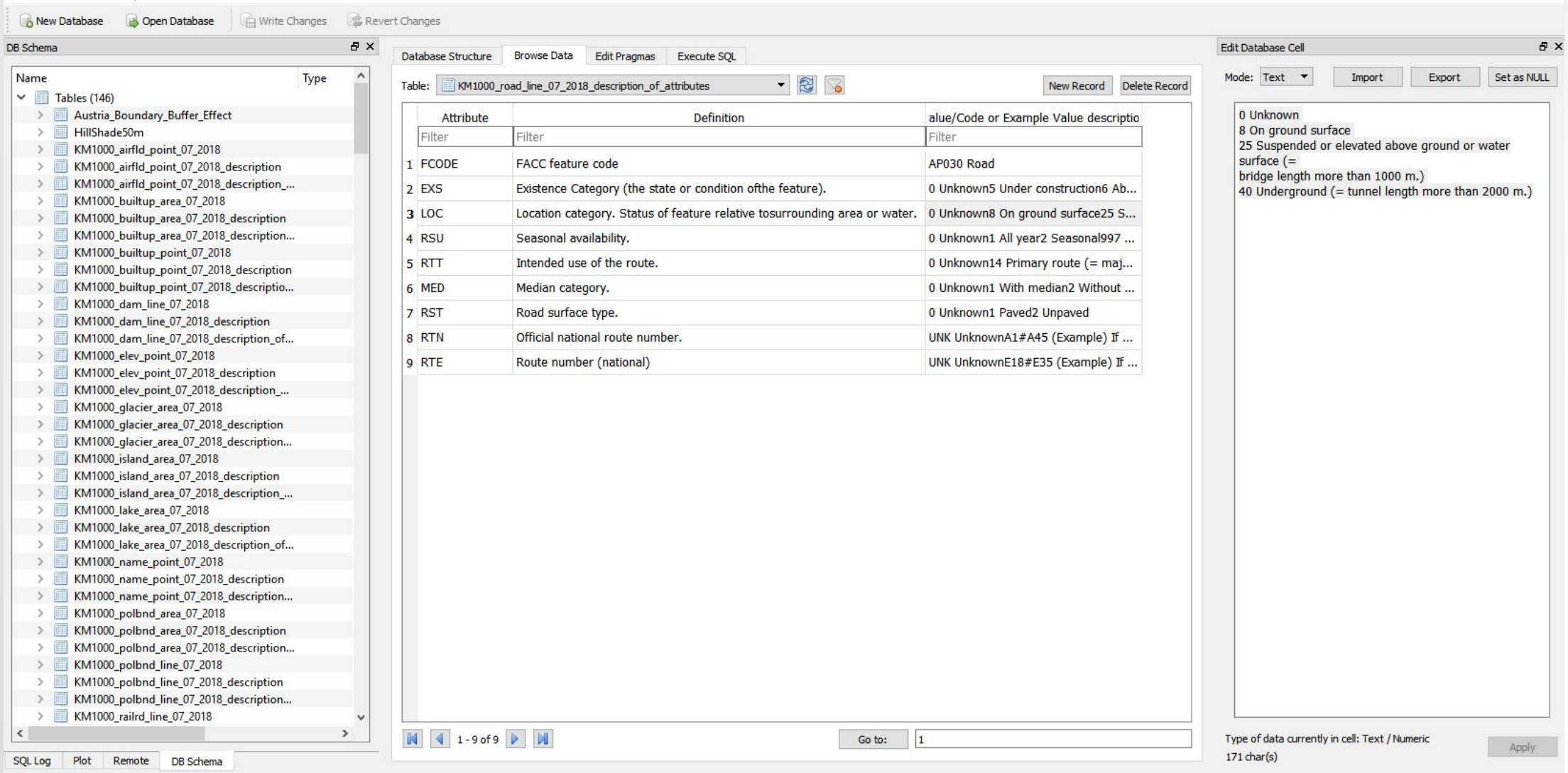
Data Size Comparison

4. Results

- Archiving Digital Maps with GeoPackage

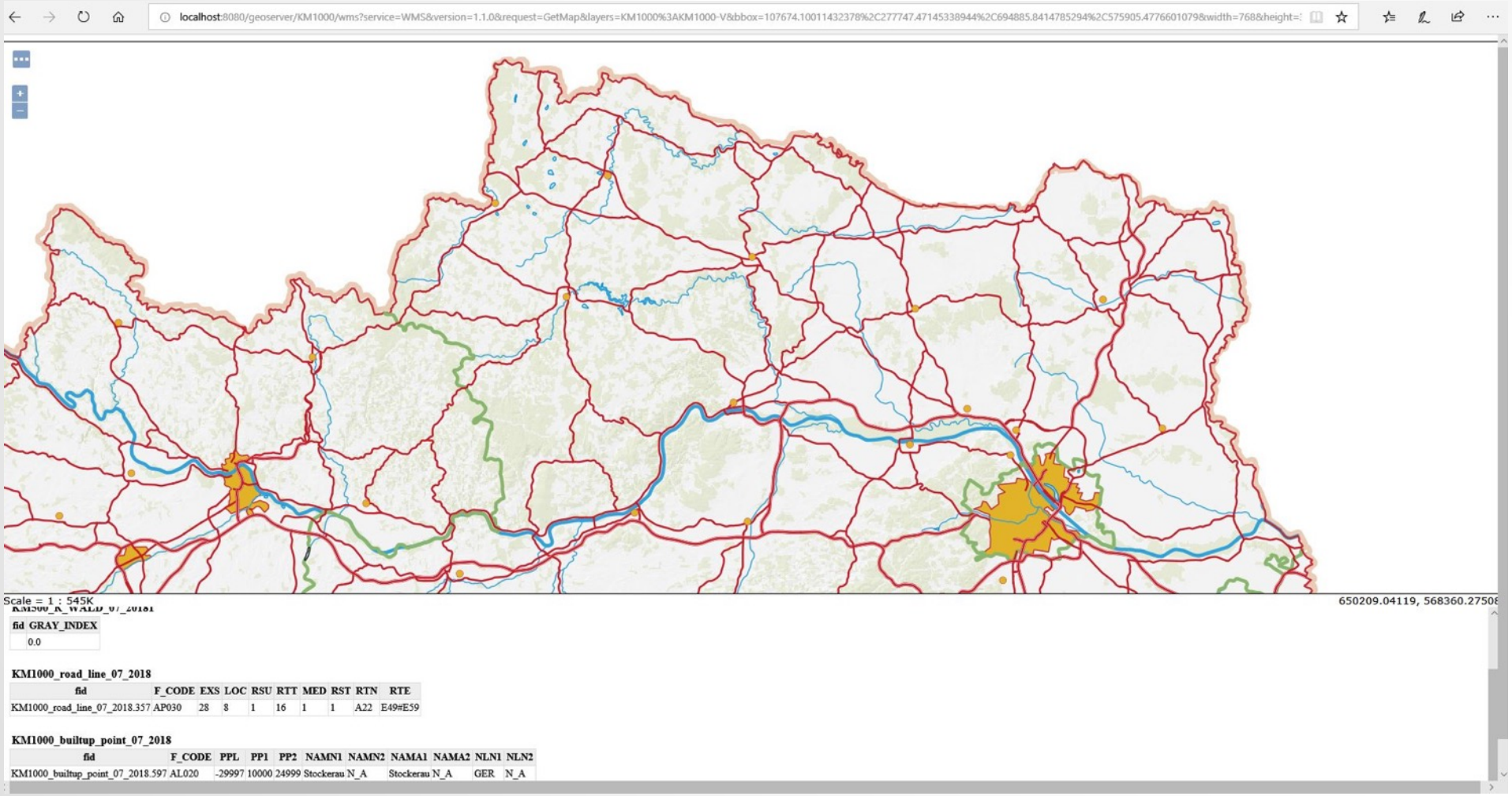
4. Results

- Archiving Digital Maps with GeoPackage



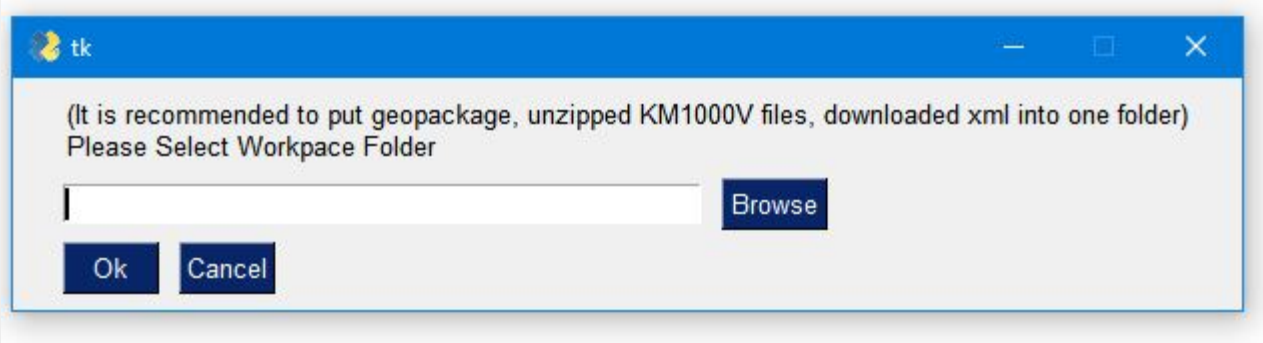
4. Results

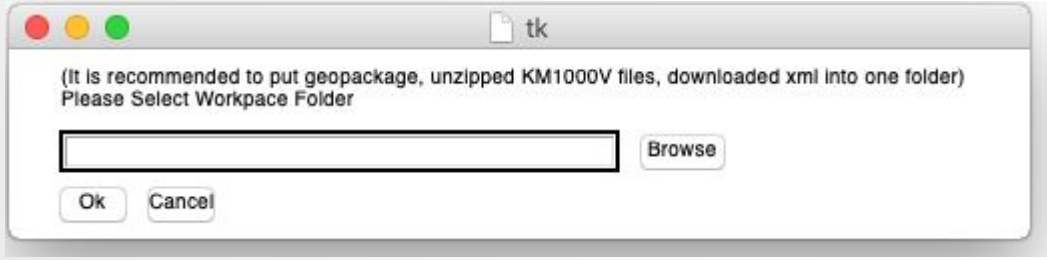
- Vector Tiles Dissemination



4. Results

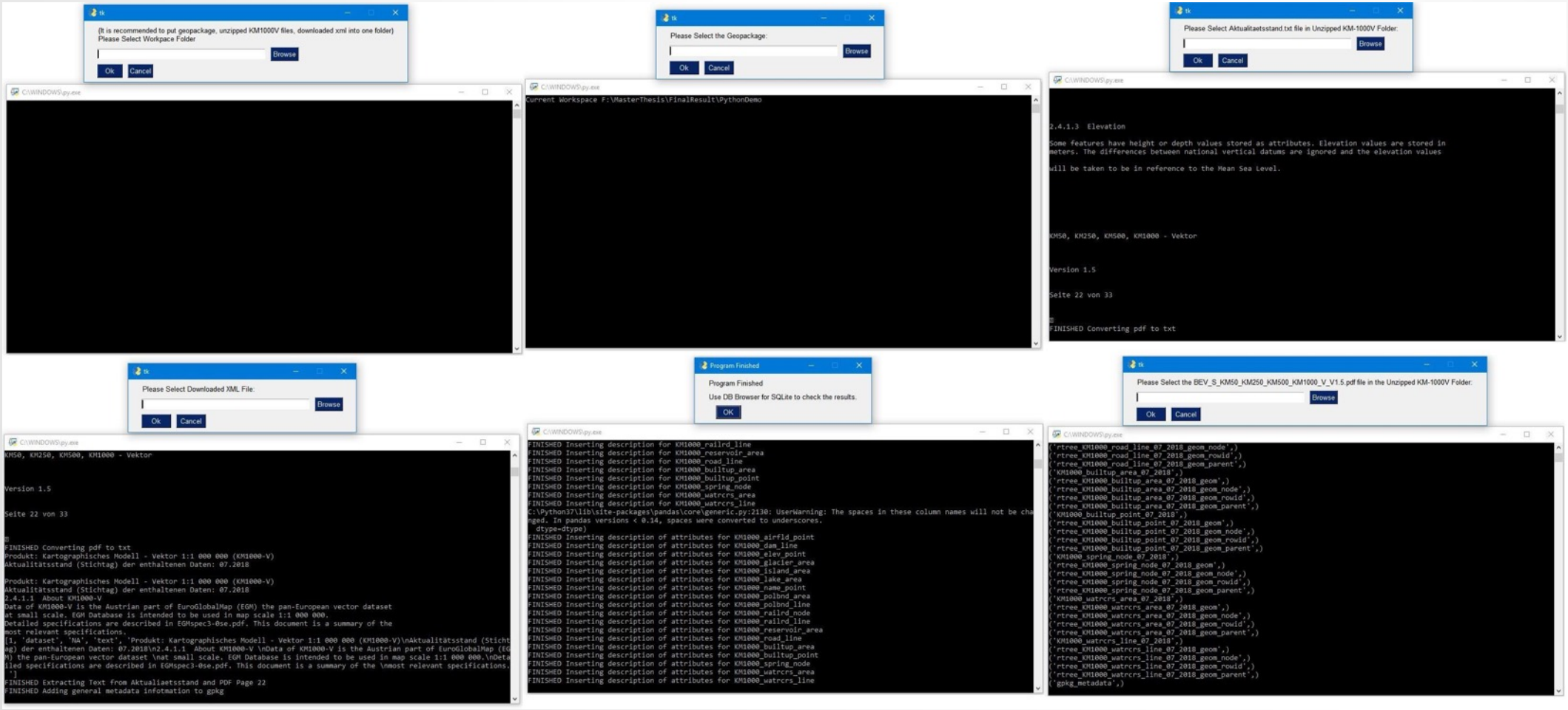
- Python Script for Automatically Embedding Metadata





4. Results

- Python Script for Automatically Embedding Metadata



5. Discussion and Future Work

- Discussion

  - Limited research concentrates on the disadvantage of existing geospatial data formats or proposes an alternative format for archiving digital maps.
  - Limited research related to GeoPackage format.
  - The default metadata table provided in GeoPackage Encoding Standard was not sufficient.
  - The method designed for the case study provides a completely free, open-source, and cross-platform solution for archiving and dissemination of digital maps.
  - This approach uses SLD format to store style information of feature layers and GeoServer to disseminate digital maps.
  - The Python script developed for this approach is exclusively designed to extract the metadata in the PDF file of BEV KM1000-V datasets.

5. Discussion and Future Work

- Character Encoding Issue

NAMN1	NAMN2	NAMA1	NAM
M²hldorf	N_A	Muehldorf	N_A
Roith	N_A	Roith	N_A
Oberweis	N_A	Oberweis	N_A
Schalchham	N_A	Schalchham	N_A
Fuschl am See	N_A	Fuschl am See	N_A
Glasenbach	N_A	Glasenbach	N_A
Plainfeld	N_A	Plainfeld	N_A
Vigaun	N_A	Vigaun	N_A
Oberndorf	N_A	Oberndorf	N_A
Endach	N_A	Endach	N_A
B²hlerwerk	N_A	Boehlerwerk	N_A
Ertl	N_A	Ertl	N_A
Reichraming	N_A	Reichraming	N_A

German Umlauts Encoding Issue

KM1000-V datasets are encoded in ISO 8859-1 (Latin -1)

5. Discussion and Future Work

- Map Labeling and Symbology Issue
- Geodata Quality Issue

BEV National Map

Final Map of this research

5. Discussion and Future Work

- Future Work

  - More detailed research is needed to study benefit and drawback of different geospatial data formats used in archiving digital maps.

  - Further research is needed to study metadata extension and set new standard for metadata tables in Geopackage.

  - Further research is needed to examine how well proprietary GIS software and other open-source GIS software support a Geopackage that contains metadata and style information.

  - It is important to study the compatibility of SLD format with other web map platforms and the conversion capability between SLD format and other style formats.

  - Future work should focus on the proper method to export and share the label placement in QGIS

6. Conclusion

- This case study, archiving the 1:1 million Austria Cartographic Model with national map style, demonstrates the possibility of using GeoPackage to embed relevant metadata from different data formats that the digital maps requires, such as general information for the whole dataset, layer description, description of attribute table, symbology, and reference system information.

- Through the last three stages in this research approach, the goal of distributing digital maps through vector-tile format has been achieved by uploading the embedded SLD style from GeoPackage to Geoserver.

- The Python script has significantly reduced the workload of embedding metadata into a GeoPackage.

In sum, this thesis proposes an innovative approach to embed relevant metadata for future usage in digital maps using GeoPackage and to embed styling from GeoPackage when distributing digital maps that use vector-tile format. It also shows the possibility of using Python scripts to automate the procedure of embedding relevant information into GeoPackage.

7. Reference

- Bleakly, D. R. (2002). Long-Term Spatial Data Preservation and Archiving: What Are the Issues? doi:10.2172/793225
- Lauriault T.P., Pulsifer P.L., Taylor D.F. (2011). The Preservation and Archiving of Geospatial Digital Data: Challenges and Opportunities for Cartographers. In: Jobst M. (eds) *Preservation in Digital Cartography. Springer, Berlin, Heidelberg*
- Shaon, Arif & Naumann, Kai & Kirstein, Michael & Roensdorf, Carsten & Mason, Paul & Bos, Marguérite & Gerber, Urs & Woolf, Andrew & Samuelsson, Göran. (2011). Long-term sustainability of spatial data infrastructures: a metadata framework and principles of geo-archiving.
- Locher, Anita-E. & Termens, Miquel(2012). *Exploring alternatives for geodata preservation*, *In 7ª Conferencia Ibérica de Sistemas y Tecnologías de Información*. CISTI 2012, Madrid
- Clark, John H (2016). The Long-Term Preservation of Digital Historical Geospatial Data: A Review of Issues and Methods, *Journal of Map & Geography Libraries*, 12:2, 187-201, DOI: 10.1080/15420353.2016.1185497
- Jobst, Markus & Gartner, Georg (2011). Structural Aspect for the Digital Cartographic Heritage. In: Jobst M. (eds) *Preservation in Digital Cartography. Springer, Berlin, Heidelberg*
- Erwin, Tracey & Sweetkind-Singer, Julie (2009) The National Geospatial Digital Archive: A Collaborative Project to Archive Geospatial Data, *Journal of Map & Geography Libraries*, 6:1, 6-25, DOI: 10.1080/15420350903432440
- North Carolina Center for Geographic Information and Analysis & North Carolina Department of Cultural Resources (2011). Geospatial Multistate Archive and Preservation Partnership Final Report
- Bogossian, C.H., Ferreira, K.R., Monteiro, A.M., & Vinhas, L. (2014). A Hybrid Architecture for Mobile Geographical Data Acquisition and Validation Systems. *GeoInfo*.
- Open Geospatial Consortium (2018) OGC GeoPackage Encoding Standard – With Corrigendum. Retrieved from https://www.opengeospatial.org/standards/GeoPackage

Thank you very much for your time!