OGC GeoPackage – nový formát pro výměnu prostorových dat

Jáchym Čepický
OpenGeoLabs s.r.o.
GIVS 2018

Shapefile: prokletí současné geoinformatiky

Jáchym Čepický
OpenGeoLabs s.r.o.
GIVS 2018

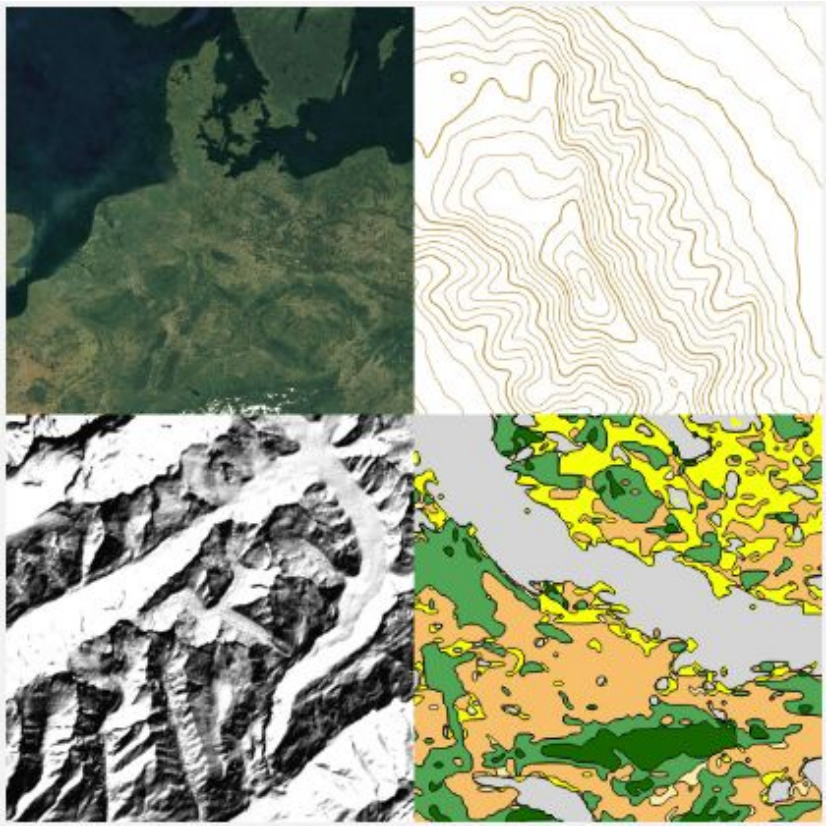
(ESRI) Shapefile

ESRI Shapefile Technical Description

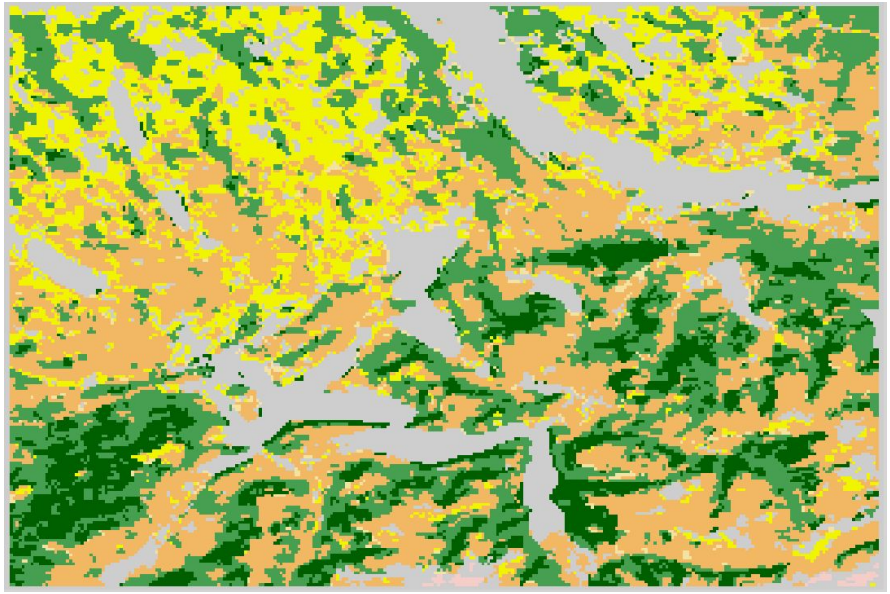
An ESRI White Paper—July 1998

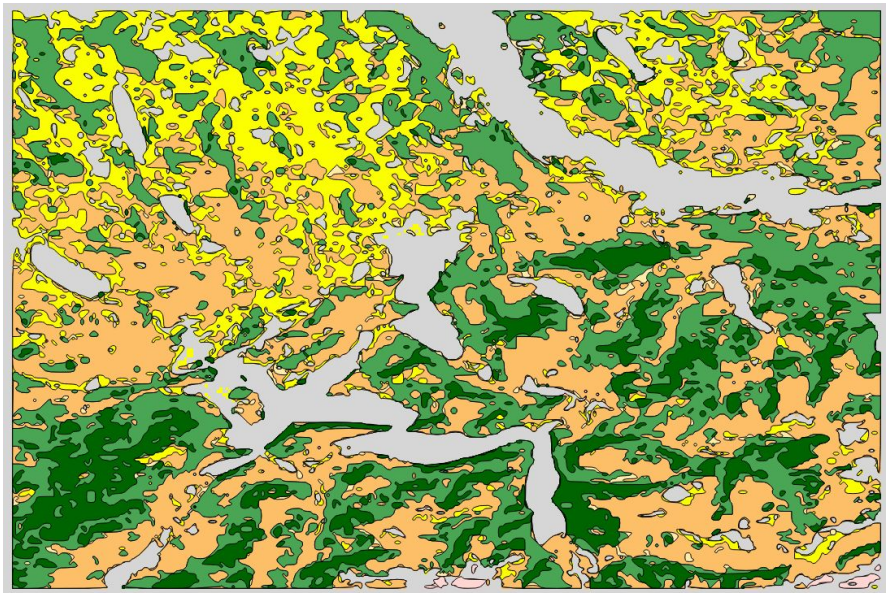
https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf

http://octogeo.cz

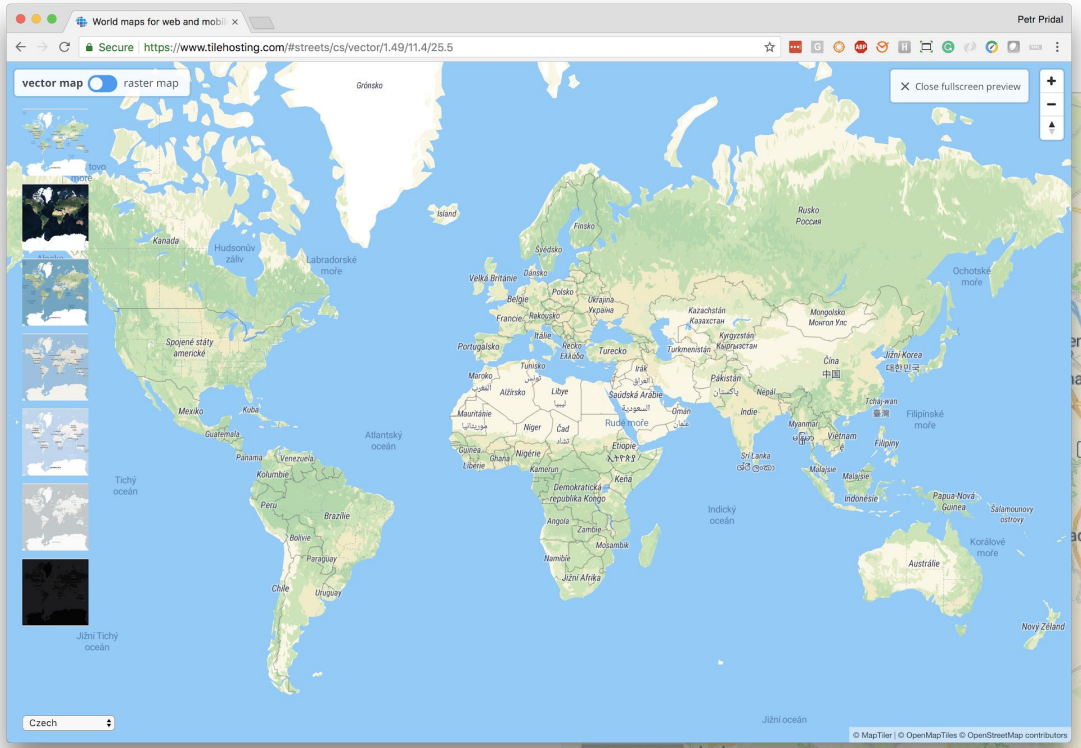


http://octogeo.org/products

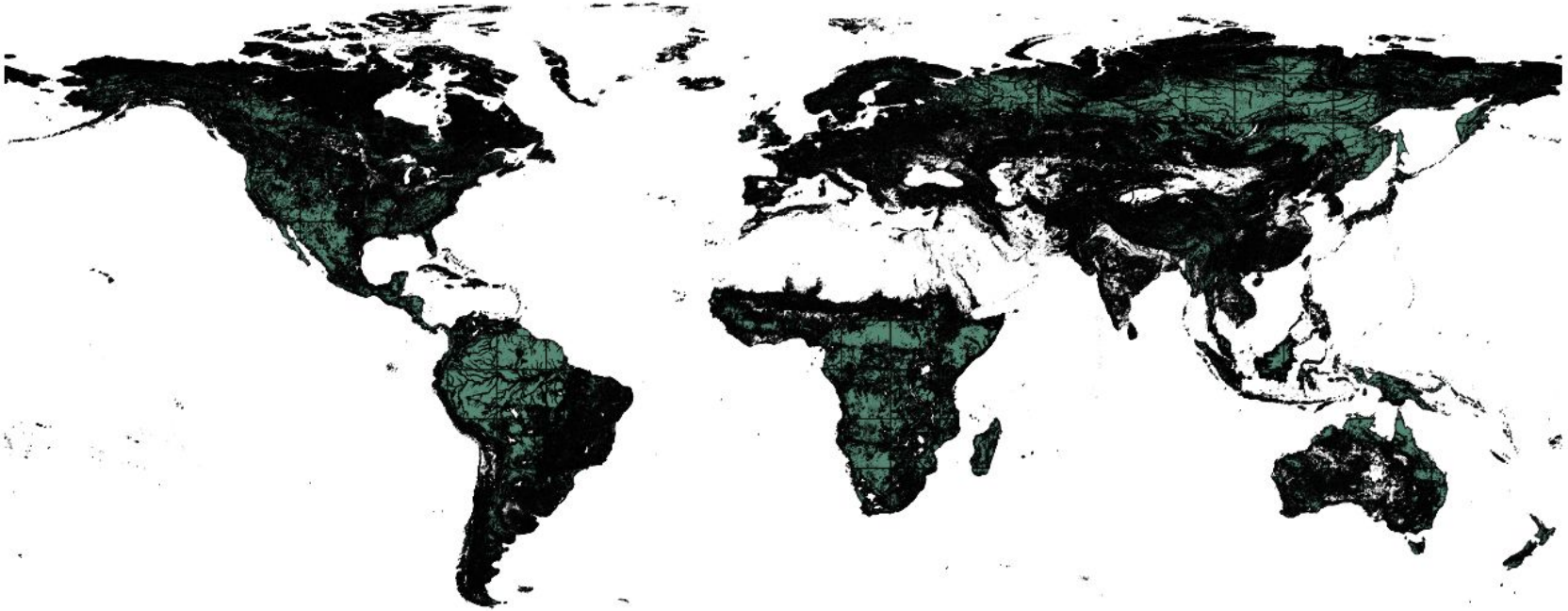




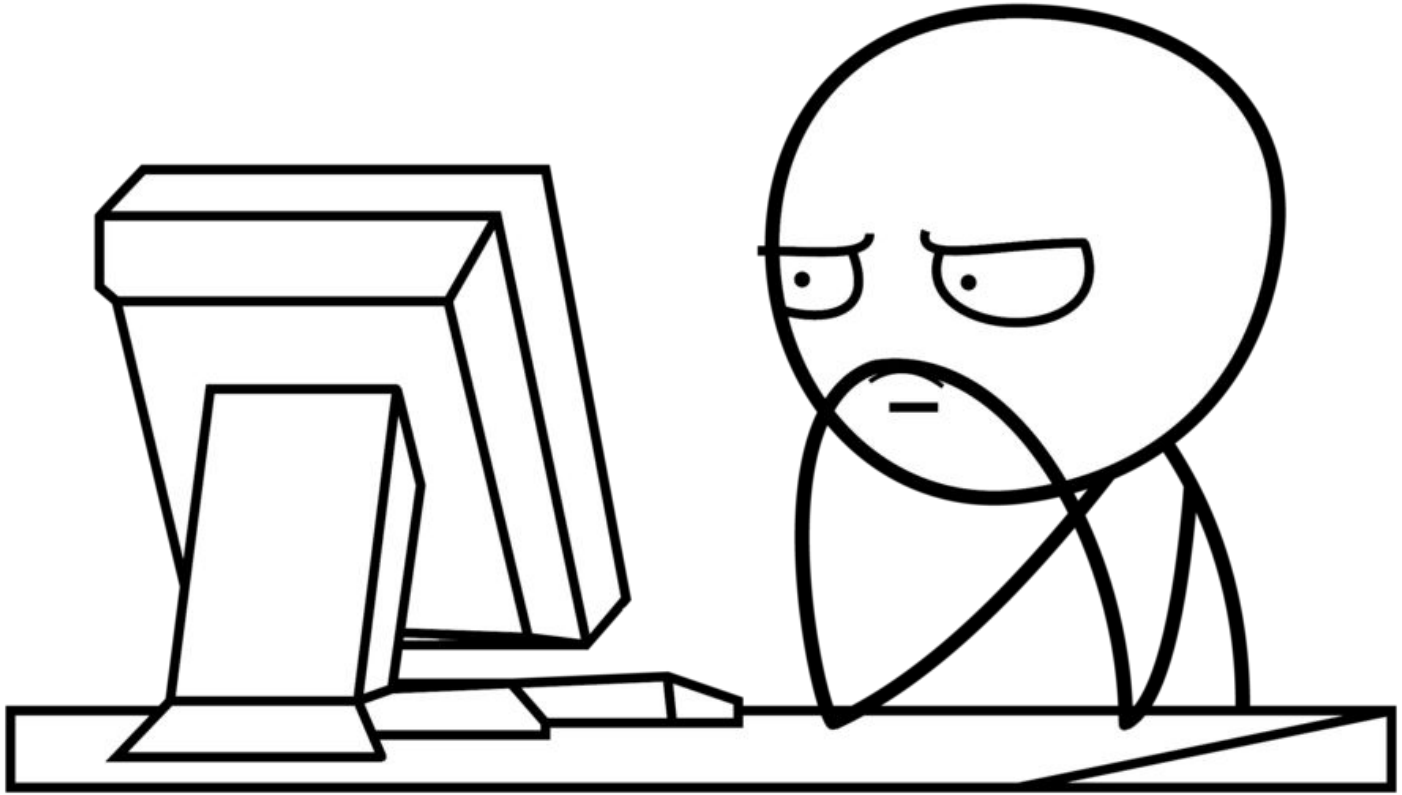
Raster → Vektor



http://tilehosting.com

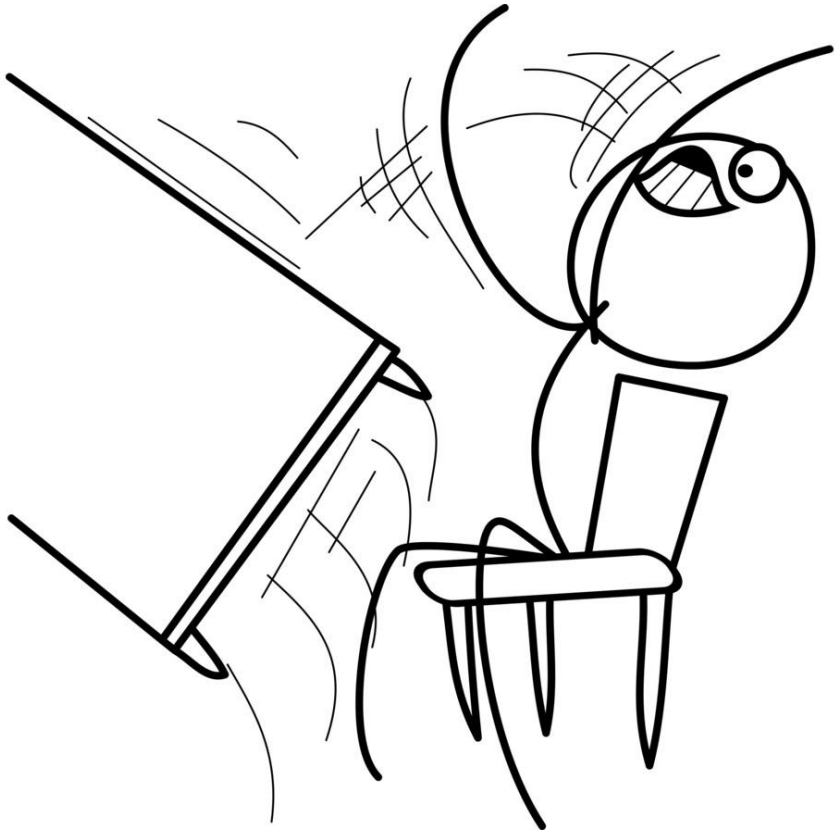


Vrstva “grass” globální mapy landuse, vektor,
cca 3 000 000 vektorových prvků



2 dny počítání…

```
ERROR 1: Failure writing DBF record 31212233.
ERROR 1: Error in psSHP->sHooks.Fwrite() while writing object
to .shp file.
ERROR 1: Terminating translation prematurely after failed
```

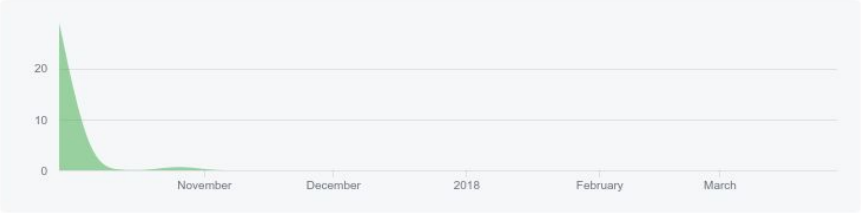


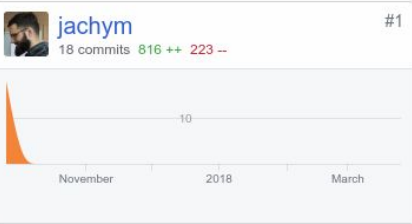
http://switchfromshapefile.org

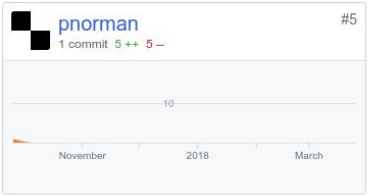
Oct 1, 2017 – Mar 31, 2018

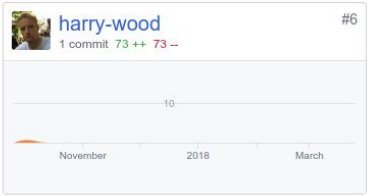
Contributions: **Commits**

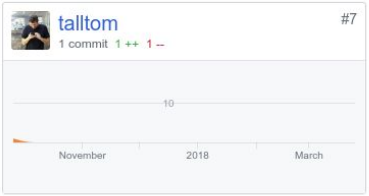
Contributions to master, excluding merge commits

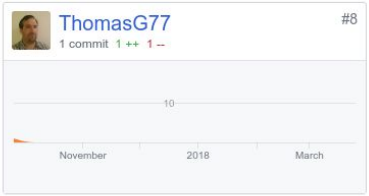


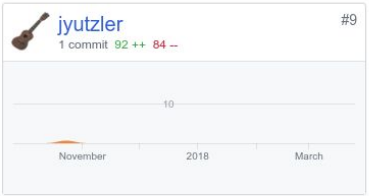


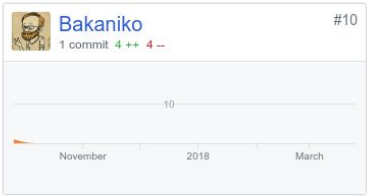












https://github.com/OpenGeoLabs/shapefilemustdie/

Shapefile - Pozitiva

Shapefile - pozitiva

Nejvíce podporovaný formát ve všech softwarech.

Shapefile - pozitiva

I když proprietární, jeho specifikace je otevřená.

Shapefile - pozitiva

For many use cases, it is good enough.

Shapefile - pozitiva

Vnitřní indexy (*.shx) umožňují optimalizovat rychlost čtení.

Shapefile - pozitiva

Co do velikosti souboru je relativně efektivní.

Shapefile - to špatné

1. Více-souborový formát.

Více-souborový formát

- shp
- shx
- dbf

Více-souborový formát

- .shp
- .shx
- .dbf
- .prj
- .sbn a .sbx
- .fbn and .fbx
- .ain a .aih
- .ixs
- .mxs
- .atx
- .shp.xml
- .cpg

2. Maximálně 10 znaků na jméno atributu

2. Maximálně 10 znaků na jméno atributu

RUIAN, KatastralniUzemi:

- Kod
- Nazev
- Nespravny
- ExistujeDigitalniMapa
- ObecKod
- PlatiOd
- PlatiDo
- IdTransakce
- GlobalniIdNavrhuZmeny
- RizeniId
- MluvnickeCharakteristikyPad2
- MluvnickeCharakteristikyPad3
- MluvnickeCharakteristikyPad4
- MluvnickeCharakteristikyPad6
- MluvnickeCharakteristikyPad7
- DatumVzniku

2. Maximálně 10 znaků na jméno atributu

RUIAN, KatastralniUzemi:

- Kod
- Nazev
- Nespravny
- ExistujeDigitalniMapa → 21
- ObecKod
- PlatiOd
- PlatiDo
- IdTransakce → 11
- GlobalniIdNavrhuZmeny → 21
- RizeniId
- MluvnickeCharakteristikyPad2 → 28
- MluvnickeCharakteristikyPad3 → 28
- MluvnickeCharakteristikyPad4 → 28
- MluvnickeCharakteristikyPad6 → 28
- MluvnickeCharakteristikyPad7 → 28
- DatumVzniku → 11

2. Maximálně 10 znaků na jméno atributu

```
Warning 6: Normalized/laundered field name: 'ExistujeDigitalniMapa' to 'ExistujeDi'

Warning 6: Normalized/laundered field name: 'IdTransakce' to 'IdTransakc'

Warning 6: Normalized/laundered field name: 'GlobalniIdNavrhuZmeny' to 'GlobalniId'

Warning 6: Normalized/laundered field name: 'MluvnickeCharakteristikyPad2' to 'MluvnickeC'

Warning 6: Normalized/laundered field name: 'MluvnickeCharakteristikyPad3' to 'Mluvnick_1'

Warning 6: Normalized/laundered field name: 'MluvnickeCharakteristikyPad4' to 'Mluvnick_2'

Warning 6: Normalized/laundered field name: 'MluvnickeCharakteristikyPad6' to 'Mluvnick_3'

Warning 6: Normalized/laundered field name: 'MluvnickeCharakteristikyPad7' to 'Mluvnick_4'

Warning 6: Normalized/laundered field name: 'DatumVzniku' to 'DatumVznik'
```

Maximálně 255 atributů v tabulce.

Plochá datová struktura.

Špatná podpora datových typů

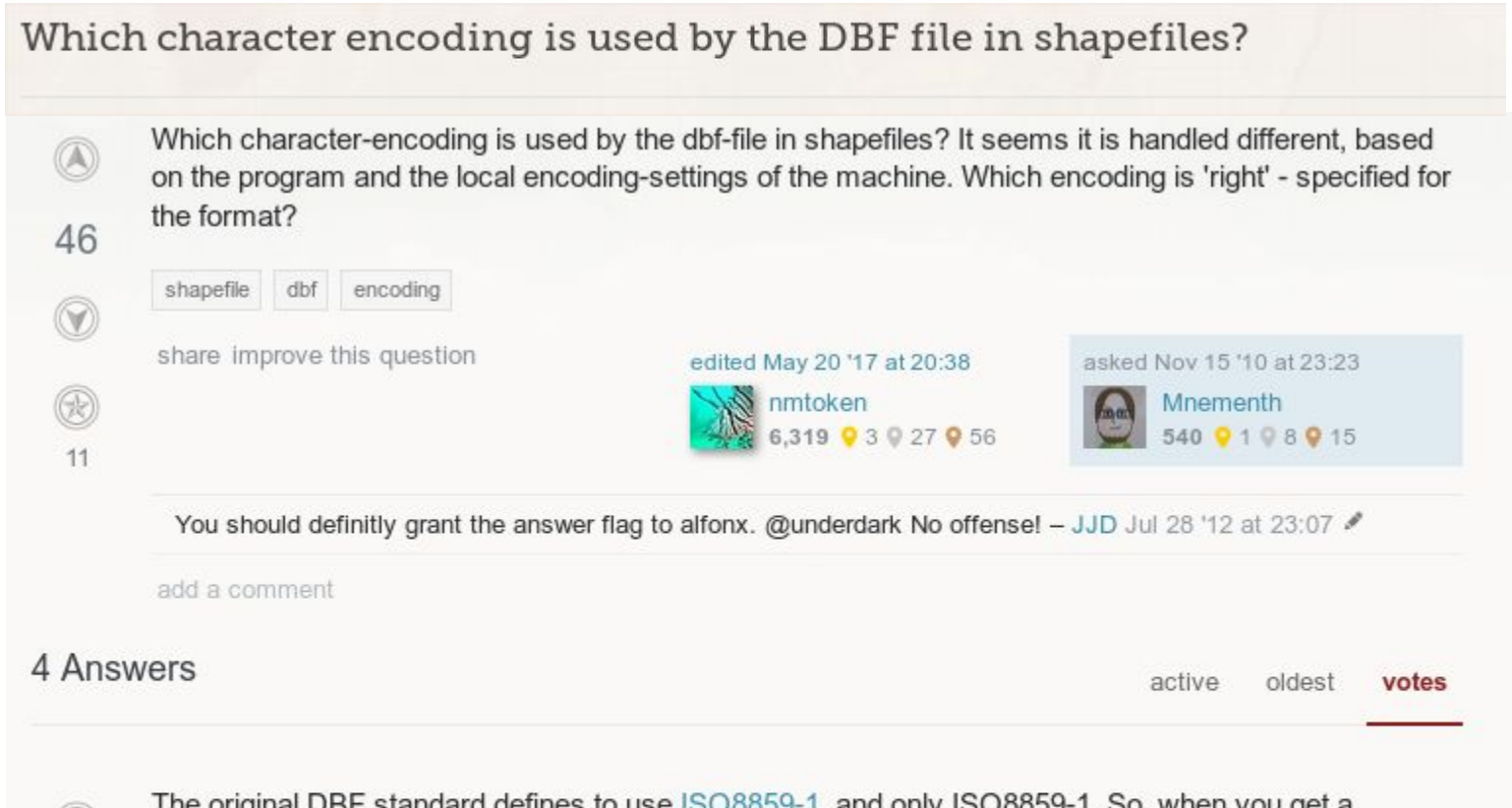
Špatná podpora datových typů

+ Float
+ Integer
+ Date
+ Character string - 255 znaků max.

- big integers
- boolean
- text
- blob
- ...

Nedefinovaná znaková sada.

Nedefinovaná znaková sada.



Which character encoding is used by the DBF file in shapefiles?

Which character-encoding is used by the dbf-file in shapefiles? It seems it is handled different, based on the program and the local encoding-settings of the machine. Which encoding is 'right' - specified for the format?

46

shapefile dbf encoding

share improve this question

edited May 20 '17 at 20:38
nmtoken
6,319 3 27 56

asked Nov 15 '10 at 23:23
Mnementh
540 1 8 15

11

You should definitly grant the answer flag to alfonx. @underdark No offense! – JJD Jul 28 '12 at 23:07

add a comment

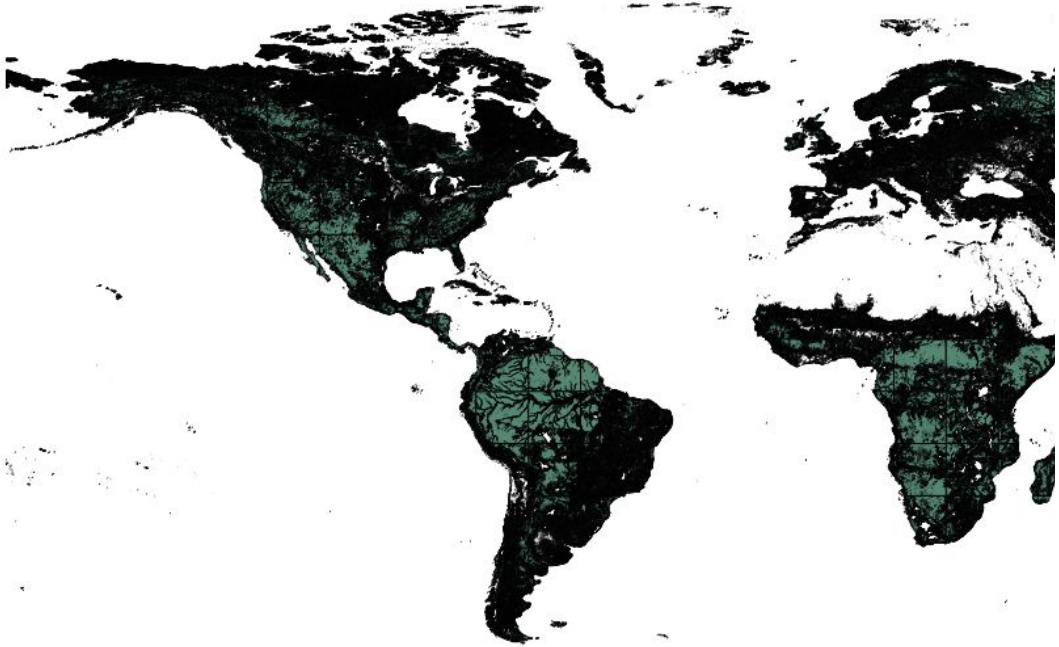
4 Answers

active oldest votes

The original DBF standard defines to use ISO8859-1, and only ISO8859-1. So, when you get a

Omezená velikost na max. 2GB.

Omezená velikost na max. 2GB.



Bez možnosti uložit topologii.

Bez možnosti uložit více geometrických typů.

Velice limitovaná podpora pro 3D.

Nekonzistentní v definici projekce.

11 hříchů Shapefilu - http://switchfromshapefile.org

Shapefile is a bad format

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

- It's a multifile format.
- Attribute names are limited to 10 characters.
- Only 255 attributes. The DBF file does not allow you to store more then 255 attribute fields.
- Limited data types. Data types are limited to float, integer, date and text with a maximum 254 characters.
- Unknown character set. There is no way to specify the character set used in the database.
- It's limited to 2GB of file size. Although some tools are able to surpass this limit, they can never exceed 4GB of data.
- No topology in the data. There is no way to describe topological relations in the format.
- Single geometry type per file. There is no way to save mixed geometry features.
- More complicated data structures are impossible to save. It's a "flat table" format.
- There is no way to store 3D data with textures or appearances such as material definitions. There is also no
  objects.
- Projections defined in the *.prj file are incompatible. Compared to definitions in EPSG / SpatialReference.org
  parameters, names and sometimes missing parameters.
- Add more ...

This repository Search Pull requests Issues Marketplace Explore

OpenGeoLabs / **shapefilemustdie**

Unwatch 7 Star 21 Fork 14

Code Issues 12 Pull requests 1 Projects 0 Wiki Insights Settings

added another item to shapefileisbad list #27

Edit

Merged **jachym** merged 1 commit into `OpenGeoLabs:master` from `jonahsullivan:patch-1` just now

Conversation 1 Commits 1 Files changed 1 +10 −7

jonahsullivan commented 2 days ago — First-time contributor

Line and polygon geometry type, single or multipart, cannot be reliably determined at the layer level, it must be determined at the individual feature level.

`added another item to shapefileisbad list` ··· Verified `0c7c42a`

jonahsullivan commented 2 days ago

https://issues.qgis.org/issues/18254

jonahsullivan closed this 2 days ago

jonahsullivan reopened this 2 days ago

Reviewers

jachym ✓

Assignees

No one—assign yourself

Labels

None yet

Projects

None yet

Milestone

No milestone

17 index.html View

```
@@ -102,13 +102,16 @@ <h2 class="alert-heading">Shapefile is a bad format</h2>
        <li><a href="#hierarchy">More complicated data structures are impossible
               to save</a>. It's a "flat table" format.</li>
        <li><a href="#3d">There is no way to store 3D data with textures or
-                       appearances such as material definitions.</a> There is also no way
-                       to store solids or parametric objects.</li>
-           <li><a href="#projection">Projections defined in the *.prj file are
-                       incompatible.</a> Compared to definitions in EPSG /
-                       SpatialReference.org, Esri WKT often has different parameters,
-                       names and sometimes missing parameters.</li>
-           <li><a href="#addmore">Add more</a> ...</li>
+           appearances such as material definitions.</a> There is also no way
+           to store solids or parametric objects.</li>
+     <li><a href="#projection">Projections defined in the *.prj file are
+           incompatible.</a> Compared to definitions in EPSG /
+           SpatialReference.org, Esri WKT often has different parameters,
+           names and sometimes missing parameters.</li>
+     <li>Line and polygon geometry type, single or multipart, cannot be reliably
+           determined at the layer level, it must be determined at the
+           individual feature level.
+     <li><a href="#addmore">Add more</a> ...</li>
    </ul>
    </div>
```

Náhrada za Shapefile?

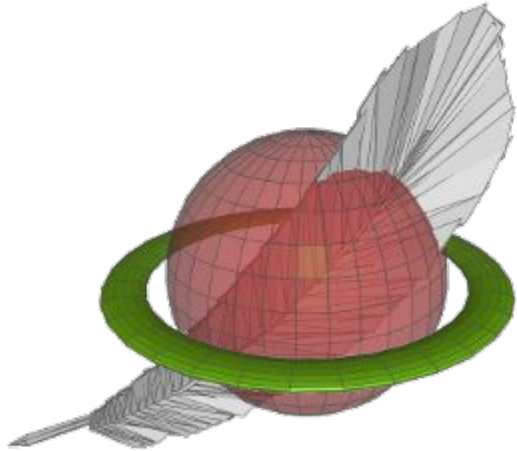
Náhrada za Shapefile

- Zvládne komplikované datové struktury
- Je rozumně velký
- Rozumně rychlý
- Podporovaný, otevřený
-

Náhrada za Shapefile

- OGC GeoPackage
- GeoJSON
- OGC GML
- SpatiaLite
- ~~CSV~~
- ~~OGC KML~~

SpatiaLite



- http://spatialite.org
- Základem je SQLite
  - Souborová SQL databáze (téměř kompletní SQL-92)
- Podporuje OGC Simple Features
- Nepodporuje rastry

OGC GML

- http://opegeospatial.org/standards/gml
- XML
- OGC
- INSPIRE

GeoJSON

- http://geojson.org
- JSON

OGC GeoPackage

- http://geopackage.org
- OGC Standard
- Postavený nad SQLite
- Podporuje i rastry
- Oficiální rozšíření

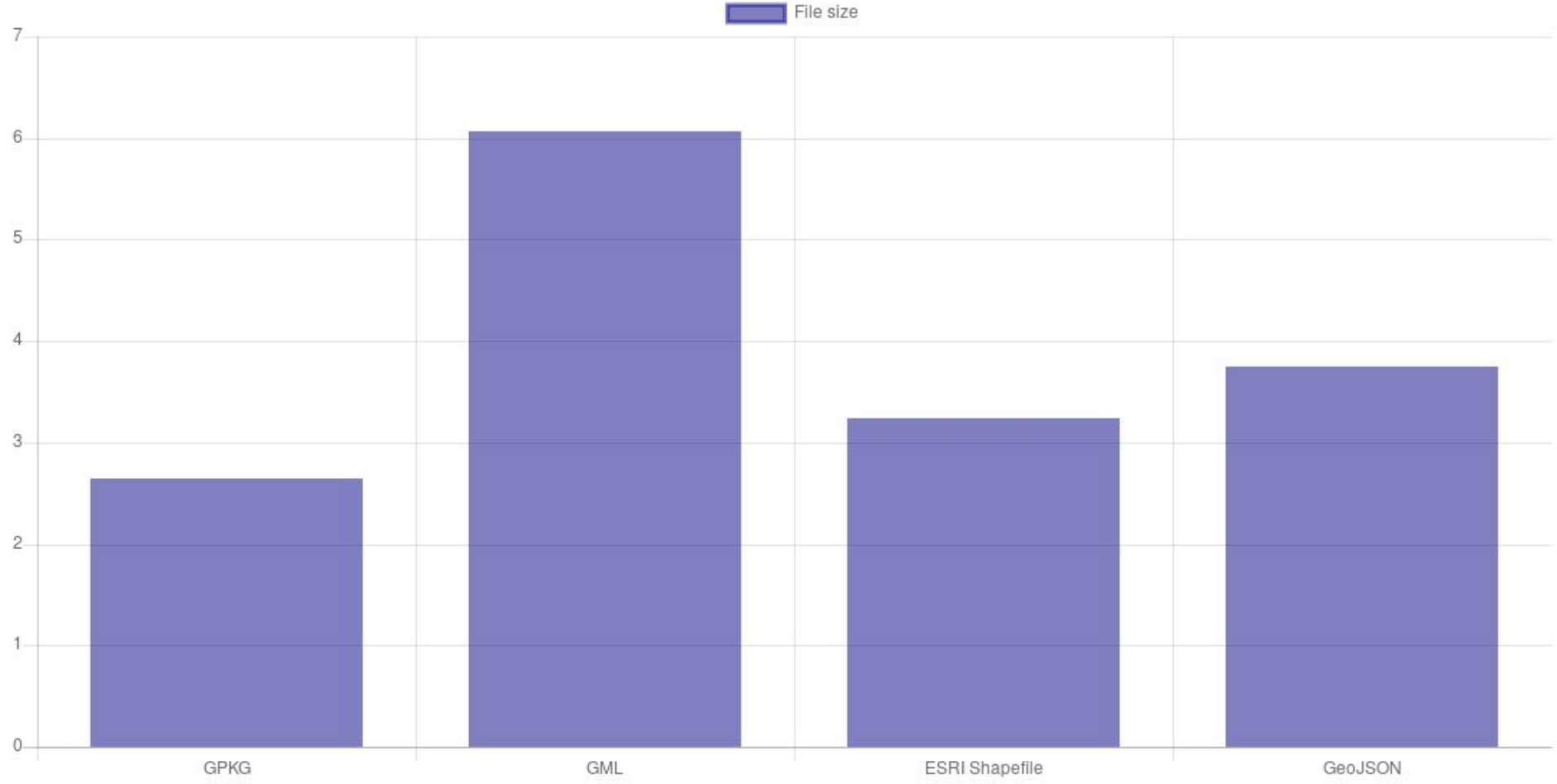
SQLite

- http://sqlite.org
- Souborová databáze
- SQL 92
- Dynamické typování, datový typ je vázaný na buňku, ne na celý sloupec
- Špatná podpora pro paralelní zápis
- fulltext
- Chrome, Opera, Safari, Firefox, Skype, Windows 10, ...

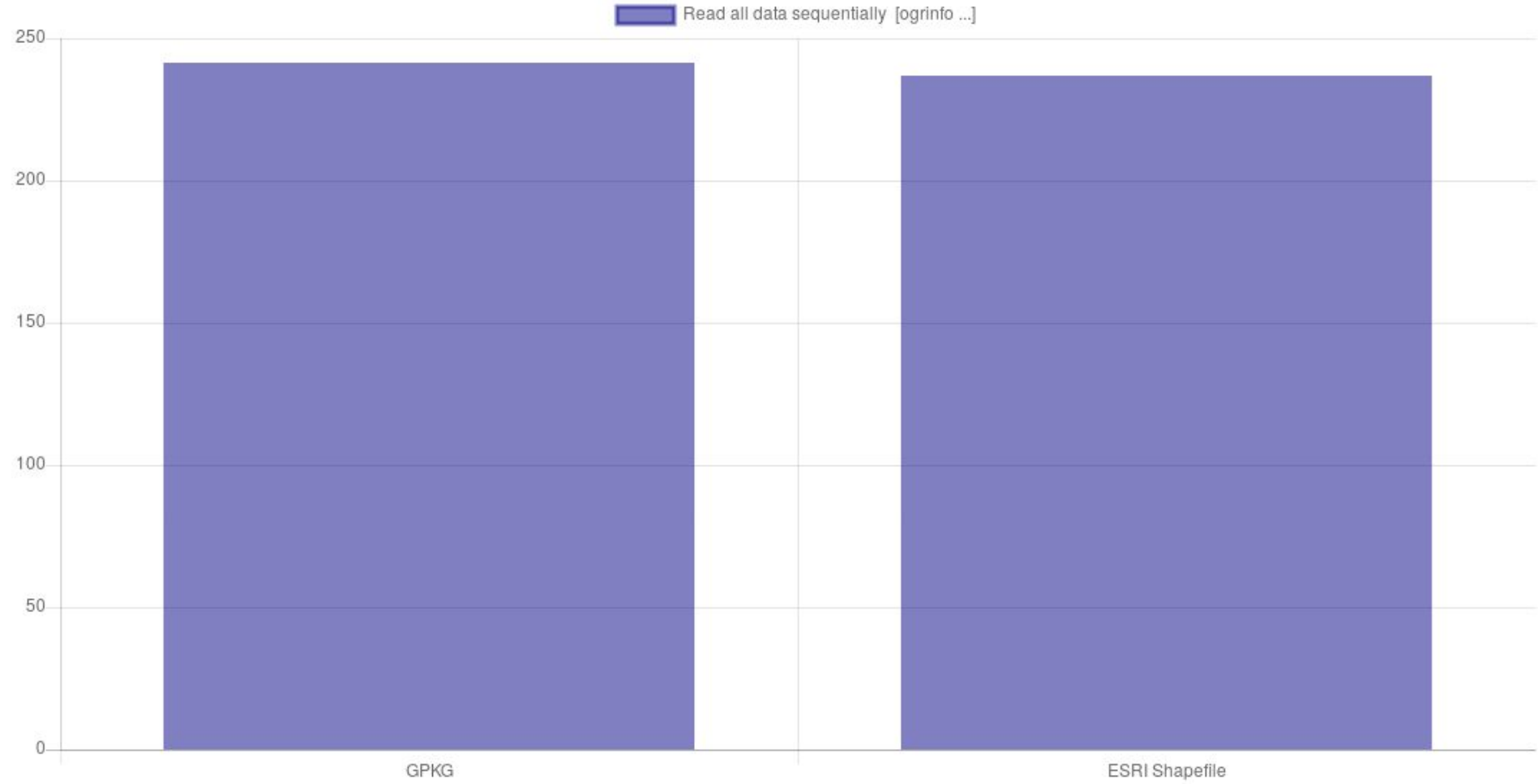
GeoPackage vs. Shapefile

http://switchfromshapefile.org/compare.html

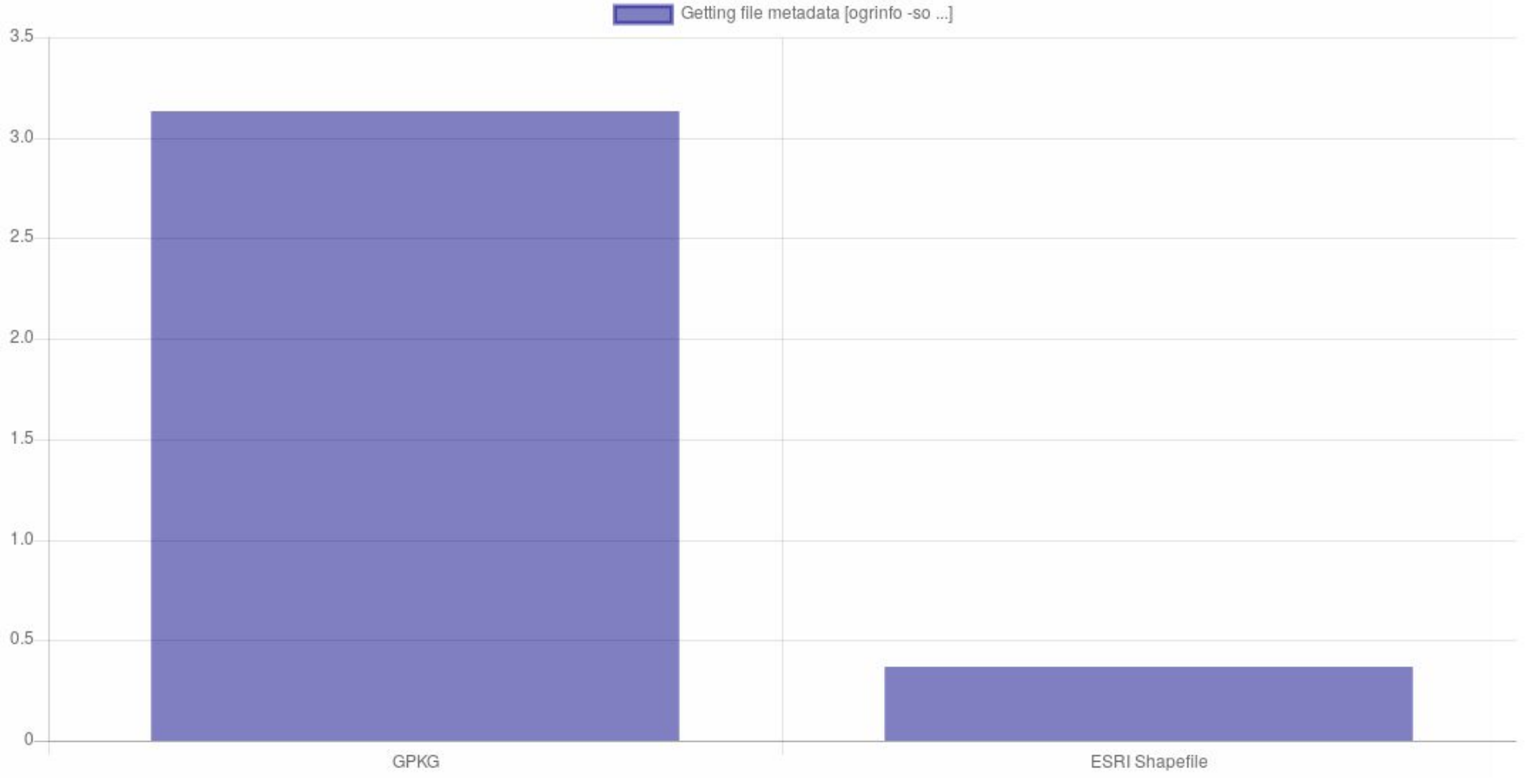
File size [GB]



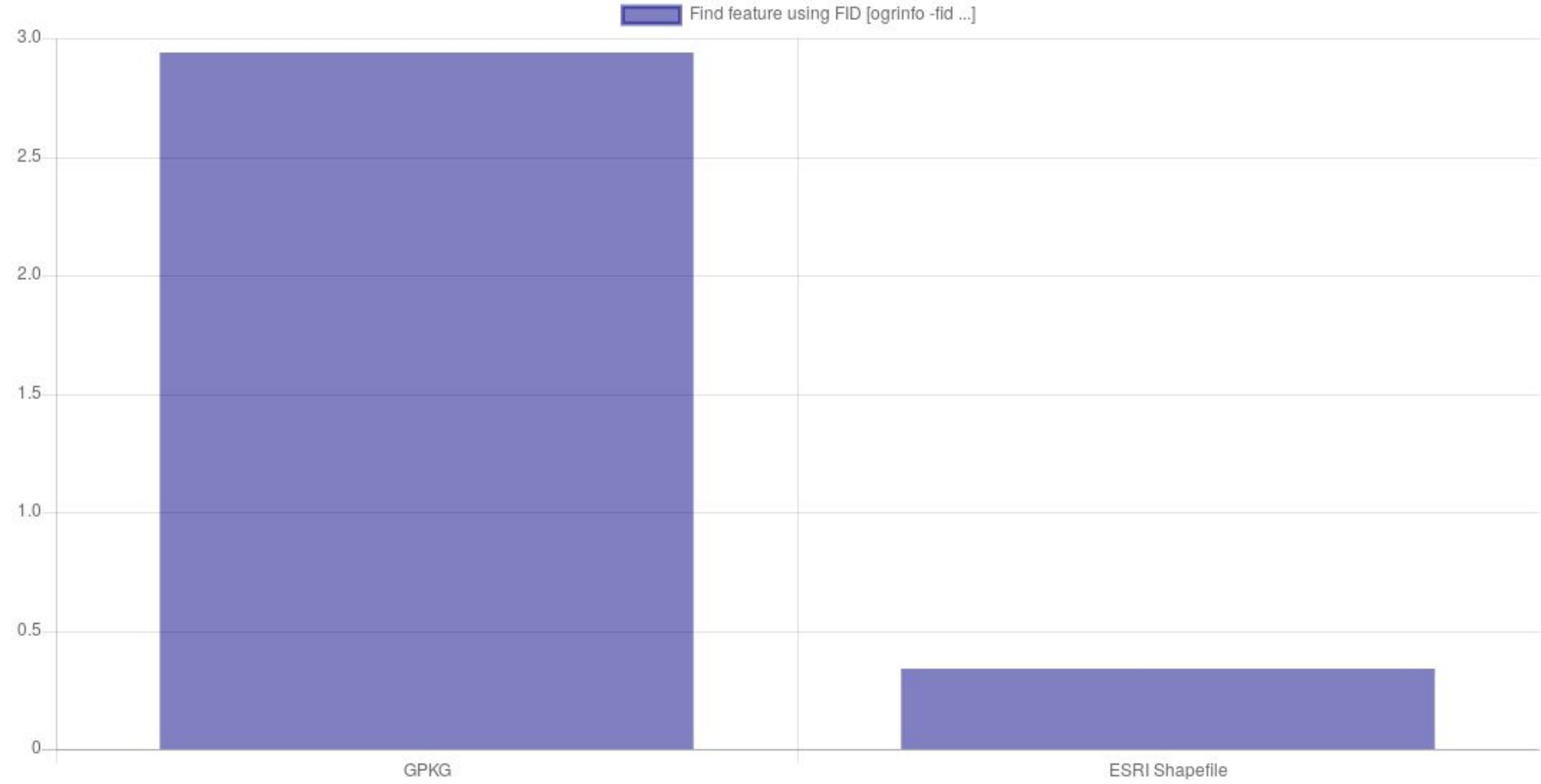
Read all data sequentially [ogrinfo ...]



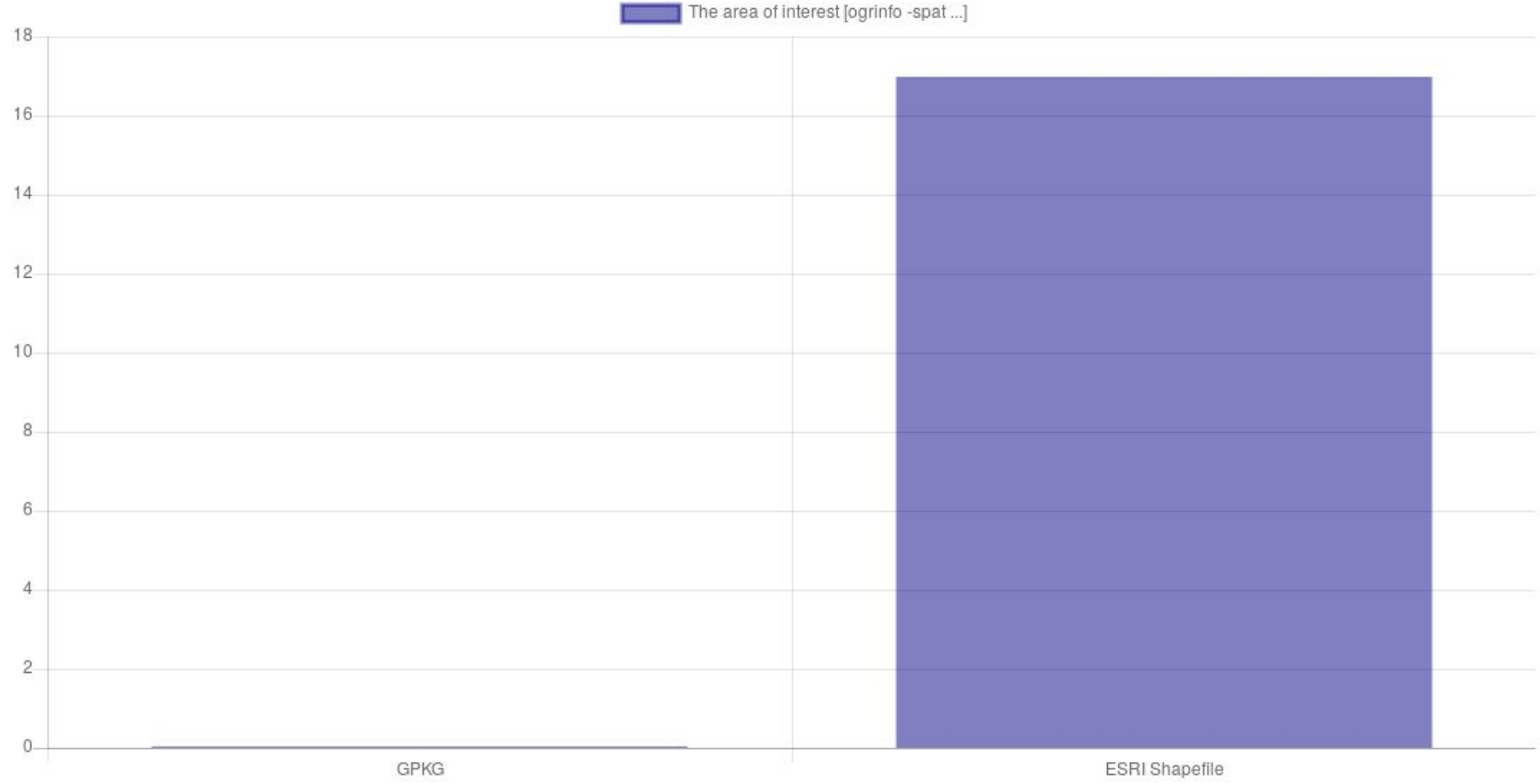
Getting file metadata [ogrinfo -so ...]



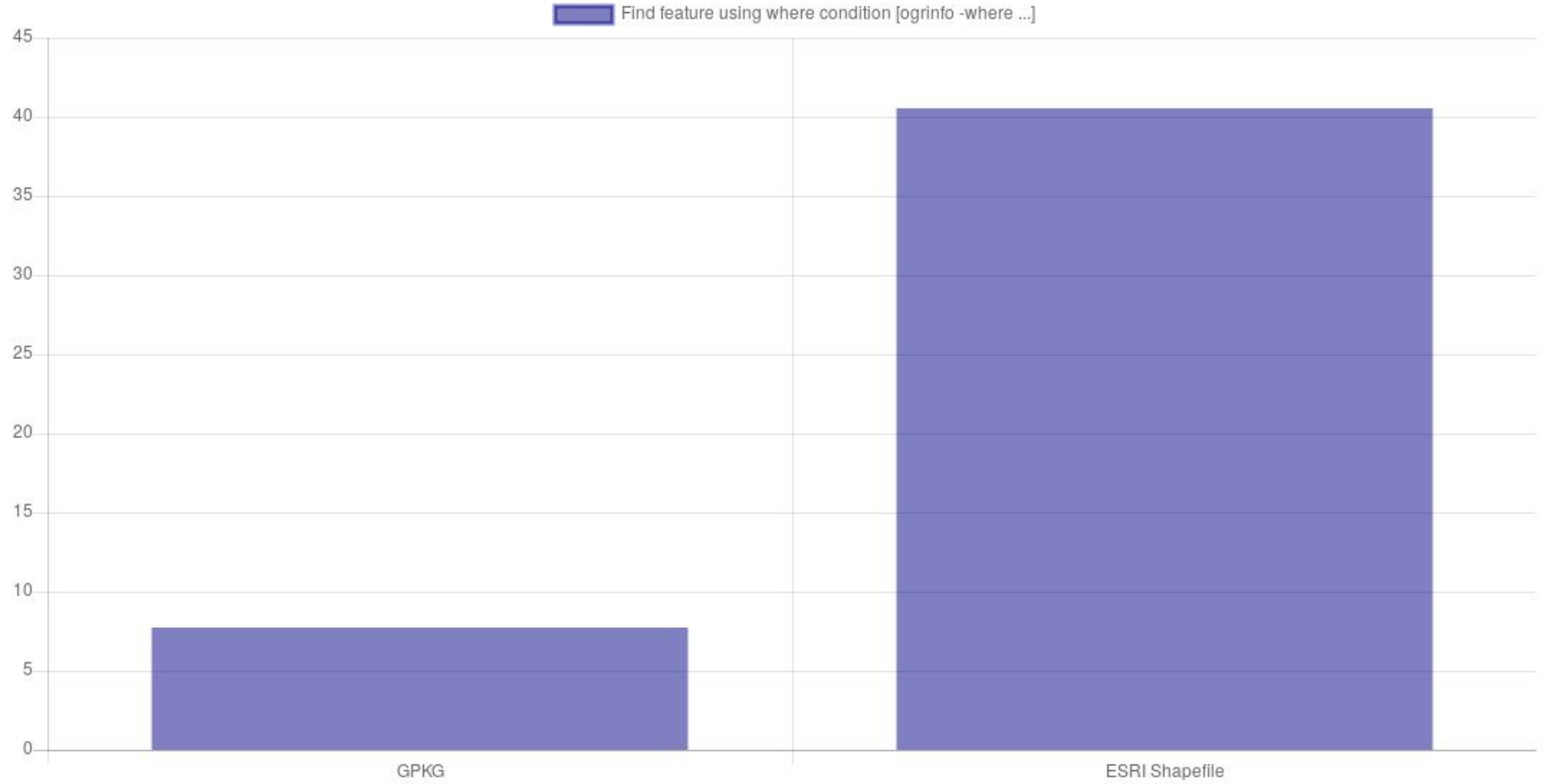
Find feature using FID [ogrinfo -fid ...]



The area of interest [ogrinfo -spat ...]



Find feature using where condition [ogrinfo -where ...]



GeoPackage vs. Shapefile - závěr

- GeoPackage je lehce úspornější
- Relativně pomalejší při jednoduchých dotazech, i když latence není nijak fatální
- Výrazně rychlejší při komplexních prostorových a databázových filtrech.

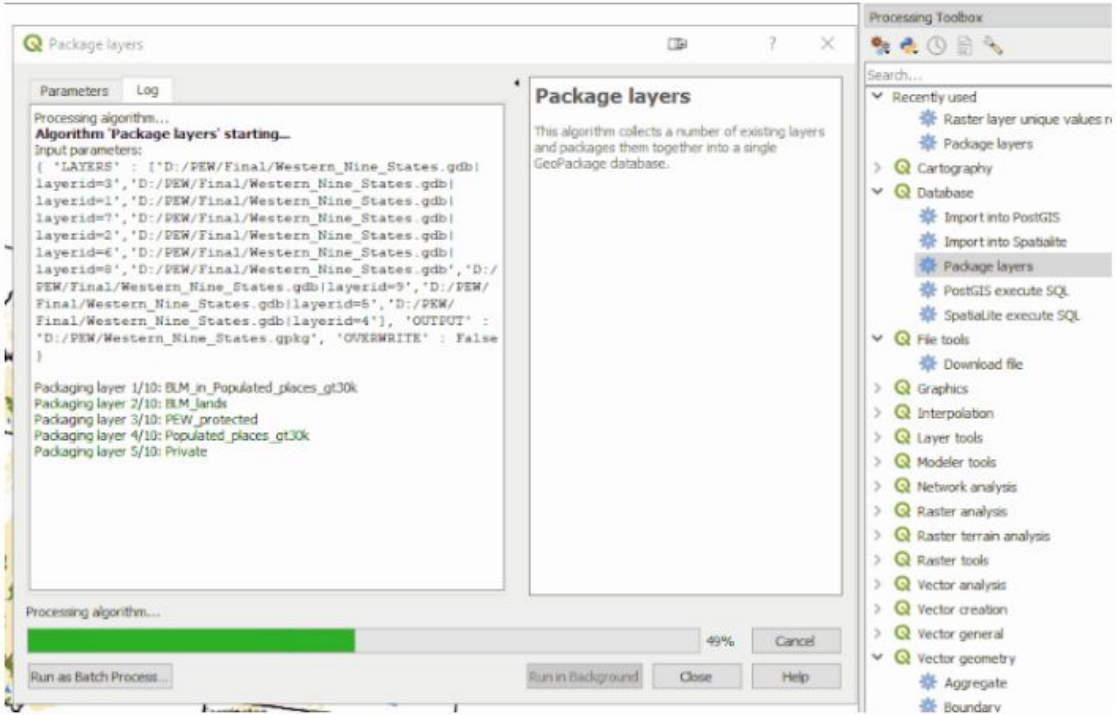
Podpora pro OGC GeoPackage

- QGIS
- GDAL
- ArcGIS
- FME
- …
- http://www.geopackage.org/implementations.html

QGIS 3

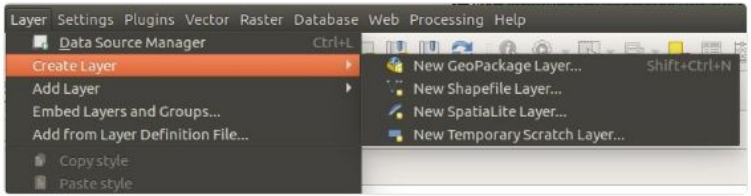
GEOPACKAGE SUPPORT!

QGIS 3 now supports the GeoPackage format as the default. This makes them really easy and convenient to work with in 3. This is probably the beginning of the end for shapefiles. You can easily save to an existing or new GeoPackage. This includes output from Geoprocessing tools. There is even a new Package layers tool that ports out all the layers in your project to a GeoPackage!



Jachym Cepicky
@jachymc

Thhis #pullrequest sets #geopackage as the preferred "new vector layer" format for @qgis github.com/qgis/QGIS/pull… will it make to master?



1:24 PM - 16 Oct 2017

8 Likes

1 8

Jachym Cepicky
@jachymc

The @GRASSGIS is moving from the @shapefile to #geopackage too trac.osgeo.org/grass/ticket/3…

Did you switchfromshapefile.org already?

5:53 AM - 26 Oct 2017

9 Retweets 12 Likes

9 12

Add another Tweet

Jachym Cepicky
@jachymc

Last piece missing?

trac.osgeo.org/gdal/ticket/71…

#gdal @shapefile -> #geopackage as default format

switchfromshapefile.org

2:37 PM - 28 Oct 2017

3 Retweets 8 Likes

1 3 8

Map Styling Information

https://github.com/pka/qgpkg/blob/master/qgis_geopackage_extension.md

Oficiální extenze

GeoPackage - shrnutí

- http://geopackage.org
- OGC Standard
- Nad SQLite
- Souborový
- Podpora pro rastry
- Extenze
- Rozšířený, aniž byste si toho možná všimli.

SHP

NO
I PREFER
GeoPackage!
SHP

Jáchym Čepický

jachym.cepicky@opengeolabs.cz

http://opengeolabs.cz | http://gismentors.cz

http://switchfromshapefile.org