

# ALPO-2014 työpaja: OGC GeoPackage

OSGeo:n paikkatietotyöpaja 27. toukokuuta 2014

Jukka Rahkonen, <http://latuviitta.org>

Lisenssi Creative Commons Attribution 3.0 Unported

Viimeksi muokattu 10. kesäkuuta 2014

## Johdanto

GeoPackage on uusi, helmikuussa 2014 hyväksytty OGC:n standardi paikkatietojen välittämiseen. Standardin tarkoituksena on paikata eräitä nykyisin yleisimmin käytettyjen siirtotiedostomuotojen ongelmia.

### ESRI Shapefile: hyvää ja huonoa

- + yleisesti hyvin tuettu, joten sopii tietojen siirtoon järjestelmästä toiseen
- + nopeasti luettava tiedostomuoto
- + sopii paitsi tietojen siirtoon myös tuotantokäyttöön sellaisenaan
- + spatiaali-indeksin luominen on mahdollista
- shapefile on tiedostonippu, johon kuuluu monta erillistä tiedostoa, jotka on monesti parasta pakata zip-arkistoksi, jonka käsittelystä tulee omat hankaluutensa
- ei tue indeksointia ominaisuustietojen perusteella
- ominaisuustiedon nimessä saa olla korkeintaan 10 merkkiä
- tukee vain päivämäärä-tietotyyppiä, mutta ei päivämäärää+aikaa
- pisteet, viivat ja polygonit on jaettava omiin shapefilenippuihinsa
- merkistökoodauksen ja projektitiedon liittämisenä usein hankaluuksia
- tiedostokoon maksimiraja .shp -osan ollessa rajoittavana tekijänä; sovelluksesta riippuen joko 2 Gt, 4 Gt tai ehdoton maksimi 8 Gt.

### MapInfo-tiedosto: hyvää ja huonoa

- + Tukee piirtotyöliien tallentamista yhdessä kohteiden kanssa
- + Samaan tiedostoon voi tallentaa pisteitä, viivoja ja polygoneja
- + Sopii paitsi tietojen siirtoon myös tuotantokäyttöön sellaisenaan
- Useita tiedostoja, joita käsitellään yhdessä nippuna, tarve zipata jakelua varten
- Usein ongelmia koordinaattijärjestelmien ja merkistökoodauksen kanssa
- Vain MapInfo tukee kunnolla, ei sovin hyvin tietojen siirtoon eri järjestelmien välillä

### GML: hyvää ja huonoa

- + Monipuolinen, GML:n avulla voidaan siirtää melkein mitä ikinä osaa kuvitella
- + Ei rajaa tiedostokoolle
- Monimutkainen, uusimmassa standardissa yli 600 sivua
- Indeksien käyttö mahdotonta
- Pelkkä siirtoformaatti, joka muunnetaan aina johonkin muuhun muotoon ennen tuotantokäyttöä

## GeoPackage: hyvää ja huonoa

- + Yksi tiedosto
- + Tiedosto pitää sisällään tietokannan, johon voi kuulua rajaton määrä erillisiä tietokantatauluja
- + Merkistönä UTF-8
- + Koordinaattijärjestelmät hyvin ja yksiselitteisesti määritelty
- + Tukee sekä spatiaali-indeksiä että tavallisia ominaisuustietoindeksejä
- + Nopea
- + Käyttää hyvin yleisesti käytettyä SQLite-tietokantaa
- + GeoPackage voi pitää sisällään vektoreiden lisäksi myös rasteriaineistot
- + Eri ohjelmistot alkavat tukea Geopaketteja ilmeisesti varsin nopeasti
- + Se on normaali tietokanta, johon voidaan tallentaa myös mitä tahansa taulukkomuotoista tietoa, mutta myös esimerkiksi kuvia, videokuvaa ja ääntä binäärimuodossa, mikä avaa hienoja uusia mahdollisuuksia sovellusten kehittäjille
- + Se on standardi, ei ole epäselvyyttä, mikä tiedostossa käytetty rakenne on oikea ja mikä on väärä
- + Tiedostokoko voi olla hyvin suuri, vaikkakaan kymmenien gigatavujen suuruisten tietokantojen tekeminen ei välttämättä ole järkevää niiden ylläpitotoimenpiteiden hankaloitumisen takia.
- Eri tyyppin geometriat on tallennettava omiin erillisiin tauluihinsa
- Kuten SQLitessä muutenkin, tietokanta on tavallaan jaettu kahteen osaan. Yhtenä osana on tiedokannan datatiedosto, jollainen GeoPackage:kin on, mutta toiseksi osaksi tarvitaan tiedokantaa käsittelevä tietokantamoottori. Tietokantamoottori on jokaisen käyttäjän omalla koneella, joskus jopa käyttävään ohjelmistoon sisäänrakennettuna, mistä voi aiheutua outoja ongelmia versioiden yhteensopimattomuuden takia (GeoPackage luotu versiolla, jossa on sellaisia ominaisuuksia, joita lukemiseen käytettävä moottori ei tue).

## **Hyödyllisiä linkkejä:**

### **GeoPackage-kotisivu**

<http://www.geopackage.org/>

### **Tietoa standardin kehittämisen lähtökohdista**

<http://www.opengeospatial.org/projects/groups/geopackageswg>

### **Vastine, joka johti muutokseen GeoPackage-standardin 1. ja 2. luonnoksen välillä**

<http://lists.opengeospatial.org/pipermail/requests/2013-February/000205.html>

### ***Kirjastoja ja ohjelmistoja (ajantasainen lista GeoPackage:n kotisivulla)***

#### **Luciad libgpkg**

<https://bitbucket.org/luciad/libgpkg>

#### **GDAL:in GeoPackage-ajuri**

[http://www.gdal.org/drv\\_geopackage.html](http://www.gdal.org/drv_geopackage.html)

#### **GDAL:in ogr2ogr-muunnosohjelman käyttöohje**

<http://gdal.org/ogr2ogr.html>

#### **OpenJUMP-blogikirjoituksia geopaketista**

<http://openjump.blogspot.fi/2014/02/openjump-plus-reads-ogc-geopackages.html>

<http://openjump.blogspot.fi/2014/02/openjump-and-geopackage-part-2-spatial.html>

<http://openjump.blogspot.fi/2014/03/openjump-and-geopackage-part-3-make-it.html>

Blogikirjoitukset saattavat olla osin vanhentuneita GDAL:in osalta koska ajuriin on tällä välin lisätty mm. tuki spatiaali-indeksien luomiseksi.

#### **GeoTools ja geopaketit**

<http://geotoolsnews.blogspot.com/2014/02/geotools-11-rc1-released.html>

#### **Spatialite ja geopaketit**

<https://groups.google.com/forum/#!topic/spatialite-users/D0MV4Z3k0-E>

#### **ESRI ja geopaketit**

<http://www.esri.com/esri-news/releases/14-2qtr/esri-supports-ogc-geopackage-encoding-standard>

## ***Työpajassa käytetyt ohjelmistot***

### **GDAL**

Ladataan palvelusta <http://gisinternals.com/sdk/>

Käytetään tuoretta eli viimeistä kehitysversiota ”MSVC2010 (Win64) -development” ja laajinta tarjolla olevaa zip-pakettia ”Compiled binaries in a single .zip package”, jonka asentamiseen ei tarvita järjestelmänvalvojan oikeuksia.

<http://gisinternals.com/sdk/Download.aspx?file=release-1600-x64-gdal-mapserver.zip>

### **OpenJUMP**

Toukokuussa 2014 julkaistu virallisen 1.7.0 Plus -versio zip-pakettina, jonka asentamiseen ei tarvita järjestelmänvalvojan oikeuksia.

<http://sourceforge.net/projects/jump-pilot/files/OpenJUMP/1.7.0/OpenJUMP-Portable-1.7.0-r3977-PLUS.zip/download>

OpenJUMP:sta on tulossa muutaman päivän sisällä työpajan jälkeen versio 1.7.1, ja ainakin Mac-käyttäjien kannattaa käyttää sitä heti kun se on saatavilla.

### **Spatialite-gui**

Versio 1.7.1 (64-bittinen)

[http://www.gaia-gis.it/gaia-sins/windows-bin-amd64/spatialite\\_gui-1.7.1-win-amd64.zip](http://www.gaia-gis.it/gaia-sins/windows-bin-amd64/spatialite_gui-1.7.1-win-amd64.zip)

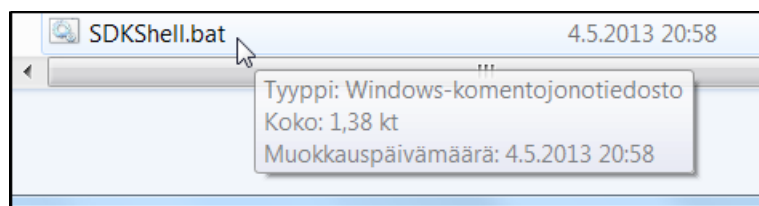
Spatialitestä on tulossa lähiaikoina työpajan jälkeen versio 4.2, joka tukee geopaketteja, joten tähän versioon kannattaa siirtyä heti kun se on mahdollista.

<https://www.gaia-gis.it/fossil/libspatialite/index>

## ***Ohjelmien asentaminen ja käynnistäminen***

Kaikki edellä mainitut ohjelmistot asennetaan samalla eli purkamalla ladattu zip-tiedosto johonkin hakemistoon kovalevyllä.

### **GDAL:n käynnistäminen**



Kaksoisnäpätys komentojonotiedoston nimeen riittää.

```
C:\Windows\system32\cmd.exe
C:\ohjelmat\GDAL_dev\
C:\ohjelmat\GDAL_dev\
Setting environment for using the GDAL and MapServer tools.
WARNING: If you encounter problems with missing oci libraries then type:
SDKShell hideoci
C:\ohjelmat\GDAL_dev>_
```

GDAL on nyt asennettu ja käytettävissä.

### GDAL:in testaus

Testataan, että GDAL-apuohjelmat toimivat ja että geopakettiajuri ”GPKG” löytyy

```
C:\Windows\system32\cmd.exe
C:\ohjelmat\GDAL_dev>ogrinfo
Usage: ogrinfo [--help-general] [-ro] [-ql] [-where restricted_where]
[-spat xmin ymin xmax ymax] [-geomfield field] [-fid fid]
[-sql statement] [-dialect sql_dialect] [-all] [-so] [-fields=
[-geom=<YES/NO/SUMMARY>] [--formats]
datasource_name [layer [layer ...]]
FAILURE: No datasource specified.
```

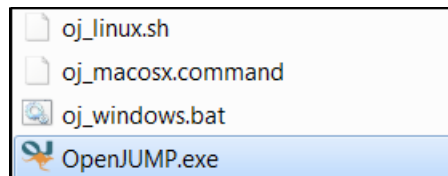
```
C:\Windows\system32\cmd.exe
C:\ohjelmat\GDAL_dev>ogrinfo --formats
Supported Formats:
-> "ESRI Shapefile" (read/write)
-> "MapInfo File" (read/write)
-> "UK .MIF" (readonly)
-> "SDTS" (readonly)
-> "TIGER" (read/write)
-> "S57" (read/write)
-> "DGN" (read/write)
-> "URT" (readonly)
-> "REC" (readonly)
-> "Memory" (read/write)
-> "BNA" (read/write)
-> "CSU" (read/write)
-> "NAS" (readonly)
-> "GML" (read/write)
-> "GPX" (read/write)
-> "LIBKML" (read/write)
-> "KML" (read/write)
-> "GeoJSON" (read/write)
-> "Interlis 1" (read/write)
-> "Interlis 2" (read/write)
-> "GMT" (read/write)
-> "GPKG" (read/write)
```

Jos tietokoneella on käytössä Oracle-ohjelmia, esimerkiksi SQL\*Plus, niin GDAL saattaa tulostaa turhia ja ärsyttäviä virheilmoituksia. Tässä tapauksessa GDAL kannattaa asentaa avaamalla ensin Windows-komentoikkuna, siirtymällä GDAL:in asennushakemistoon ja antamalla käynnistyskomento muodossa

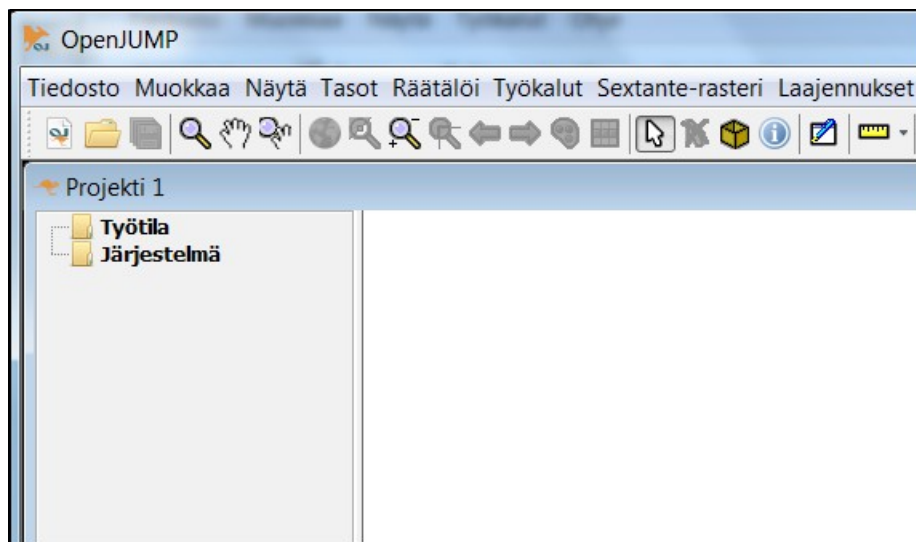
```
sdkshell hideoci
```

## OpenJUMP:n käynnistäminen

OpenJUMP on java-ohjelma ja sen käyttämiseksi tietokoneelta on löydettävä java. Jos java on koneella, niin käyttöjärjestelmästä riippuen ohjelma käynnistetään kaksoisklikkaamalla OpenJUMP:in asennushakemiston alla olevasta bin-hakemistosta löytyvää käynnistyskomentoa: oj\_linus.sh (Linux), oj\_macosx.command (Max OS X) tai oj\_windows.bat. Windows:lla käynnistäminen onnistuu myös kaksoisnäpäyttämällä kengurukuvalla varustettua OpenJUMP.exe-tiedostoa. Exe-ohjelma itse asiassa vain suorittaa oj\_windows.bat -komentojonoa, joten nämä käynnistystavat ovat samanveroiset.



Onnistuneen käynnistyksen jälkeen OpenJUMP:in aloitusnäkyä aukeaa näytölle.



## OpenJUMP:in käynnistysongelmien ratkaiseminen

Jos OpenJUMP ei käynnisty, niin tavallisesti syynä on se, ettei käynnistysjono yrityksistään huolimatta onnistu löytämään koneelta javan asennushakemistoa. Luonnollisesti jos java puuttuu koneelta kokonaan, niin se on ensin asennettava, mutta jos java on jo asennettu, mutta se ei vain löydy, niin OpenJUMP:in käynnistyskomentojonoa on muokattava. Windows:in komentojono on \bin-hakemistosta löytyvä oj\_windows.bat. Avataan se tekstitiedostona ja muokataan seuraavaa riviä

```
rem -- uncomment to manually set java home -
```

Muutetaan riviä niin, että se osoittaa oikeaa omalta koneelta löytyvää javan asennushakemistoa, esimerkiksi

```
set JAVA_HOME=C:\ohjelmat\Java\jre6
```

## **GeoPackage-tiedoston tekeminen GDAL:lla**

GeoPackage-tiedostoja voidaan tehdä GDAL-apuohjelmalla ogr2ogr. Perusmuodossa komennolle tarvitsee määrittää vain tulostiedoston tiedostomuoto (GPKG), tulostiedoston nimi ja lähtötiedoston nimi.

```
ogr2ogr -f GPKG aineisto.gpkg aineisto.shp -progress
```

Ogr2ogr on hyvin monipuolinen ohjelmisto, ja sen ominaisuuksiin kuuluu mm.

- projektiomuunnokset järjestelmästä toiseen
- tietokantojen tuki lähtö- ja tulosaineistoille (Oracle, PostgreSQL/PostGIS, MS SQL, MySQL, SQLite/Spatialite)
- lähtöaineiston haku verkkopalveluista (WFS, REST, tiedosto verkkolevyllä)
- muunnettavan aineiston rajausta aluerajauksella
- muunnettavan aineiston rajausta SQL-kyselyllä

Kannattaa lukea käyttöohjeita.

## **GeoPackage-tietokannan tarkasteleminen ogrinfo-työkalulla**

GDAL-apuohjelma ”ogrinfo” <http://www.gdal.org/ogrinfo.html> on perustyökalu kaikkien vektoritiedostojen ominaisuuksien tutkimiseen.

Peruskomento on

```
ogrinfo aineistolähde
```

```
F:\GPKG>ogrinfo plohkot.shp
INFO: Open of 'plohkot.shp'
       using driver 'ESRI Shapefile' successful.
1: plohkot (Polygon)
```

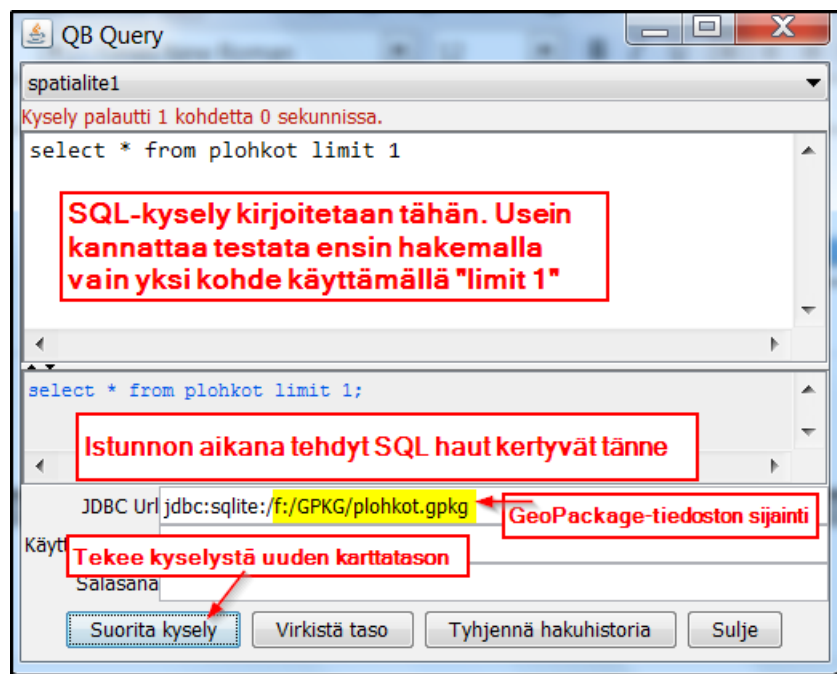
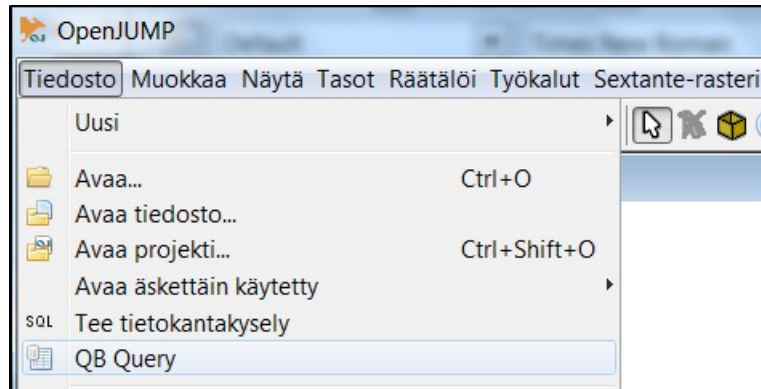
Komento listaa aineistolähteeseen sisältyvät karttatasot. Tietoa aineistotasosta saadaan seuraavalla komennolla, johon on lisätty hyödyllinen parametri -so (summary only)

```
ogrinfo aineistolähde aineistotaso -so
```





## GeoPackage-tietokannan lukeminen OpenJUMP:lla



## Testataan nopeutta

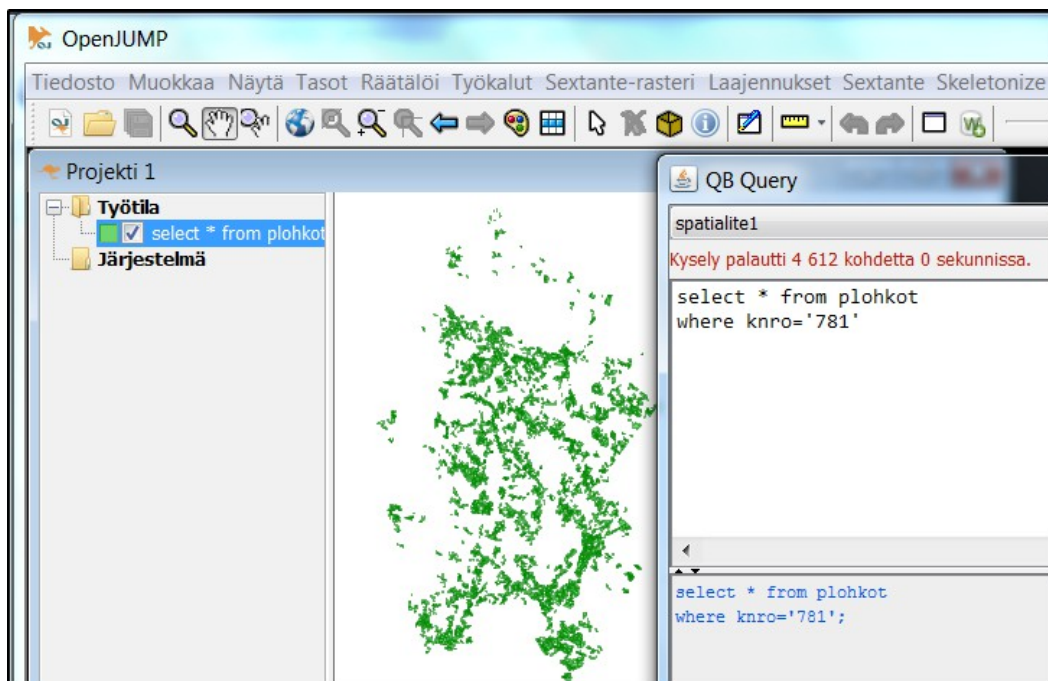
Suurentamalla limit-arvoa saa hyvän käsityksen siitä, kuinka nopeaa tiedon haku GeoPackage-tietokannasta on silloin, kun kyselyyn ei sisälly hakuehtoja. Testiaineistona taulu, jossa on polygoneja ja 7 attribuuttia.

Limit	32-bit java	64-bit java
1	Alle 1 s	Alle 1 s
10	Alle 1 s	Alle 1 s
100	Alle 1 s	Alle 1 s
1000	Alle 1 s	Alle 1 s
10000	Alle 1 s	Alle 1 s
100000	3 s	2 s
300000	8 s	4 s
600000	Muisti (1 Gt) ei riitä	13 s
1122732	Muisti (1 Gt) ei riitä	39 s

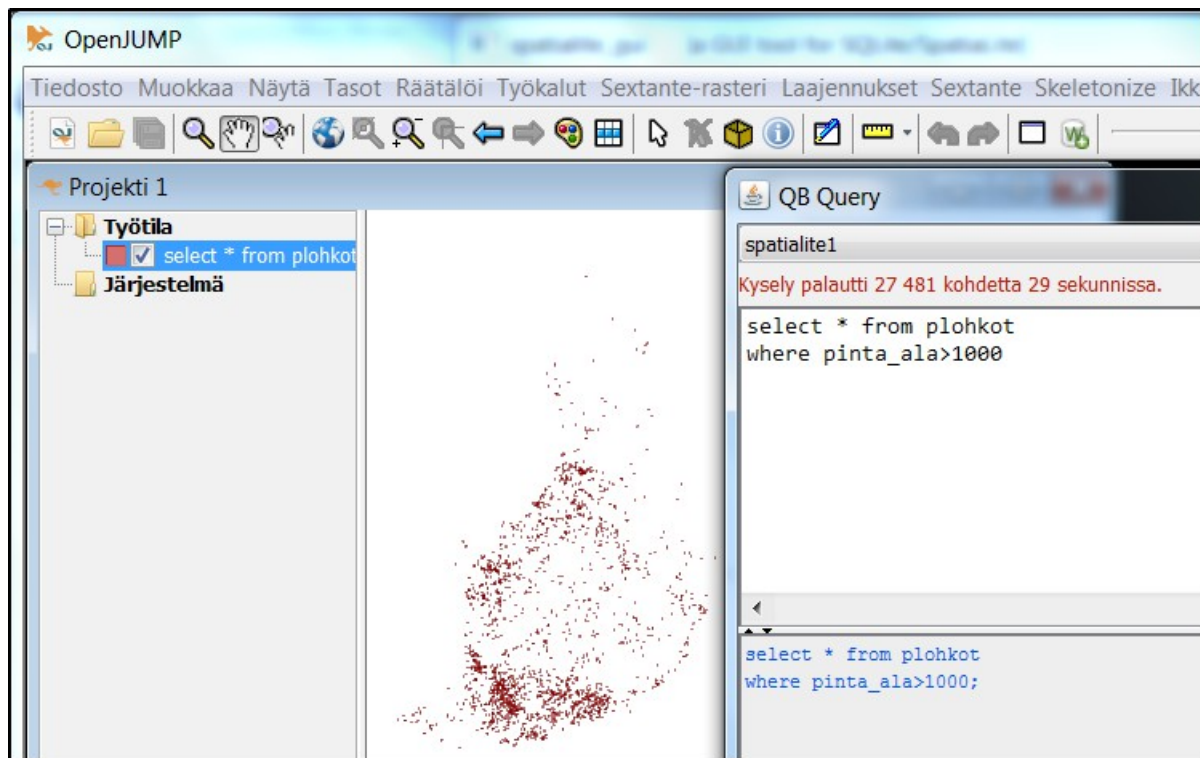
Tuloksista voi painaa mieleensä nyrkkisäännön:

**GeoPackage-tietokannasta voi olettaa saavansa ulos 100000 kohdetta noin 5 sekunnissa. Jos ei saa, niin jokin osa järjestelmää ei toimi niin kuin pitäisi.**

GeoPackage on tietokanta, joten karttatasoja ei ole mikään pakko lukea kokonaisina, vaan niistä voidaan SQL-kyselyllä valita juuri halutut kohteet. Esimerkiksi koko maan peltoaineistosta vain Saarijärven peltolohkot.



Tai kaikki yli 1000 aarin eli yli 10 hehtaarin suuruiset peltolohkot.



Ajat on jätetty nykyviin tarkoituksella. Ilman indeksiä pinta-alan mukaan tehty kysely kesti puoli minuuttia. Se ei ole erityisen hidas, mutta indeksin avulla hakua voidaan nopeuttaa. Indeksit voidaan tehdä esimerkiksi OpenJUMP:in SQL-kyselyikkunasta suorittamalla komento

```
create index plokkot_knro_idx on plokkot (knro);
```

Indeksi voidaan tehdä myös ogrinfo-ohjelmalla

```
ogrinfo plokkot.gpkg -sql "create index plokkot_pinta_ala_idx on plokkot (pinta_ala)"
```

Lopuksi vielä spatiaali-indeksiä käyttävä SQL-kysely. Kysely on konstikas eikä tavallisen käyttäjän voi olettaa kirjoittavan sitä itse, vaan GeoPackage:a tukevan ohjelmiston tehtävänä on hallita kyselyn teko, mutta silti on hyvä tietää, miten kyselyn voi tarpeen vaatiessa rakentaa.

```
select * from plokkot
where rowid in
(select id from rtree_plokkot_geom
where
minx<500000 and
maxx>400000 and
miny<6900000 and
maxy>6800000);
```

