

# Digiroad-aineiston 2018\_01 haltuunotto

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

Lisenssi <https://creativecommons.org/licenses/by/4.0/>

Viimeksi muokattu 24. toukokuuta 2018

Tämä on päivitetty versio vuonna 2016 kirjoitetusta ohjeesta

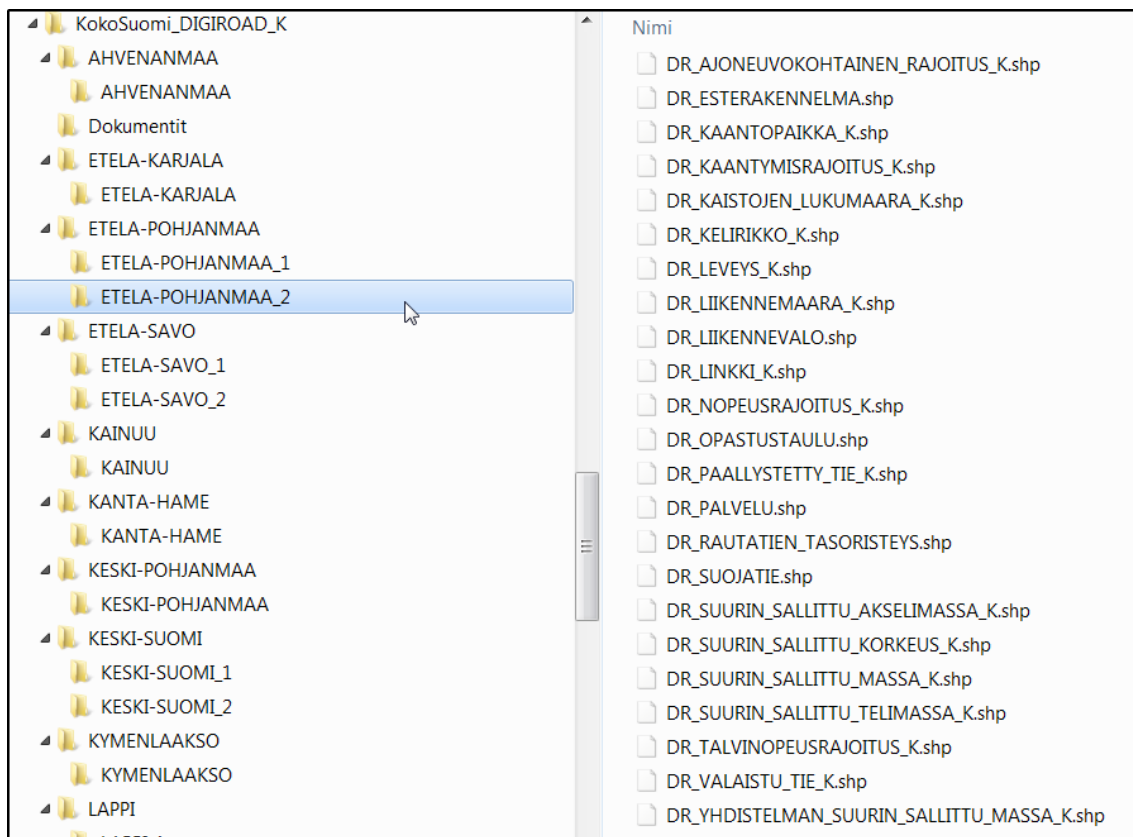
[http://latuviitta.org/documents/Digiroad-aineiston\\_haltuunotto.pdf](http://latuviitta.org/documents/Digiroad-aineiston_haltuunotto.pdf)

Tärkeimmät erot vanhaan ohjeeseen verrattuna ovat ohjeen sovittaminen Didiroad-jakeluun 2018\_01, zip-tiedostojen käyttö suoraan ilman että ne purettaisiin ensin levyille, XYZM-koordinaatien yksinkertaistaminen XY-muotoon sekä jakelussa 2018\_01 esiintyneen merkistöongelman ratkaiseminen. Indeksien luontilauseet on jätetty pois, mutta vuoden 2016 ohjeen pitäisi toimia niiden osalta.

## Yleistä Digiroad-aineistosta

Digiroadin version 2018\_01 jakelumuoto on zip-arkisto. Tiedot on jaettu maakunnittain 19 maakuntahakemistoon, joista 10 jakaantuu kahteen alihakemistoon ja 1 maakunta kolmeen alihakemistoon. Hakemistoja on siten yhteensä 32 kappaletta. Näihin hakemistoihin on tallennettu Digiroadin tiedot ESRI:n shapefile-tiedostomuodossa.

Maakunnittaisissa hakemistossa on 15-29 shapefileä, joista jokainen koostuu 6 osasta. Yhdessä hakemistossa on siten 90-174 tiedostoa. Yhteensä Digiroadin versiossa 2018\_1 on 818 shapefileä, jotka koostuvat 4908:sta tiedostosta.



Aineistossa esiintyvät vektoritasot ja niiden geometriatyypit:

```
1: DR_AJONEUVOKOHTAINEN_RAJOITUS_K (3D Measured Line String)
2: DR_ESTERAKENNELMA (3D Point)
3: DR_KAANTYMISRAJOITUS_K (3D Measured Line String)
4: DR_KAISTOJEN_LUKUMAARA_K (3D Measured Line String)
5: DR_LEVEYS_K (3D Measured Line String)
6: DR_LIIKENNEVALO (3D Point)
7: DR_LINKKI_K (3D Measured Line String)
8: DR_NOPEUSRAJOITUS_K (3D Measured Line String)
9: DR_PAALLYSTETTY_TIE_K (3D Measured Line String)
10: DR_PALVELU (3D Point)
11: DR_SUOJATIE (3D Point)
12: DR_SUURIN_SALLITTU_AKSELIMASSA_K (3D Measured Line String)
13: DR_SUURIN_SALLITTU_KORKEUS_K (3D Measured Line String)
14: DR_SUURIN_SALLITTU_MASSA_K (3D Measured Line String)
15: DR_YHDISTELMAN_SUURIN_SALLITTU_MASSA_K (3D Measured Line String)
16: DR_EUROOPPATIENRO_K (3D Measured Line String)
17: DR_JOUKKOLIIKENNEKAISTA_K (3D Measured Line String)
18: DR_KAANTOPAikka_K (3D Point)
19: DR_KELIRIKKO_K (3D Measured Line String)
20: DR_LIIKENNEMAARA_K (3D Measured Line String)
21: DR_LIITTYMANRO_K (3D Measured Line String)
22: DR_OPASTUSTAULU (3D Point)
23: DR_RAUTATIEN_TASORISTEYS (3D Point)
24: DR_SUURIN_SALLITTU_LEVEYS_K (3D Measured Line String)
25: DR_SUURIN_SALLITTU_PITUUS_K (3D Measured Line String)
26: DR_SUURIN_SALLITTU_TELIMASSA_K (3D Measured Line String)
27: DR_TALVINOPEUSRAJOITUS_K (3D Measured Line String)
28: DR_VALAISTU_TIE_K (3D Measured Line String)
29: DR_VAK_RAJOITUS_K (3D Measured Line String)
```

## Ongelmat

Jos Digiroadia on tarkoitus käyttää koko Suomen kattavana aineistona, niin se, että jakelupaketissa tieaineisto on jaettu 19 maakuntaan, 32 hakemistoon, 818 shapefileen ja 4908 tiedostoon on ilmeinen ongelma. Toinen asia, joka käyttötapauksesta riippuen voi olla ongelma, liittyy XYZM-tyyppisiin geometrioihin. M-koordinaattia käytetään Digiroad-aineistossa lineaariseen referointiin eli kertomaan, kuinka kaukana taitepiste on viivan alkupisteestä. Tästä tiedosta on hyötyä vain ohjelmille, jotka hallitsevat lineaarisen referoinnin, ja sellaiset ohjelmat ovat harvassa. Muille ohjelmille M-koordinaatti on tarpeeton. M-koordinaatista ei välttämättä ole myöskään haittaa, vaan ohjelmat käyttäytyvät niin kuin sitä ei olisi olemassa. Geometrisia funktioita käytettäessä tulos voi kuitenkin arvaamaton. Esimerkiksi polygoniksi puskuroidulle viivalle ei ole mahdollista interpoloida M-arvoja, ja ohjelma saattaa luoda esimerkiksi POLYGON ZM -tyyppisen geometrian, jossa Z-arvona on aina nolla (SpatiaLite) tai kaksiulotteisen POLYGON-geometrian (PostGIS). Ikävämpi tilanne on silloin, jos ohjelma ei avaa M-koordinaatteja sisältävää shapefile-tiedostoa ollenkaan, vaan antaa esimerkiksi virheen "Tunnistamaton shape type: 23 PolyLineM" tai jotain vastaavaa. Nyrkkisääntönä voi sanoa, että jos käyttäjä ei tiedä tarvitsevasa M-koordinaattia, niin sen voi huoletta pudottaa pois.

Se mitä edellä sanottiin M-koordinaatista, pätee yleensä myös Z-koordinaattiin. Suurimmassa osassa Digiroadin aineistoissa Z-koordinaatti on sitä paitsi 0, joten siitä ei senkään takia ole mitään hyötyä. Tämä johtunee shapefile-formaatista, sillä M-koordinaattia ei ole mahdollista tallentaa shapefileen ilman, että samalla tallennetaan myös Z-koordinaatti. DR\_LINKKI-aineistossa kuitenkin on mukana oikeat Z-koordinaatit.

## Ratkaisu: GeoPackage

Tuhansien tiedostojen aiheuttama vaiva voidaan välttää kirjoittamalla koko Digiroad-aineisto GeoPackage-tietokantaan. Tämä onnistuu GDAL-apuohjelmalla ogr2ogr <http://www.gdal.org/ogr2ogr.html> käyttämällä hyväksi GeoPackage-vektoriajuria [http://www.gdal.org/drv\\_geopackage.html](http://www.gdal.org/drv_geopackage.html). Digiroadin ZIP-jakelupaketin yhteydessä on lisäksi kätevä käyttää hyväksi GDAL:in kykyä lukea tiedot suoraan ZIP-arkistosta ilman, että se tarvitsisi ensin purkaa levyille [http://gdal.org/gdal\\_virtual\\_file\\_systems.html](http://gdal.org/gdal_virtual_file_systems.html).

## Näin se käy

Ogr2ogr-komento yhden Digiroad-zipissä olevan hakemiston muuntamiseksi

```
ogr2ogr -f gpkg digiroad_k_ahvenanmaa.gpkg  
/vsizip/KokoSuomi_DIGIROAD_K_EUREF-  
FIN.zip/KokoSuomi_DIGIROAD_K/AHVENANMAA/AHVENANMAA
```

- ogr2ogr ohjelman nimi
- -f gpkg luodaan GeoPackage
- digiroad\_k\_ahvenanmaa.gpkg  
luotavan GeoPackage-tiedoston nimi
- /vsizip/KokoSuomi\_DIGIROAD\_K\_EUREF-  
FIN.zip/KokoSuomi\_DIGIROAD\_K/AHVENANMAA/AHVENANMAA  
lähtöaineiston sisältö, ”kaikki mitä GDAL ymmärtää zip-arkiston  
KokoSuomi\_DIGIROAD\_K\_EUREF-FIN.zip sisältämästä hakemistosta  
KokoSuomi\_DIGIROAD\_K/AHVENANMAA/AHVENANMAA”

Ogr2ogr-komento, jolla voidaan lisätä tiedot yhdestä Digiroad-zipin hakemistosta jo olemassa olevaan GeoPackage-tietokantaan.

```
ogr2ogr -f gpkg -update -append -addfields -dim 2 digiroad_k_xy.gpkg  
/vsizip/KokoSuomi_DIGIROAD_K_EUREF-  
FIN.zip/KokoSuomi_DIGIROAD_K/AHVENANMAA/AHVENANMAA -oo encoding="ISO-  
8859-1"
```

- ogr2ogr ohjelman nimi
- -f gpkg luodaan GeoPackage
- -update päivitetään olemassa olevaa GeoPackage-tietokantaa eikä kirjoiteta yli
- -append lisätään tietoja olemassa olevaan tauluun eikä kirjoiteta yli
- -addfields lisätään tauluun uusia attribuutteja, jos niitä tulee datassa vastaan
- -dim 2 pudotetaan geometrioista pois Z- ja M-koordinaatit
- digiroad\_k\_xy.gpkg  
luotavan GeoPackage-tiedoston nimi
- /vsizip/KokoSuomi\_DIGIROAD\_K\_EUREF-  
FIN.zip/KokoSuomi\_DIGIROAD\_K/AHVENANMAA/AHVENANMAA  
lähtöaineiston sisältö, ”kaikki mitä GDAL ymmärtää zip-arkiston  
KokoSuomi\_DIGIROAD\_K\_EUREF-FIN.zip sisältämästä hakemistosta  
KokoSuomi\_DIGIROAD\_K/AHVENANMAA/AHVENANMAA”
- -oo encoding="ISO-8859-1"  
tämä ”open option” oli tarpeen Digiroad 2018\_01 -julkaisun kanssa, koska muuten  
Kainuun maakunnan teiden nimien merkistön kanssa tuli ongelmia, eivätkä äät ja ööt  
muuntuneet oikein GeoPackagen käyttämään UTF-8-merkistöön.

Jos halutaan säilyttää Z- ja M-koordinaatit, niin ”-dim 2” -parametri jätetään pois. Optio -oo encoding="ISO-8858-1" on normaalisti tarpeeton. Vanhemmat GDAL-versiot eivät tue tätä optiota.

## Koko maan aineiston muunnos komentojonolla

GDAL:in vsizip-ajuri ei kykene lukemaan listaa zip-arkistoon kuuluvista alihakemistoista. Siitä syystä koko maan Digiroad-aineiston muuntaminen GeoPackage-muotoon ei onnistu yksirivisellä komentojonolla. Windows-ympäristössä toimivaksi testattu tapa on kirjoittaa kaksi komentojonoa, joista ensimmäinen sisältää käsityönä kirjoitetun listan hakemistoista joka kutsuu sitten varsinaista muunnoskomentoa hakemisto kerrallaan.

Komentojono 1.

kokoa\_digiroad\_k.bat  
=====

```
TIME /t
call digiroad_k2gpkg.bat AHVENANMAA/AHVENANMAA
call digiroad_k2gpkg.bat ETELA-KARJALA/ETELA-KARJALA
call digiroad_k2gpkg.bat ETELA-POHJANMAA/ETELA-POHJANMAA_1
call digiroad_k2gpkg.bat ETELA-POHJANMAA/ETELA-POHJANMAA_2
call digiroad_k2gpkg.bat ETELA-SAVO/ETELA-SAVO_1
call digiroad_k2gpkg.bat ETELA-SAVO/ETELA-SAVO_2
call digiroad_k2gpkg.bat KAINUU/KAINUU
call digiroad_k2gpkg.bat KANTA-HAME/KANTA-HAME
call digiroad_k2gpkg.bat KESKI-POHJANMAA/KESKI-POHJANMAA
call digiroad_k2gpkg.bat KESKI-SUOMI/KESKI-SUOMI_1
call digiroad_k2gpkg.bat KESKI-SUOMI/KESKI-SUOMI_2
call digiroad_k2gpkg.bat KYMENLAAKSO/KYMENLAAKSO
call digiroad_k2gpkg.bat LAPPI/LAPPI_1
call digiroad_k2gpkg.bat LAPPI/LAPPI_2
call digiroad_k2gpkg.bat LAPPI/LAPPI_3
call digiroad_k2gpkg.bat PAIJAT-HAME/PAIJAT-HAME
call digiroad_k2gpkg.bat PIRKANMAA/PIRKANMAA_1
call digiroad_k2gpkg.bat PIRKANMAA/PIRKANMAA_2
call digiroad_k2gpkg.bat POHJANMAA/POHJANMAA
call digiroad_k2gpkg.bat POHJOIS-KARJALA/POHJOIS-KARJALA_1
call digiroad_k2gpkg.bat POHJOIS-KARJALA/POHJOIS-KARJALA_2
call digiroad_k2gpkg.bat POHJOIS-POHJANMAA/POHJOIS-POHJANMAA_1
call digiroad_k2gpkg.bat POHJOIS-POHJANMAA/POHJOIS-POHJANMAA_2
call digiroad_k2gpkg.bat POHJOIS-SAVO/POHJOIS-SAVO_1
call digiroad_k2gpkg.bat POHJOIS-SAVO/POHJOIS-SAVO_2
call digiroad_k2gpkg.bat SATAKUNTA/SATAKUNTA_1
call digiroad_k2gpkg.bat SATAKUNTA/SATAKUNTA_2
call digiroad_k2gpkg.bat UUSIMAA/ITA-UUSIMAA
call digiroad_k2gpkg.bat UUSIMAA/UUSIMAA_1
call digiroad_k2gpkg.bat UUSIMAA/UUSIMAA_2
call digiroad_k2gpkg.bat VARSINAIS-SUOMI/VARSINAIS-SUOMI_1
call digiroad_k2gpkg.bat VARSINAIS-SUOMI/VARSINAIS-SUOMI_2
TIME /t
```

Komentojono 2.

kokoa\_digiroad\_k2gpkg.bat  
=====

```
ogr2ogr -f gpkg -update -append -addfields -dim 2 digiroad_k.gpkg
/vszip/KokoSuomi_DIGIROAD_K_EUREF-FIN.zip/KokoSuomi_DIGIROAD_K/%1 -oo
encoding="ISO-8859-1"
```

## Tulosten tarkastelu

### Muunnokseen tarvittava aika

Laitteisto: Kannettava Windows 7 tietokone Intel i7-prosessorilla ja SSD-kovalevyllä  
Muunnosaika XY-geometrioilla: **16 minuuttia**

### Tiedostokoot

KokoSuomi\_DIGIROAD\_K\_EUREF-FIN.zip: 2,12 Gt  
Zip purettuna: 7,86 Gt  
GeoPackage XYZM-koordinaateilla: 5,56 Gt  
GeoPackage XY-koordinaateilla: 3,53 Gt

### Kohteiden määrät aineistossa

1	DR_AJONEUVOKOHTAINEN_RAJOITUS_K	15523
2	DR_ESTERAKENNELMA	44070
3	DR_KAANTYMISSRAJOITUS_K	3607
4	DR_KAISTOJEN_LUKUMAARA_K	408125
5	DR_LEVEYS_K	2167016
6	DR_LIIKENNEVALO	6610
7	DR_LINKKI_K	3500112
8	DR_NOPEUSRAJOITUS_K	2633367
9	DR_PAALLYSTETTY_TIE_K	824244
10	DR_PALVELU	6194
11	DR_SUOJATIE	31574
12	DR_SUURIN_SALLITTU_AKSELIMASSA_K	296
13	DR_SUURIN_SALLITTU_KORKEUS_K	1230
14	DR_SUURIN_SALLITTU_MASSA_K	1360
15	DR_YHDISTELMAN_SUURIN_SALLITTU_MASSA_K	383
16	DR_EUROOPPATIENRO_K	20532
17	DR_JOUKKOLIIKENNEKAISTA_K	904
18	DR_KAANTOPAikka_K	4335
19	DR_KELIRIKKO_K	89228
20	DR_LIIKENNEMAARA_K	344990
21	DR_LIITTYMANRO_K	698
22	DR_OPASTUSTAULU	8959
23	DR_RAUTATIEN_TASORISTEYS	1693
24	DR_SUURIN_SALLITTU_LEVEYS_K	505
25	DR_SUURIN_SALLITTU_PITUUS_K	827
26	DR_SUURIN_SALLITTU_TELIMASSA_K	638
27	DR_TALVINOPEUSRAJOITUS_K	32768
28	DR_VALAISTU_TIE_K	176266
29	DR_VAK_RAJOITUS_K	23386