

Tietokanta-aineistojen siirron suunnitelma

Tähän dokumenttiin on kirjattu Gispo Oy:n huomiota tietojen siirtoon olemassa olevista tietokannoista avoimen tietokannan mukaiseen tietokantaan.

Muutoshistoria

Päivä	Muutos	Tekijä
08.06.2018	yk048_0401_2018 (keskustan strateginen osayleiskaava) käyty läpi ja lisätty dokumenttiin	Erno Mäkinen (erno)
09.06.2018	yk2040_19122017 (Kantakaupungin yleiskaava 2040) käyty läpi ja lisätty dokumenttiin ja lisätty luvut "Tiedonsiirto olemassa olevista tietokannoista" ja "näkömien luonti" sekä siistitty rakennetta	erno
14.6.2018	yk049_testiaineisto:n mukaan otto suunnitelmaan	erno
18.6.2018	Siirretty tietokannasta toiseen siirtämiseen käytetyt SQL-kyselyt omiin .sql-tiedostoihinsa	erno
8.8.2018	Siistitty ja tarkennettu sisältöä	erno
13.8.2018	Päivitetty Virrassa oleva SQL-skripti-tiedostolinkki	erno
3.9.2018	Selkeytetty dokumenttia Sanna Jokelan kommenttien pohjalta	erno

Sisällysluettelo

[Johdanto](#)

[Tietokantojen ja tiedonsiirron valmistelu](#)

[Lähdetietokanta-aineiston tuonti PostgreSQL-tietokantaan](#)

[Kohdetietokannan luonti](#)

[Lähdetietokannan taulujen kopiointi kohdetietokantaan](#)

[Tiedonsiirto lähdetietokannasta kohdetietokantaan](#)

[Tiedonsiirtoon käytettävien SQL-lauseiden ajaminen](#)

[Yhteenveto](#)

Johdanto

Avoimen tietokannan suunnittelussa on on pyritty noudattamaan seuraavia periaatteita.

- Tiedonsiirto tietokantojen välillä riittävän helppoa
- Paperisten ja digitaalisten yleiskaava-karttojen luonti mahdollista
 - SLD-tiedostojen luonti riittävän helppoa (GeoServerillä/QGIS:ssä)
- Inspire-yhteensopiva
- Tukee kaavoittajien suunnittelu- ja muuta työtä
- Käsitemallin, tulevaisuuden yleiskaavoituksen toiveiden, prosessikaavion, toimijat-, jne. dokumenttien sisältö huomioitu

Lisäksi huomioitavia asioita tiedonsiirrossa ovat seuraavat ja lähtöaineisto vaikuttaa kaikkiin muihin paitsi indeksien luontiin:

- Lähde- ja kohde-tietokantojen geometriatyyppeiden mahdolliset erot
- Lähde- ja kohde-tietokantojen koordinaattijärjestelmät
- Lähde- ja kohde-tietokantojen merkistöt
- Spatiaaliset indeksit tulee aina luoda kohdekantaan siirretylle aineistolle

Seuraavassa luvussa kuvataan miten olemassa oleva tietokanta-aineisto voidaan valmistella ja avoimen tietomallin mukainen tietokanta luoda ja valmistella, jotta olemassa oleva tietokanta-aineisto saadaan siirrettyä avoimen tietokantamallin mukaisen tietokannan tauluihin. Aineistojen siirto vanhasta tietokannasta uuteen tietokantaan on kuvattu tämän jälkeen omassa luvussa.

Tässä oletetaan, että käyttäjällä on perusosaaminen PostgreSQL-tietokannan hallinnasta ja käytöstä sekä psql-komentorivityökalulla että pgAdmin-sovelluksella ja myös PostgreSQL-tietokanta, sen PostGIS-paikkatietolaajennos sekä PostgreSQL:n komentorivityökalut asennettu ja löytyvät PATH-ympäristömuuttujasta. Myös pgAdmin-työkalu tulee olla asennettuna.

Tietokantojen ja tiedonsiirron valmistelu

Lähdetietokanta-aineiston tuonti PostgreSQL-tietokantaan

Tämä riippuu lähdeaineistosta. Tampereen kaupungin testiaineistolla, joka löytyy GitHubista osoitteesta <https://github.com/GispoCoding/Tampere-KDYK>, tuonti onnistuu pgAdmin 4 -työkalun Tools | Restore... -toiminnolla:

1. Käynnistetään pgAdmin 4
2. Avataan yhteys tietokantapalvelimeen, vähintään PostgreSQL v9.6, jolle lähde- ja kohdetietokannat tulevat. Palvelimella tulee olla PostGIS-laajennos.
3. Luodaan lähdetietokantaa varten tyhjä tietokanta nimeltä `yk049_testiaineisto`
4. Aja luodulle tietokannalle kyselyt:
 - a. `CREATE EXTENSION POSTGIS;`
 - b. `CREATE ROLE jussitah;`
5. Valitaan Tools-valikosta Restore...-toiminto
6. Avautuvan dialogin Filename-kohtaan tulee Virrasta haetun ja puretun zipin sisältämä `yk049_testiaineisto.backup`-tiedosto, rooliksi valitaan Postgres ja muut asetukset voi jättää oletusarvoihin. Napautetaan Restore-painiketta

Kohdetietokannan luonti

1. Haetaan GitHubista, <https://github.com/GispoCoding/Tampere-KDYK>, viimeisin avoimen tietokannan .sql-tiedosto

2. Käynnistetään pgAdmin 4 tai psql-komentorivityökalu, avataan yhteys samaan tietokantapalvelimeen kuin aiemmin ja luodaan tietokanta, esim.
2018_06_18_avoin_yleiskaava ja sille PostGIS-laajennos: CREATE EXTENSION POSTGIS;
3. Ajetaan luodulle tietokannalle GitHubista haetun SQL-tiedoston kyselyt, joko pgAdminin kyselytyökalun kautta tai psql-komentorivityökalulla (esim. Windowsin komentoikkunassa navigoi ensin PostgreSQL-kansioon kohtaan bin) esim. komennolla
`psql -p 5432 -d 2018_06_18_avoin_yleiskaava -f 2018-06-18-avoin_yleiskaava.sql`

Testiaineiston tiedonsiirron kannalta tarpeelliset enumeraatio-tyypit arvoineen luodaan automaattisesti, mutta jos kyseessä on jokin muu aineisto tai tietokantaa ylläpidetään, niin tarvittaessa enumeraatio-tyyppien arvojen luonti tapahtuu kätevimmin FME-työkalulla hyödyntäen CSV-tiedostoja, jotka saa INSPIRE-tietomallin tapauksessa CSV-muodossa osan suomenkielisenä EU:n sivustolta. Esimerkiksi täydentävien kaavamerkintöjen lista löytyy useilla eri kielillä ja useissa eri muodoissa sivulta:

<http://inspire.ec.europa.eu/codelist/SupplementaryRegulationValue>.

Lähdetietokannan taulujen kopiointi kohdetietokantaan

Tietokanta-aineistojen siirto lähdetietokannan tauluista onnistuu psql- ja pgAdmin-työkalussa SQL-kyselyillä, kun lähde- ja kohdetietokannan taulut ovat samassa tietokannassa ja siksi lähdetietokannan taulut siirretään kohdetietokantaan ennen tietokanta-aineistojen siirtoa. Seuraava komentorivi-komento toimii lähdetietokannan taulujen siirtämisessä kohdetietokantaan mikäli lähde- ja kohdetietokanta on nimetty kuten edellä on ehdotettu, tietokantapalvelin sijaitsee oletusportin 5432 takana ja pg_dump- sekä psql-komentorivityökalut on lisätty järjestelmän polkuun (Windowsissa mahdollisesti sijainti on C:\> cd Program Files (86x)\PostgreSQL\9.6\bin):

```
pg_dump -U postgres -p 5432 -O -n yk049_testiaineisto yk049_testiaineisto | psql -U postgres  
-d 2018_06_18_avoin_yleiskaava -U postgres -p 5432
```

Huom! Käyttäjän postgres salasana pitää syöttää kaksi kertaa peräkkäin ko. komennon ajamisen aikana heti ensimmäisen kerran jälkeen vaikka siitä ei erillistä ilmoitusta komentoikkunaan tulekaan. Ko. komento siis kopioi yk049_testiaineisto-tietokannan yk049_testiaineisto-skeeman sisältöineen 2018_06_18_avoin_yleiskaava-tietokantaan. Parametrin -O takia ei kohdetietokannassa omistajuutta palauta samaksi kuin lähdetietokannassa.

Tiedonsiirto lähdetietokannasta kohdetietokantaan

Testauksessa käytetään yk049_testiaineisto-tietokantaa. Tiedonsiirto tehdään psql- tai pgAdmin-työkalun kautta SQL-kyselyillä.

Tiedonsiirtoon käytettävien SQL-lauseiden ajaminen

Kutakin lähdetietokannan taulua kohden on oma sql-tiedosto, jonka avulla tiedonsiirron saa tehtyä testiaineistosta, joka on valmisteltu edellä kuvatusti avoimen tietomallin mukaiseen kohdetietokantaan. Testiaineistoa varten luodut SQL-tiedostot löytyvät GitHubista: <https://github.com/GispoCoding/Tampere-KDYK>. Tiedonsiirroissa käytettävät SQL-tiedostot on nimetty numeroituvasti 01, 02, jne. ja ensiksi tulee siis ajaa 01-alkuisen tiedoston sisältö, sitten 02-alkuisen tiedoston sisältö, jne. Ne saa ajettua kohdetietokannalle psql-komentorivityökalun kautta antamalla työkalulle polun ensimmäiseen SQL-tiedostoon parametrina ja toistamalla tämän muiden tiedostojen kanssa tai pgAdmin-sovelluksen kautta kopioimalla ensimmäisen tiedoston sisällön kysely-työkaluun, ajamalla kyselyt ja toistamalla tämän muille SQL-tiedostoille. Komennon muoto on esimerkiksi:

```
psql -p 5432 -d 2018_06_18_avoin_yleiskaava -f 01_kaava_alueet_polygon.sql
```

Yhteenveto

Edellä kuvattujen vaiheiden jälkeen alkuperäinen lähtöaineisto on siirretty avoimen tietomallin mukaiseen kohdetietokantaan. Tämän jälkeen ko. aineistoa voi hyödyntää esimerkiksi GeoServer- ja QGIS-ohjelmistojen avulla. Hyödyntämisessä oleellista on ymmärtää tietokannan taulujen, taulujen kenttien ja taulujen välisten relaatioiden merkitykset. Näistä kuvaukset tuotetaan omaan dokumenttiinsa.