

# Kokeiluympäristöissä syntyy uusia ratkaisuja paikkatietoalalle

Open Geospatial Consortium (OGC) -järjestön kokeiluympäristöissä syntyy uusia ratkaisuja, kun alan huiput pääsevät tutkimaan ja kokeilemaan ideoitaan yhdessä. Mitä tapahtuikaan viimeisimmän projektin aikana?

JARI REINI

**Kokeiluympäristöt (testbed)** ovat tärkeä osa OGC:n innovaatio-ohjelmaa (Innovation Program). Vajaan vuoden kestävässä projekteissa paikkatietoalan huiput kokoontuvat yhteen ja miettivät uusia ratkaisuja alan polttaviin kysymyksiin. Tutkimustyön ja kokeilujen kautta osa tuloksista päättyy standardeiksi ja osa synnyttää arvokasta tietoa siitä, mikä on mahdollista ja mikä ei. Kaikki tulokset julkaistaan OGC:n verkkosivuilla.

Jo ensimmäinen kokeiluympäristö vuonna 1999 oli menestys, jonka tuloksena syntyi OGC:n ehkä käytetyin standardi ”Web Map Service” (WMS). Sen jälkeen kokeiluympäristöjä on ollut käynnissä lähes vuosittain. Viimeisin OGC Testbed-16 päättyi vuoden vaihteessa, ja uusi OGC Testbed-17 on jo käynnistymässä tarjouskilpailun myötä.

Mitä sitten tapahtui vuodenvaihteessa päättyneessä projektissa? Lue tästä jutusta tärkeimmät kohokohdat. Tarkemmin tuloksia on kuvattu pro-

jektin verkkosivuilla ([www.ogc.org/testbed16](http://www.ogc.org/testbed16)), josta löytyy muun muassa videotallenne ja materiaalit projektin tuloksia esitelleestä verkkoseminaarista.

## GeoPackage-tiedostot löytyvät ja toimivat nopeammin

GeoPackage on avoin, standardisoitu ja alustariippumaton paikkatietokanta, joka koostuu kokonaan yhdestä tiedostosta. Geopackagen haasteina ovat olleet puuttuvat metatiedot, kuvaustekniikan puute sekä heikko suorituskky erittäin suurilla vektoriaineistoilla. Näihin Testbed-16-projekti kehitti ratkaisuja.

GeoPackage-tietokannoille luotiin metatietoprofiili, jonka avulla niiden löydettävyys parani. Samalla niille kehitettiin tyylilaajennos (style extension), jossa GeoPackage-tietokannan tauluihin pystyttiin liittämään kuvaustekniikkaa.

GeoPackage-tietokannoissa käytetään SQLite-tietokantamootoria. Sen suorituskky riittää hyvin kohtuul-

lisen kokoisilla tietomassoilla, mutta ongelmia ilmeni, kun tietoauneiston koko oli kymmeniä gigatavuja tai tätä enemmän.

Tiedostojen kokoa pienennettiin käyttämällä koodilistoja (enumerations). Koodilistat tallennettiin gpkg\_schema-laajennosta hyödyntäen lukuarvoina tai yksittäisinä merkkeinä, ja niiden ihmisluevat vastineet löytyvät gpkg\_data\_column\_constraints-taulusta.

Tiedostokoon lisäksi suorituskkyä pyrittiin parantamaan segmentoimalla tietoja, koodaamalla geometria uudella tavalla, yleistämällä tietoja ja järjestämällä tietokantataulujen sisältöä uudella tavalla.

## Tiedoille lisää turvaa API-rajapinnalta

Monet tiedot ovat nykyään tarjolla pilviympäristössä. Tämä aiheuttaa haasteen tietosuojalle. Jakelukanavien viidakossa tietosuoja ei voi enää perustua jakelukanavan tietosuojaan palomuureineen, vaan suoja on oltava

tiedossa itsessään. Lisäksi valtuutettujen käyttäjien on pystyttävä lukemaan tietoja myös silloin, kun verkkoyhteyttä ei ole käytettävissä.

Testbed-16-projekti jatkoi datakeskeisen turvallisuuden (Data Centric Security) kehittämistä. Kokeiluympäristössä käyttötapauksina olivat suoratoisto (streaming) sekä suojatun tiedon kysely ja käyttö ilman verkkoyhteyttä (offline).

Projekti määritteli uuden arkkitehtuurisen Key Management Server (KMS) -komponentin, jota käytetään API-rajapinnan kautta. Rajapintaa käytettiin avainten rekisteröintiin tietojen suojausta varten sekä myös asiakkaiden tarpeisiin tietojen tarkistamiseksi ja salauksen purkamiseksi.

### **Koneoppimisesta apua paikkatietojen hyödyntämiseen?**

Koneoppimisen (machine learning) avulla pilvialustoilla voidaan kerätä ja ottaa käyttöön suuria määriä paikkatietoaineistoja. Tämä vaatii kuitenkin, että koneoppimisen työkaluja yhdistetään OGC-standardien mukaisesti metatietoihin sekä tiedon jakeluun ja prosessointiin.

Testbed-16-projektin tavoitteena oli löytää ja käyttää uudelleen ennustavissa koneoppimisen malleissa käytettyjä tietoja sekä integroida koneoppimisen mallit standardeihin perustuvaan (paikka)tietoinfrastruktuuriin.

Kokeilussa pohdittiin muun muassa näitä kysymyksiä: Edellyttääkö koneoppiminen tietojen yhteentoimivuutta vai voiko se mahdollistaa yhteentoimivuuden? Tarvitseeko koneoppimisessa käytetyn data olla valmista analysoitavaksi (Analysis Ready Data, ARD)? Mikä merkitys datakuutioilla (datacubes) on koneoppimisessa?

## **Kaikesta ei synny standardeja, mutta jo ratkaisujen etsiminen yhdessä on arvokasta.**

Lisäksi projekti tutki, miten Map Markup Language (MapML) -tekniikkaa voitaisiin hyödyntää kustannustehokkaasti visualisoinneissa ja hauissa.

### **Mukaan kehittämään uusia ratkaisuja**

Kokeiluympäristöissä syntyy paljon kaikenlaista. Tässä on mainittu vain muutama esimerkki Testbed-16-projektissa syntyneistä tuloksista. Kaikista niistä ei synny standardeja, mutta pohdinta ja ratkaisujen etsiminen yhteisiin paikkatieto-ongelmiin on usein jos sinänsä arvokasta.

Jos haluat mukaan projekteihin tai olisit kiinnostunut rahoittamaan toimintaa, ota yhteyttä allekirjoittaneeseen.

### **Lisätietoa**

Open Geospatial Consortium (OGC) on kansainvälinen vuonna 1994 perustettu standardisointijärjestö, jonka tehtävänä on tehdä paikkatiedoista ja paikkatietopalveluista löydettäviä, saavutettavia, yhteentoimivia ja uudelleenkäytettäviä (FAIR - Findable, Accessible, Interoperable, and Reusable).

OGC Testbed-16-projekti:  
[www.ogc.org/testbed16](http://www.ogc.org/testbed16)

---

Kirjoittaja työskentelee johtavana asiantuntijana Maanmittauslaitoksella. Hän on ollut vuosia mukana OGC:n eri työryhmissä Maanmittauslaitoksen edustajana.  
ETUNIMI.SUKUNIMI@  
MAANMITTAUSLAITOS.FI

---