

# Ontologioilla paikkatietojen sisältö hallintaan

KUVA: RIIKKA HENRIKSSON

**Kuinka saada tietokone tulkitsemaan käsitteitä ja suhteita? Ratkaisu löytyy ontologiasta. Ontologia on keino tulkita paikkatietoaineiston sisältöä ja merkitystä käyttäjän käsitteillä tiedontuottajan luokittelusta riippumatta.**

Semanttisessa webissä mallinnetaan semantiikkaa koneiden käyttöön toisin kuin perinteisessä webissä, jonka dokumentit on suunnattu ihmisille. Tietämyksen tarkassa tallennuksessa ja valtavien tietomäärien käsittelyssä käytetään apuna ontologioita.

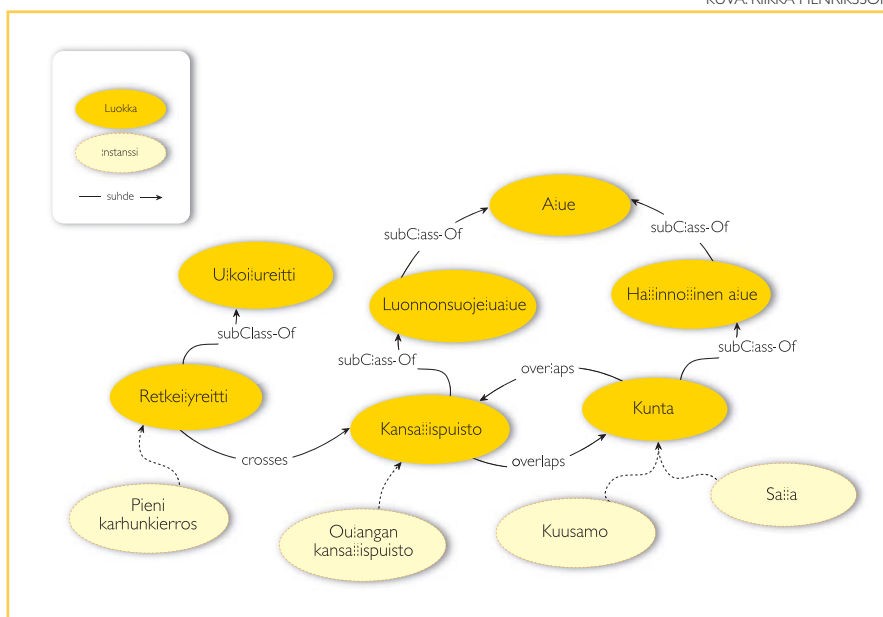
Ontologioiden ja niihin liittyvien menetelmien tutkimus ja sovellus on maailmalla tällä hetkellä laajaa. Suomessakin ollaan aktiivisia: esimerkiksi Tekesin rahoittamassa FinnONTO-projektin osahankkeessa ”Suomalaiset paikkaontologiat” tutkitaan paikkatietoon liittyvän tietämyksen esittämistä ja päättelymekanismeja käyttäen hyväksi semanttisen webin ontologiatekniikoita.

## Maantieteellisen todellisuuden käsitteet

Ontologia tarkoittaa formaalia, tietokoneen tulkittavissa olevaa kuvausta tietyn sovellusalueen käsitteistä ja sellaisten käsitteiden välisistä suhteista, jonka jokin toimijajoukko keskenään jakaa.

Paikkaontologialla tarkoitetaan siten maantieteellisesti jäsenyneen todellisuuden käsitteitä (esimerkiksi kaupunki, järvi, saari ja retkeilyalue) eli käsitteitä, joilla viitataan mm. ihmisten muodostamiin territoriaalisiin ja sosiospatiaalisiin tiloihin ja paikkoihin tai luonnon määrittelemiä alueisiin ja niiden välisiin suhteisiin.

Ontologiassa luokat edustavat käsitteitä, jotka on usein järjestetty hierarkiaan. Hierarkia mahdollistaa yliluokalle määriteltävien ominaisuuksien ja suhteiden periytymisen aliluok-



*Ontologia tarkoittaa tietokoneen tulkittavissa olevaa kuvausta käsitteistä ja niiden välisistä suhteista. Ontologiassa luokat edustavat käsitteitä, jotka on usein järjestetty hierarkiaan.*

kalle. Luokilla voi olla ilmentymiä, joita kutsutaan instansseiksi. Suhteet edustavat luokkien välisiä assosiaatioita. Esimerkiksi periytymishierarkian rakentamiseen voidaan käyttää suhdetta subClass-Of. Instansseille – kuten myös niitä edustaville luokille – tulee myös määritellä globaalisti yksilöivä URI-tunnus (Universal Resource Identifier).

Paikkaontologioita rakennettaessa luokkien välille kuvataan perinteisten hierarkkisten suhteiden lisäksi erilaisia spatiaalisia suhteita. Välttämättömiä ovat topologiset suhteet kuten päällekkäisyys (overlaps), leikkaavuus (crosses) ja erillisuus (disjoint), sekä mereologiset osa-kokonaisuus-suhteet (is-part-of). Lisäksi voidaan kuvata esimerkiksi seuraavia paikkojen välisiä kvalitatiivisia ja kvantitatiivisia suhteita:

- 1) peittävyys-suhteita (covers), joille voidaan antaa peittävyys arvo esimerkiksi prosentteina,
- 2) etäisyys-suhteita ilmaistuna esimerkiksi kilometreissä tai matkustusajassa,

3) sumeita etäisyys-suhteita (near, far, jne.) ja

4) orientaatiota ilmaisevia suhteita (north-of, south-of, jne.).

Paikkaontologioissa instanssit ovat paikkatietokohteita eli kohteita, joiden sijainti Maan suhteen tunnetaan. Instanssien georeferointi voi yksinkertaisimmillaan tapahtua pelkän paikannimen perusteella, mutta nimet ovat harvoin yksilöllisiä ja voivat muuttua tai jopa hävitä historian saatossa. Tästä johtuen paikkainstansseille on syytä ilmoittaa ominaisuutena sen koordinaattitieto. Koordinaattitieto esitetään instanssia kuvaavan geometrian kautta – eli joko pisteenä, käyränä tai polygonina.

Paikkaontologioiden keskeisiä lähteitä ovat maantieteelliset paikannimihakemistot ja -rekisterit. Suomessa tällaista lähdeä edustaa mm. Maanmittauslaitoksen tuottama ja ylläpitämä Paikannimirekisteri, joka sisältää yli 800 000 paikannimeä koordinaattitietoineen. Tämä aineisto toimii instanssiaineistona

KUVA: ANTERO AALTONEN



*Paikkaontologialla tarkoitetaan maantieteellisesti jäsenytyneen todellisuuden käsitteitä (esimerkiksi kaupunki, järvi, saari ja retkeilyalue) eli käsitteitä, joilla viitataan ihmisten muodostamiin tiloihin ja paikkoihin tai luonnon määrittelemiin alueisiin ja niiden välisiin suhteisiin.*

FinnONTO-projektin puitteissa rakennettavassa Suomalaisessa paikkaontologiassa (SUO). SUO sisältää kaikkiaan yli 150 hierarkkisesti jäsenettyä, maantieteellisiin tai geometrisiin käsitteisiin viittaavaa luokkaa. Luokkien välisistä spatiaalisista suhteista on esitetty muun muassa topologiset ja mereologiset suhteet.

#### **Miten muutokset hallintaa?**

Paikkaontologioissa ongelmaiseksi muodostuu helposti se, että maailma ja sen myötä sitä kuvaava maantieteellinen käsitteistö muuttuu. Kun esimerkiksi vuoden 2007 alussa useat kunnat yhdistyivät ja kuntien kokonaismäärä väheni, niin myös paikkaontologiaa on päivitettävä.

Muutosten mallinnukseen ja hallintaan voidaan tehdä ontologia-aikasarja muutossiltojen avulla. Suomen ajallinen paikkaontologia (SAPO) käyttää näitä menetelmiä tallentamaan kattavasti viimeisen noin 150 vuoden aikana tapahtuneet muutokset suomalaisissa kunnissa.

SAPO:n ja päättelykoneen avulla

saadaan muun muassa selville, että nykyinen Lappeenranta peittää 1900-luvun vaihteen Viipurista yli kymmenen prosenttia. Tämä peittävyys on ontologiassa ilmaistu coversuhteen avulla. Esimerkiksi sen ajan Viipurilla indeksoidut tiedot, kuten eri kylien kartat tai valokuvat ovat siis todennäköisesti osittain peräisin nykyisen Suomen alueelta. Tämä on tietoa, jota voidaan hyödyntää aineistojen hallinnassa ja haussa.

Paikkaontologioista on hyötyä myös monessa muussa yhteydessä. Paikkatietoaineistojen tuottajien näkökulmasta katsottuna ontologioiden avulla voidaan välittää mahdollisille käyttäjille tietoaineiston semanttikka eli merkityssisältö riippumatta siitä, onko sijaintitiedon esitystapa vektori- vai rasterimuotoista. Kattavat metatietokuvailutkaan eivät tähän vielä riitä.

Ontologian käsitteistöä voidaan käyttää helpottamaan erilaisten resurssien metatietojen dokumentoimista aina paikkatietoaineistoista museoiden kokoelmiin asti (esi-

#### **Lisätietoja:**

FinnONTO-projektin osahankkeeseen ”Suomalaiset paikkaontologiat” etsitään seuraavaan vaiheeseen (2008-) esimerkkiaineistoja. Ontologiapohjaisten menetelmien ja työkalujen käytöstä kiinnostuneita organisaatioita pyydetään ottamaan yhteyttä kirjoittajiin (Riikka.Henriksson@tkk.fi tai Tomi.Kauppinen@tkk.fi). Lisätietoa FinnONTO-projektista löytyy Semanttisen laskennan tutkimusryhmän sivuilta osoitteesta: <http://www.seco.tkk.fi/>.

merkiksi esineiden löytöpaikka). Paikkaontologiat tarjoavat myös uuden tavan tietoaineistojen paitsi keskinäiseen vertailemiseen, myös sisällölliseen ja käsitteelliseen harmonisointiin. Eri järjestelmiä voidaan myös integroida tehokkaammin, kun mahdollisimman moni toimija jakaa saman käsitteistön.

RIIKKA HENRIKSSON  
TOMI KAUPPINEN

Tutkija Riikka Henriksson työskentelee Teknillisen korkeakoulun Geoinformaatio- ja paikannustekniikan laboratoriossa FinnONTO-projektissa.  
Sähköposti: Riikka.Henriksson@tkk.fi

Tutkija Tomi Kauppinen työskentelee Teknillisen korkeakoulun viestintätekniikan laboratoriossa ja Helsingin yliopiston tietojenkäsittelytieteen laitoksella FinnONTO-projektissa.  
Sähköposti: Tomi.Kauppinen@tkk.fi