

PAIKKATIETOALUSTA

Selvitys paikkatietoaineiston metatietoprofiilin tarpeista ja käytöstä

Jukka Lepistö

Siili Solutions Oyj | toimeksiantaja Maanmittauslaitos | Marraskuu 2018

SISÄLLYS

1	YHTEENVETO	2
2	JOHDANTO	2
2.1	NYKYTILANNE.....	3
2.2	TAVOITETILA.....	3
3	TEORETTINEN VIITEKEHYS JA TEHTÄVÄ	3
4	MENETELMÄT	3
4.1	HAASTATELUUN OSALLISTUNEET ORGANISAATIOT JA HAASTATELUPÄIVÄMÄÄRÄT	4
4.2	OSALLISTUNEET HENKILÖT ORGANISAATIOITTAIN	4
4.3	HAASTATELUIDEN AIHEALUEET JA KYSYMYKSET.....	5
5	ANALYYSI	7
5.1	PAIKKATietoaineistojen tarve ja käyttö omassa organisaatiossa.....	7
5.2	AINEISTOJEN LÖYTÄMISEN, HANKINNAN JA AJAN TASALLA PITÄMISEN KÄYTTÖTAPAUKSET	8
5.3	KATTAVIEN PAIKKATietoaineistojen hyödyllisyys ja tiedon jalostus.....	10
5.4	TUNNETUT JA KÄYTTÖKELPOISET AINEISTOJEN LUOKITTELUTAVAT, ASIANSANASTOT JA AVAINSANAT	11
5.5	MUUT HUOMIOT	13
6	JOHTOPÄÄTÖKSET	14
6.1	PAIKKATietoaineiston etsimiseen ja löytämiseen liittyvät johtopäätökset.....	14
6.2	PAIKKATietoaineiston soveltuvuuteen liittyvät johtopäätökset.....	15
	LÄHTEET	17

1 YHTEENVETO

Selvityksen tavoitteena oli selvittää paikkatietoaineiston tuottajien ja hyödyntäjien tarpeita liittyen metatietoprofiiliin. Tämän perusteella on tarkoitus muodostaa yhteinen paikkatietoaineiston metatietoprofiili, jonka avulla voidaan kehittää aineistojen löydettävyyttä sekä parantaa aineistojen soveltuvuuden arviointia.

Selvityksen yhteydessä haastateltiin puolistrukturoituina haastatteluina paikkatietoaineiston tuottajia ja kuluttajia Rambollin, Sitowisen, Forum Viriumin, Suomen Ympäristökeskuksen, Tilastokeskuksen sekä Ympäristöministeriön organisaatioista. Kaikille haastateltaville esitettiin samat kysymykset liittyen paikkatietoaineiston tarpeeseen ja käyttöön organisaatiossa, aineistojen löytämiseen, hankinnan ja ajan tasalla pitämiseen, kattavien paikkatietoaineistojen hyödyllisyyteen ja tiedon jalostukseen sekä tunnettuihin ja käyttökelpoisiin aineistojen luokittelutapoihin, asiasanastoihin ja avainsanoihin liittyen.

Haastatteluiden johtopäätöksenä voidaan todeta, että metatietoprofiililla on paikkatietoaineistojen suhteen kaksi toisestaan poikkeavaa käyttötarkoitusta. Ensiksi metatietoprofiilin on tärkeässä roolissa etsimiseen ja löytämiseen liittyvissä käyttötapaüksissa, jolloin metatietoprofiilissa arvioitavia tietoja arvioidaan objektiivisesti. Toiseksi sen avulla tulisi voida arvioida löydetyn paikkatietoaineiston soveltuvuutta subjektiivisesti omaan käyttötarkoitukseen.

Digitalisoinumisen myötä nousseet uudet vaatimukset niin paikkatietoaineistolle kuin sen metatietoprofiilille, edesauttaisivat paikkatietoaineistojen käsittelyä tehokkaammin automaattisesti, koneavusteisesti sekä ohjelmistojen avulla, sillä muun muassa aineistojen yhdistelyyn ja käsittelyyn liittyy myös kaikilla haastatteluihin osallistuneilla organisaatiolla manuaalista työtä, jonka automatisoimisessa ajan säästön ja kustannusten näkökulmasta. Myös yksityiskohtaisuuden tasolla todettiin olevan vaikutusta silloin kun erilaisia aineistoja tulisi yhdistellä.

Haastateltavilla tuntuu olevan hieman epäselvyyksiä siitä, mitä paikkatietoalusta sisältää ja mitä kautta saatavaan aineistoon pääsee käsiksi (saatavuus). Lisäksi ontologiaan liittyvät haasteet nousivat esille useammassa haastattelussa organisaatiossa. Haasteet liittyvät aineiston löytämiseen haettaessa esimerkiksi avainsanalla ”joki”, odotetaan hakutuloksen pitävän sisällään myös ”vesistöt” ja muut avainsanaan liittyvät aineistot. Lisäksi koodistojen kuten luokittelutapojen ja asiasanastojen hyödyntämisen suhteen toivotaan yhtenäisiä käytäntöjä.

Haastatteluista nousi esille myös, että aineiston soveltuvuus voidaan usein varmistaa vasta jälkikäteen ja on aikaa vievää, joten mahdollisimman aikaisessa vaiheessa tapahtuva soveltuvuuden arviointi olisi hyödyllistä. Lisäksi paikkatietoaineiston soveltuvuudessa laatu ja aineiston ajantasaisuus ovat oleellista ja myös aineiston homogeenisuuteen kiinnitettiin huomioita.

2 JOHDANTO

Paikkatietoalusta-hanke yhtenäistää valtion, maakuntien ja kuntien paikkatiedot ja tuo ne myös yritysten ja yhteisöjen saataville. Paikkatietoalusta eli Julkisen hallinnon yhteinen paikkatietoalusta -hanke on osa hallituksen Digitalisoidaan julkiset palvelut -kärkihankekokonaisuutta.

Tavoitteena on yhtenäistää ja parantaa julkisen hallinnon sähköisiä palveluita, parantaa ja tehdä läpinäkyvämmäksi tietopohjaista päätöksentekoa sekä säästää kustannuksia julkisessa hallinnossa. Hanke koostuu kahdeksasta osahankkeesta, joista Tekninen alusta -osahankkeen puitteissa tehtiin tämä selvitys.

2.1 NYKYTILANNE

Suomen paikkatietoinfrastruktuurin palvelujen ja aineistojen löydettävyyks on tällä hetkellä korkeintaan kohtalainen. Keskitettyjä palveluita, kuten metatietohakemistoa ja luettelopalvelua ei hyödynnetä niin hyvin kuin olisi mahdollista. Paikkatietopalvelujen ja -aineistojen tilannekuva on puutteellinen eivätkä tiedon hyödyntäjät saa tietoa niistä kootusti ja helpopolukuisesti.

Useiden tahojen ylläpitämien tietojen, kuten maankäyttöpäätösten, osoitteiden tai erilaisten aluejakojen yhdistäminen kansallisiksi tai INSPIRE-tietotuotteiksi vaatii erityisiä ponnisteluja. Olemassa olevien palvelurajapintojen tuki standardeille on joissakin tapauksissa puutteellinen eikä käytettävissä ole mekanismeja toteutusten validointiin. Nykyiset alustat eivät ole keskenään yhteen toimivia tai niiden välille ei ole muodostettu yhteyksiä.

2.2 TAVOITETILA

Tavoitetilassa paikkatiedot ja palvelut ovat löydettävissä älykkään, myös hakukoneisiin kytkeytyvän haun kautta. Kaikkia aineistoja ja palveluja ei ole tarkoituksenmukaista keskittää yhdelle alustalle, vaan on mahdollistettu yhteen toimivien alustojen toimiminen hajautettuna verkostona. Metatieto- ja luettelopalveluihin on viety tarvittavat tiedot tukemaan aineistojen, palvelujen ja muiden alustojen nykyistä monipuolisempaa hyödyntämistä.

Aineistojen, palvelujen ja erilaisten alustojen liittäminen paikkatietoinfrastruktuuriin tapahtuu toimivan ja selkeästi ohjeistetun palveluketjun avulla. Paikkatiedot ja palvelut ovat löydettävissä intuitiivisen hakutoiminnon avulla.

Paikkatietoalusta on portti toimivaan paikkatietoinfrastruktuuriin, jonka päälle voi rakentaa erilaisia paikkatietoa hyödyntäviä ekosysteemejä ja innovatiivisia lisäarvopalveluita.

3 TEOREETTINEN VIITEKEHYS JA TEHTÄVÄ

Älykkään haun yksi kulmakivistä on taustalla olevan paikkatietoaineiston selkeästi määritelty ja hyödynnetty metatietoprofiili ja sen kuvaus. Tavoitteena oli selvittää paikkatietoaineiston tuottajien ja hyödyntäjien tarpeita liittyen metatietoprofiiliin.

Paikkatiedon metatietoa on käsitellyt Paula Ahonen-Rainio väitöskirjassaan *Visualization of Geospatial Metadata for Selecting Geographic Datasets*.

4 MENETELMÄT

Selvitys tehtiin puolistrukturoituna haastatteluina paikkatietoaineiston tuottajien ja hyödyntäjien organisaation asiantuntijoille. Organisaation asiantuntijoita saattoi osallistua useita yhtä aikaa samaan

haastatteluun. Haastattelut suoritettiin yhtä lukuun ottamatta organisaation omissa tiloissa. Haastatteluissa suurin osa haastateltavista oli läsnä, mutta Sitowisellä haastateltavia oli myös etänä puheyhteyden päässä.

4.1 HAASTATTELUUN OSALLISTUNEET ORGANISAATIOIT JA HAASTATTELUPÄIVÄMÄÄRÄT

1. Ramboll Oy - 29.8.2018
2. Sitowise Oy - 31.8.2018
3. Forum Virium Oy (Helsingin kaupunki) - 31.8.2018
4. Suomen ympäristökeskus - 4.9.2018
5. Uudenmaan ELY-keskus - 7.9.2018
6. Tilastokeskus - 24.9.2018
7. Ympäristöministeriö - 1.10.2018

4.2 OSALLISTUNEET HENKILÖT ORGANISAATIOITTAIN

- **Ramboll**
 - Veli-Pekka Koskela, GEO-yksikkö, paikkatiedot, tietomallinnus
 - Antti Miettinen, GEO-yksikkö, paikkatieto, GIS
 - Samuli Massinen, GEO-yksikkö, GIS, paikkatietoasiantuntijan tehtävät
 - Markus Hytönen, Vesi ja ympäristö, kaavoitus (paikkatieto-orientoitunut)
- **Sitowise**
 - Juha Liukas, infra ja tietomallinnuksen asiantuntija, taustaa softa ja geo-puolesta, alan standardin kehittämisessä mukana
 - Teemu Aaltio, pp kartio ja maastomallinnus, käytetään keilaus, ilmakuvia yms.
 - Taina Klinga, ympäristön ja maankäytön asiantuntija, tiehankkeet ja voimahankkeet, paikkatiedon päivittäinen käyttäjä, avointa dataa
 - Jan-Erik Berg, Tiekatusuunnittelu
 - Mika Virtanen, Paikkatietopuolen aineistotiimi, julkaisu ja kartta
 - Hannu Lammi, tiedonhallinnan palvelut infra-alalla
 - Pauliina Nissi, tiesuunnittelija, tiedonhallinta ja mallinnus, MML paikkatietoaineistoa käytetään, peruskartta ja rasterit, keilausaineisto ja maastotietomalli
- **Forum Virium (Helsingin kaupunki)**
 - Timo Ruohomäki, semanttinen kaupunkitietomalli, cityGML, BIM, WMS datan käyttöä
- **Suomen ympäristökeskus**
 - Kaisu Harju, vetää paikkatietopalvelut -ryhmää (vastaa Syken paikkatietopalveluiden järjestämisestä, softat, aineisto, metatieto)
 - Markus Törmä, geoinformatiikka tutkimusryhmästä. Vetää PTA-satelliittikuvat osahanketta.
 - Suvi Hatunen, Kaisun ryhmässä. Paikkatietoaineiston ja metatiedon ylläpitoa.
- **Uudenmaan Ely-keskus**
 - Anna Elf, projektipäällikkö, suunnittelun hankinta, tie- ja rakennussuunnittelu
 - Ari Tuppurainen, tierekisteri ja paikkatiedot. Pääsääntöisesti aineiston käyttäjä (vektori muotoisten aineistojen käyttöä)
 - Jaakko Kuha, suunnittelun hankinta, tie- ja rakennussuunnittelu
- **Tilastokeskus**

- Rina Tammisto, IT-erityispalveluiden esimies tietohallintoyksikössä. Paikkatiedon teknologia. Juuri tehty sijaintitiedon viitearkkitehtuuri, joka liittyy tietoarkkitehtuurityöhön (joka taas liittyy kokonaisarkkitehtuurityöhön).
- Mikko Saloila, metatietoasiantuntija tuotannon kehittäminen -yksikössä. Paikkatietoihin liittyvää työtä yleisesti metatietojen kautta.
- Tuukka Saranpää, kehittämisspäällikkö. Tilastotuotannon tietoarkkitehtuurin kehittäminen.
- **Ympäristöministeriö**
 - Topi Tjukanov, projektiasiantuntija PTA-maankäyttöpäätökset osahankkeessa

4.3 HAASTATTELUIDEN AIHEALUEET JA KYSYMYKSET

Haastatteluiden alussa kultakin haastateltavalta kysyttiin kuka olet ja mitä teet. Haastatteluiden kulku oli jaettu seuraaviin aihealueisiin ja kysymyksiin.

- Paikkatietoaineistojen tarve ja käyttö omassa organisaatiossa
 - Millaisten paikkatietoaineistojen käyttö on välttämätöntä tai erittäin hyödyllistä omassa tai lähioorganisaatiossi jokapäiväisessä työssä?
 - Tuleeko mieleen aineistoja, joiden käyttö tehostaisi tai helpottaisi arkea, mikäli ne olisivat käytettävissä? Jos, niin mikä rajoittaa ko. aineistojen käyttöä?
 - Tuleeko mieleen aineistoja, joiden käyttö on hankalaa tai tehotonta? Jos, niin miksi?
 - Anna yksi konkreettinen esimerkki omalle työllesi tarpeellisesta paikkatietoaineistosta. Miten työ muuttuisi, mikäli ko. aineistoa ei olisi saatavilla, tai sen laatuun ei voisi luottaa riittävästi?
 - Minkälaisia työkaluja ja välineitä käytät itse tai lähioorganisaatiossasi käytetään paikkatietojen käsittelyyn, analysointiin ja visualisointiin tai paikkatietopohjaiseen suunnitteluun ja raportointiin?
- Aineistojen löytämisen, hankinnan ja ajan tasalla pitämisen käyttötapaukset
 - Miten organisaatiossasi löydetään ja hankitaan uusia paikkatietoaineistoja?

Kuka tunnistaa tarpeet, miten hankintapäätös tehdään ja millä perusteella, miten hankinta konkreettisesti tehtiin jossain konkreettisessa tapauksessa? Mitkä tiedot ovat tärkeitä hankintapäätöksenteossa?
 - Kuinka suuri osa organisaatiossasi käytössä olevista aineistoista on käyttökelpoisia useissa eri hankkeissa ja/tai projekteissa, tai toisaalta vain jossain tietyssä projektissa?
 - Miten paikkatietoaineistojen laatua arvioidaan ennen hankintaa ja hankinnan jälkeen?

Miten ja mihin laatuarvioita laaditaan, tallennetaan ja käytetään organisaation sisällä?
Millaisiin toimenpiteisiin havaitut laatupoikkeamat johtavat?

- Miten hankittujen aineistojen ajantasaisuutta arvioidaan, ja miten tehdään päätös päivitetyn aineiston hankinnasta?

Kuvaile aineiston päivittämiseen liittyviä tehtäviä ja päätöksentekoa jossain tietyssä konkreettisessa tapauksessa.

- Miten paikkatietoaineistoja säilytetään ja jaellaan organisaatiossasi?

Onko teillä käytössä rajapintapalveluja aineistojen katseluun ja lataamiseen?

Ovatko nämä palvelut sisäisiä vai Internetin yli käytettäviä?

Mistä löydät tietoa saatavilla / käytettävissä olevista, tiettyyn käyttöön soveltuvista aineistoista ja miten jaat tätä tietoa kollegojesi kanssa?

- Kattavien paikkatietoaineistojen hyödyllisyys ja tiedon jalostus
 - Kun ajattelet organisaatiossasi tällä hetkellä käytössä olevia yksittäisiä paikkatietoaineistoja, kuinka suuri osa kattaa jo hankittaessa koko tarvittavan alueen ja/tai aikavälin?

Jos aineistoja pitää yhdistellä kattavuussyistä, jotta ne olisivat käyttökelpoisia organisaatiosi tarpeisiin, miten yhdistely tehdään ja miten helppoa se on?
 - Kuvaile konkreettisesti jotakin tapausta, jossa oli tarpeen tehdä eri aineistojen yhdistelyä, yhtenäistämistä tai muuta esikäsittelyä.

Onko ko. käsitelty aineisto ollut käyttökelpoinen organisaatiossasi laajemminkin tämän jälkeen?
 - Tarjoaako organisaatiosi itse keräämiänne ja/tai jalostamianne paikkatietoaineistoja tuotteina ulkopuolisille avoimena tietona tai maksullisina tuotteina? Jos, niin millä perusteilla oman käsityksesi mukaan nämä ulkopuoliset tiedon käyttäjät arvioivat tarjoamianne aineistoja?

- Tunnetut ja käyttökelpoiset aineistojen luokittelutavat, asiasanastot ja avainsanat
 - Minkälaisia luokittelutapoja ja asiasanastoja tai avainsanoja itse käytät tai lähiorganisaatiosi käyttää paikkatietoaineistojen organisointiin ja relevanttien aineistojen etsintään?

Missä ja miten nämä asia- tai avainsanastot on kuvattu? Ovatko asiasanastot monikielisiä vai olemassa ainoastaan esimerkiksi suomeksi tai englanniksi?
 - Tukevatko käytössäsi olevat työvälineet avainsanojen käyttöä relevanttien aineistojen löytämiseksi?

- Tuleeko mieleesi ongelmia tai kehittämistarpeita paikkatietoaineistojen luokittelun tai asiasanastojen käytön suhteen? Hyviä tai huonoja esimerkkejä?
- Muita huomioita
 - Paikkatietoalusta-hankkeen eräs keskeinen tavoite on saada kansallisesti kattavia, ajantasaisia ja laatuvarmennettuja paikkatietoaineistoja julkisesti saataville yhtenäisessä muodossa ja yhtenäisten rajapintojen kautta.

Minkälaisia odotuksia tai mahdollisia epäilyksiä sinulla tai edustamallasi organisaatiolla on Paikkatietoalustan toteuttamisen suhteen?

- Kansallisesti merkittävien paikkatietoaineistojen metatiedon sisällön ja sen muotoilun nykyistä tarkemmilla vaatimuksilla voidaan helpottaa relevanttien aineistojen löytämistä ja niiden soveltuvuuden arviointia tiettyihin käyttötapauksiin. Metatietojen tarkentaminen vaatii kuitenkin tiedon tuottajilta entistä enemmän huomiota ja parempia työvälineitä.

Miten arvelet metatietojen tarkentumisen vaikuttavan omaan työhösi? Minkä metatietojen tarkentaminen ja ajantasaisena pitäminen olisi oman kokemuksesi perusteella kaikkein hyödyllisintä ja miksi?

5 ANALYYSI

5.1 PAIKKATIETOAINIESTOJEN TARVE JA KÄYTTÖ OMASSA ORGANISAATIOSSA

Haastatteluiden perusteella välttämättömiä tai erittäin hyödyllisiä aineistoja ovat peruskartta, ilmakuvat sekä ortokuvat. Sen lisäksi rakennus ja huoneistorekisterin aineisto nähdään niin ikään erittäin tarpeellisena. Uudenmaan ELY-keskus, Suomen Ympäristökeskus sekä Ympäristöministeriö kaikki toteavat maankäyttöön ja kaavoitukseen liittyvät aineisto välttämättöminä. Tilastokeskus mainitsee myös Maanmittauslaitoksen kuntarajat-aineisto välttämättömäksi.

Kysyttäessä aineistoista, joiden käyttö tehostaisi tai helpottaisi arkea, mikäli ne olisivat käytettävissä, todettiin Rambollilta, että erityisesti rajapinnoista saatavalle aineistolle on tarvetta. Sitowiseltä vastaavasti todetaan, että suunnitteluasteella olevien maankäytön vektoraineisto olisi erittäin hyvä. Tilastokeskuksen osalta mainittiin mobiilipaikannusdataan (mobiililaitteiden sijainnit), kiinteistöjen ominaisuustietoihin (ml. postinumerotieto) sekä äänestysalueisiin liittyvät aineistot. Ympäristöministeriön maankäytön osastolta todettiin asemakaavat tällaiseksi aineistoksi, sillä vain osa on saatavilla ja nekin, jotka ovat, niin ovat nyt hajallaan eri kunnissa.

Kysyttäessä aineistoista, joiden käyttö on tällä hetkellä hankalaa tai tehotonta, todetaan Suomen ympäristökeskukselta, että Maastotietokanta (MTK) ja kiinteistörajat aineistojen olevan liian pieninä ruutuina, jolloin se aiheuttaa melko paljon käsityötä. Rambollilta mainitaan, että aineiston tarjoajan päässä voi ilmetä ongelmia, joten vaihtoehtoinen tapa esimerkiksi tiedostojen lataus tulisi olla sekä mainitaan, että tietyissä tapauksissa rajoituksena on kuntien aineistopalvelut, joihin pitää kirjautua (käyttäjätunnistautuminen). Sitowisen osalta mainitaan myös kirjautumiseen liittyvästä hankaluudesta, mutta myös siitä, että kuntien datat ovat todella vaihtelevia, kuten että joskus aineistoa ei

paikkatietoaineistona, attribuutit puuttuvat tai metatiedot puuttuvat. Sitowiseltä mainitaan myös, että kiinteistörajat-aineiston käyttö on todella tarpeellista, mutta sen käyttö on hankalaa, koska se saadaan paloina. Suomen ympäristökeskuksen osalta todettiin ilmakuviin osalta se, että on hyvin hankalaa tai mahdotonta on saada selville miltä ajalta (päivämäärä) kuva on otettu. Tilastokeskukselta vastaavasti todetaan Kuntakartta-aineistoihin liittyvä käsittely hankalaksi, sillä aineistot saadaan sähköpostilla ja työ on manuaalista ja monivaiheista työtä. Ympäristöministeriöstä (maankäytön osasto) mainittiin, että käytännössä kaikki aineistot vaativat ihmistyötä. Eräs syy tähän mainittiin olevan juuri kaavoituksen vapaussa (tavassa tehdä) ja siitä, että ei ole yhtenäisiä sääntöjä, vaan lähes kaikkialla yksilöllistä toimintaa. Lisäksi Ympäristöministeriöstä mainittiin, että vapaamuotoinen tekstille ei ole mitään määrittäjä tai rakennetta mitä siinä tulisi kertoa ja johtavan tämän siihen, että koneellinen tulkinta mahdotonta.

Kuitenkin tarjolla oleva paikkatietoaineisto on merkittävästi tehostanut organisaatioiden toimintaa, sillä muun muassa Ramboll ja Sitowise toteavat, että mikäli nykyisiä aineistoja ei olisi saatavilla tai sen laatuun ei voisi luottaa riittävästi, tarkoittaisi tämä aineistojen pyytämistä jopa paikan päältä ja siltä osin paluuta 10 vuotta menneisyyteen ja, että muutos olisi merkittävä ja merkitsisi paluuta ikään kuin menneisyyteen ja huomattavaan määrään manuaalista työtä. Suomen ympäristökeskukselta todetaan, että käytännössä jäisi työt tekemättä. Vastaavasti Tilastokeskukselle tämä vaikuttaisi oleellisesti tuotetarjontaan ja tilastoinnin laatuun. Ympäristöministeriöstä todetaan, että käytännössä tilanne on parhaillaan tämän kaltainen - aineiston laatuun ei voi luottaa riittävästi. Erityisesti kaavojen osalta paikkatietoaineistoa ei ole saatavilla kattavasti.

Valtaosa haastatelluista organisaatioista käyttää paikkatietoaineistojen käsittelyyn, analysointiin ja visualisointiin Arcgis, QGIS sekä ESRI:n ohjelmistoja. Selainpohjaisista palveluista mainittiin käytettävän erityisesti seuraavia: Paikkatietoikkuna, Tiemappi, Liikenneviraston katselu ja latauspalvelu, kaupunkien ja kuntien omat karttapalvelut, Oskari sekä Liiteri.

5.2 AINEISTOJEN LÖYTÄMISEN, HANKINNAN JA AJAN TASALLA PITÄMISEN KÄYTTÖTAPAUKSET

Rambollilla ja Sitowisellä paikkatietoaineistojen hankinta perustuu pitkälti projektien tarpeisiin ja usein aineisto saadaan projektin tilaajalta. Sellaiset aineistot, jotka ovat yleishyödyllisiä haetaan avoimista lähteistä tai jaellaan muuten organisaation sisällä. Sitowiseltä mainitaan, että vähemmän käytetään paikkatietoalustaa, sillä on epäselvää mitä se sisältää ja se, että osaan aineistoon pääsystä liittyy käyttöoikeudet, joiden saaminen on epäselvää. Tällä tarkoitetaan sitä, että vaikka kyseessä on avoin data tulisi se tarkoittaa helposti saatavaa. Suomen ympäristökeskuksella siinä vaiheessa, kun tiedetään mitä aineistoa halutaan, niin Suomen sisäisistä aineistoista tiedetään jo melko hyvin ne aineistoa tarjoavat organisaatiot ja jopa ihmiset, joilta aineistoa voi hakea. Suomen ympäristökeskuksessa tarve uusille paikkatietoaineistolla nousee Syken keskuksista käsin, ELY-keskuksesta sekä käyttäjäkyselyiden perusteella nousevista tarpeista. Uudenmaan ELY-keskukselta kerrotaan, että aineistojen tarve liittyy tiettyihin selvityksiin ja tehtäviin, mutta että harvoin on tarvetta uusiin paikkatietoaineistojen hankintoihin. Tilastokeskukselta mainitaan, että tämä asia (miten hankitaan uusia paikkatietoaineistoja) on parhaillaan hieman auki ja muutama asiantuntija tutkii parhaillaan asiaa. Yleisesti ottaen kuitenkin tutkitaan mitä on olemassa, luetaan artikkeleita sekä kuullaan mitä muut käyttävät. Tärkeimpiä tietoja hankintapäätöksenteossa Tilastokeskuksella todettiin olevan aineiston soveltuvuus, aineiston laatu, jotta sitä voidaan pohjata tilastotuotantoon sekä se, että käytetään samaa aineistoa eri tilastotuotannoissa. Muidenkin haastateltujen organisaatioiden tarpeissa korostuu myös aineiston soveltuvuus ja laatu.

Siihen kysymykseen kuinka suuri osa organisaatiossa käytössä olevista aineistoista on käyttökelpoisia useissa eri hankkeissa tai projekteissa, kerrotaan, että Suomen ympäristökeskukselta, että iso osa aineistoista on käyttökelpoisia useissa eri tilanteissa organisaation sisällä. Uudenmaan ELY-keskukselta taas kerrotaan, että käytännössä tarvittava aineisto on aina hanke kohtaista, joka johtuu useimmiten siitä, että paikka (kohta) joihin hanke tai selvitys liittyy, on eri. Tilastokeskukselta todettiin, että lähes kaikki aineisto on uudelleenkäytettävissä ja Ympäristöministeriöltä, että perusaineistot kuten kaavatiedot hyödyllisiä laajalti sekä myös kiinteistöraajat.

Paikkatietoaineiston laadun arviointiin liittyvässä kysymyksessä Rambolilta todettiin, että tunnetusti laadukasta aineistoa on saatavilla tahoilta MML, Syke, Ympäristö. Lisäksi todettiin, ettei ole varsinaista systeemiä laadun tarkkailuun. Sitowisen osalta todettiin, että paikkatietoaineiston laadun näkee oikeastaan vasta sitä tutkimalla ja että voi olla ihan sattuman kauppaa, soveltuuko aineisto käytettäväksi. Erityisesti resursseja syö maankäytönsuunnitelmat, jotka voi olla tehty, vaikka CAP-pohjaisilla tuotteilla mutta eivät ole kunnan paikkatietodataa. Lisäksi Sitowiselta kerrotaan, että laserkeilausaineistossa ajantasaisuus vaihtelee ja hankaluutta aiheuttaa se, että aineisto täytyy ladata, jonka jälkeen vasta näkee metatiedot. Suomen ympäristökeskukselta todetaan myös, että arvioitaessa paikkatietoaineistojen laatua ennen hankintaa, olisi hyvä, jos metatiedot olisivat kunnolliset niin, että voisi perehtyä miten aineisto tehty. Paikkatietohakemiston metatietokuvaukset ovat Suomen ympäristökeskuksen mukaan aika suppeita. Selostusta siitä mitä jokin attribuuttikenttä tarkoittaa ei ole kuvattu. Suomen ympäristökeskukselta todetaan lisäksi, että mittakaava vaikuttaa, kun arvioidaan aineiston laatua. Tilastokeskukselta kerrottiin, että aina etukäteen arvioidaan aineiston käyttökelpoisuutta ja tähän liittyen esimerkkinä mainittiin digiroad tarjoamat aineistot. Ympäristöministeriön osalta todetaan, että paikkatietoaineistojen laadun arviointi liittyy enempi Suomen ympäristökeskuksen tekemiseen sekä siitä, että kaava-aineistossa hyvin vaihtelevaa ja työhön liittyy paljon sekavan aineiston tulkintaa. Lisäksi todetaan, että mikäli halutaan automatisoida prosesseja niin ilman säännönmukaista aineistoa se ei onnistu.

aineiston tuottaja

Hankittujen aineistojen ajantasaisuutta ja päivitystarvetta arvioitaessa Suomen ympäristökeskukselta kerrotaan, että heillä seurataan vuosikelloa, jonka mukaan koordinoidaan minkä organisaatioiden aineistot haetaan. Tiettyjen aineistojen osalta aineisto haetaan vain joka toinen vuosi. Organisaation sisällä on arvioitu, kuinka usein aineisto on tarve päivittää. Tämän arvion tekevät paikkatietoaineiston vastuhenkilöt. Uudenmaan ELY-keskukselta todettiin lyhyesti, että konsultit arvioivat. Tilastokeskukselta mainittiin, että kerran vuodessa päivitetään rakennus ja huoneistorekisterin aineistot ja tieaineisto päivitetään, jos osoitteet eivät enää kohdistu. Yleisesti ottaen Tilastokeskukselta todettiin, että aika vähän tehdään hankintoja vaan niitä käytetään latauspalveluiden tai rajapintojen yli. Arvioitaessa hankittujen aineistojen ajantasaisuutta, Ympäristöministeriöltä mainittiin, että Liiterin yleiskaava -palvelu päivittyy neljä kertaa vuodessa. Asemakaava -palvelun osalta ei ole varmaa tietoa miltä ajalta se on, sillä kerrotaan vain, että tieto tulee kuntien rajapinnoista, mutta ei sen tarkempaa. Sitowisen ja Rambollin osalta aineistot ladataan useimmiten uutena, jotta vältetään riski, että projektissa tai hankkeessa käytetty aineisto olisi vanhentunutta.

Kysyttäessä miten paikkatietoaineistoja säilytetään ja jaellaan organisaatiossa, kerrottiin Rambollilla, että aineiston säilytys on hyvin projektikohtaisia ja yleisperiaate on se, että vanhoja aineistoja ei käytetä. Kuitenkin todettiin, että ruutu- ja tieaineisto on yleisellä palvelimella saatavilla organisaatiossa. Aineiston rajapintapalveluissa todettiin olevan yksikkökohtaisia eroja, ja osa aineistosta tallennetaan ArcGIS Online alustalle ja on siten käytettävissä ESRI:n ohjelmistoilla. Sitowisella aineistot ovat usein yksittäisen henkilöiden

tiedossa ja että yleisesti käytettävät aineistot ovat verkkolevyllä organisaation sisällä jaettavissa. Suomen ympäristökeskuksella aineiston jakelu tapahtuu keskitetyn paikkatietoinfran ja tarkemmin sanottuna ArcGis:n kautta siten, että vuosikellon mukaan viedään keskitettyihin paikkatietokantoihin aineistot, joista käsin Syken eri keskuskeskukset voivat hakea aineistoa. Lisäksi on tehty myös käyttöliittymä, jonka avulla aineistoa voi hakea ja varsinainen aineisto katseltavissa käyttöliittymän kautta. Uudenmaan ELY-keskuksesta kerrotaan, että paikkatietoaineistoa säilytetään Livin palveluissa ja näiden käyttö tapahtuu internetin yli siten, että aineistoa voi sekä katsella että ladata. Tilastokeskus tarjoilee aineistoa sekä sisäiseen käyttöön että ulkoisesti. Sisäisesti aineistoa tarjoillaan SQL Serverin, OGC-rajapintojen (WFS, WMS) kautta. Vastaavasti ulkoisia tarjotaan OGC-rajapintojen kautta myös (WFS, WMS), verkkolevyjakeluna (valmiita tuotteita, kuten ruututietokanta) sekä Statfin tilastopalvelusta (ilman paikkatietomuotoa). Ympäristöministeriössä paikkatietoaineiston jakelu ja käyttö tapahtuu käytännössä Liiteri-palvelun kautta. Lisäksi mainittiin, että joskus maakuntakaavojen kanssa työskennellään sähköpostilla saatavien shape-tiedostojen kanssa ja, että rajapintoja ei hyödynnetä. Lisäksi Ympäristöministeriöstä todetaan, että luultavasti paikkatietohakemistoa ei ole käytetty ja, että osa aineistosta siellä on avointa ja osa taas ei ja joihinkin on vaan yhteystiedot. Mainitaan, että on vaikea saada käsitys minkälaisesta aineistosta kulloinkin on kyse tai maksaako se jos aineiston haluaa ja jos maksaa, niin mitä. Tähän liittyen mainittiin esimerkkeinä aineistot ”Korkeusmalli 10” sekä ”Pälkäneen WMS palvelu”. Lisäksi todettiin, että tulisi olla suoraan linkki itse aineistoon.

5.3 KATTAVIEN PAIKKATietoAINEISTOJEN HYÖDYLLISYYS JA TIEDON JALOSTUS

Kysyttäessä kattavatko organisaatioon hankitut paikkatietoaineistot koko tarvittavan alueen, aikavälin tai muun, todetaan Rambollilta että kiinteistörajojen osalta tilanne on hyvä, mutta tilanne riippuu hyvin paljon projektista ja esimerkiksi pohjoisessa yksityiskohtaisuuden taso heikkenee, joka tekee yhdistelystä ja jalostuksesta vaikeaa. Ramboll mainitsee myös joskus olevan tarvetta vanhoille peruskartoille tai ilmakuville, jotta voidaan verrata ja tutkia muun muassa ympäristön muuttumista. Tilastokeskukselle tarve kattavuudelle perustuu heidän toiminnan luonteeseen - tilastollisista tarpeista johtuen on välttämätöntä, että kattavuustarve käsittää aina koko Suomen. Ympäristöministeriön kerrotaan, että maankäytön-osaston kannalta oleellinen asemakaava-palvelu kattaa vain puolet Suomen kunnista.

Aineiston yhdistely on yleistä kaikissa haastatelluissa organisaatioissa. Tämän todettiin johtuvan pääasiassa kahdesta syystä. Käsitellään laajaa aluetta, jonka osalta aineisto ei kata koko aluetta vaan on joukko palasia. Toisaalta tyyppillinen syy yhdistelyle on yhdistellä aineistoa useasta eri lähteistä ja muodostaa näistä käsiteltävä kokonaisuus. Tässä tilanteessa on oleellista, että aineistot ovat toistensa kanssa yhteensopivia. Kummassakin tapauksessa yhdistely on manuaalista ja yksi paikkatietoaineiston käsittelyssä olevia eniten kuormittavia tehtäviä.

Rambollilta mainitaan, että pahimmillaan yksi rakennus saattaa olla neljän kartan keskellä, mutta yleisin (perus) esikäsittely toimenpide on aineiston geokoodausta, jossa muun muassa tehdään koordinaattimuunnoksia. Toisaalta todettiin myös, että rakennus ja huoneistorekisterin ja maastotietokannan eroavan, kun näitä yhdistellään. Rakennusluokat eivät ole samat kuten esimerkiksi sauna toisessa aineistossa tarkoittaa lomarakennukseen. Käytännössä tämä vaatii aineistoon tutustumista ja sen ymmärtämistä. Suomen ympäristökeskuksen osalta todettiin, että erityisesti Corine aineiston muodostuksessa tehdään yhdistelyä, mutta lisäksi todettiin, että aineistojen yhdistely osaltaan kuuluu Suomen ympäristökeskuksen tehtäviin.

Tilastokeskus mainitsee aineistojen yhdistelyyn liittyen, että rakennuksille johdettava aluetieto(tunnus) on haettava toisesta paikkatietoaineistosta, kun muodostetaan kuntien osa-alueita. Tähän liittyen mainittiin,

että nämä tiedot tulevat suoraan kunnista ja muutoksia näihin tietoihin tulee noin 20-25 kappaletta vuodessa. Sen lisäksi todettiin, että tiedostomuoto vaihtelee merkittävästi eikä välttämättä ole paikkatietoaineistoa vaan esimerkiksi pdf-tiedostoja.

Haastateltavilta kysyttiin myös, tarjoaako organisaationne itse keräämiänne ja/tai jalostamianne paikkatietoaineistoja tuotteina ulkopuolisille avoimena tietona tai maksullisina tuotteina ja jos tarjoaa, niin millä perusteilla ulkopuoliset tiedon käyttäjät voivat arvioida tarjoamianne aineistoja. Ramboll, Sitowise, Suomen ympäristökeskus sekä Tilastokeskus tarjoaa aineistoja ulkopuolisille. Rambollin osalta käytännössä tuotettu aineisto luovutetaan tilaajalle, sillä usein aineisto on projektikohtaista ja ei välttämättä voida näin ollen käyttää toisessa projektissa. Sitowise tarjoaa paikkatietoaineistotuotteita kuten Louhi-palvelu ja Aino-palvelu. Suomen ympäristökeskus tarjoaa osaa aineistoa avoimena www-sivuillaan sekä sen lisäksi tarjolla on myös Suomen ympäristökeskuksen metatieto-palvelu, josta on nähtävissä aineistojen metatietoprofiili aineistojen arvioimiseen. Uudenmaan ELY-keskus mainitsee harvoin tarjoavansa itse keräämiä ja/tai jalostamia paikkatietoaineistoja tuotteina ulkopuolisille avoimena tietona tai maksullisina tuotteina. Tilastokeskus taas tarjoaa paikkatietoaineistoja tuotteina ulkopuolisille sekä avoimena tietona että maksullisina tuotteina, mutta mainitsee painopisteen olevan nykyään avoimessa datassa. Lisäksi Tilastokeskukselta kerrotaan, että ulkopuolisilta tiedon käyttäjille lähetetään kyselyitä näihin liittyen aineistojen laatuun liittyen ja kerätään näin palautetta. Ympäristöministeriöstä mainitaan, että Ympäristöministeriä ei sinänsä tarjoa paikkatietoaineistoa.

5.4 TUNNETUT JA KÄYTTÖKELPOISET AINEISTOJEN LUOKITTELUTAVAT, ASIASANASTOT JA AVAINSANAT

Organisaatioilta tiedusteltiin myös luokittelutapojen, asiasanastojen ja avainsanojen käytöstä paikkatietoaineistojen ja relevanttien aineistojen etsintään liittyen. Ramboll hyödyntää yleisten inframallinnus vaatimusten (YIV) sisältävää yleisiä luokituksia sekä INSPIRE luokituksia jossain määrin online-rajapintapalveluissa. Myös Sitowise hyödyntää yleisten inframallinnus vaatimusten mukaista luokittelua, mutta toteaa sen lisäksi, että vaikka luokittelu on olemassa, niin nimeäminen ei aina kohtaa, sillä miten mikäkin on nimetty ja mitkä ovat käytännöt, on usein epäselvää. Luokitteluun liittyen Suomen ympäristökeskukselta osalta todettiin, että kaikille käyttöliittymän kautta tarjottaville aineistoille pitää olla INSPIRE luokittelu. Toisaalta todettiin myös, että osa aineistoista ei kuulu Inspire raportoinnin piiriin, jolloin näiden osalta luokittelun valinnassa täytyy käyttää hieman luovuutta.

Uudenmaan ELY-keskukselta mainittiin, että aineistojen löytämisen kannalta on oleellista, että aineistojen jaottelu ohjaisi käyttäjää löytämään muita vastaavia relevantteja aineistoja. Aineistoja jaottelu voisi pohjautua luokkiin kuten "liikenne", "maankäyttö", "johtotiedot, joka tarjoaisi kaikkea luokkaan liittyvää aineistoa. Uudenmaan ELY-keskukselta mainittiin myös, että lisäksi hyödynnetään paikkatietoikkunan luokittelua sekä Liikenneviraston latauspalvelussa olevaa luokittelua. Kaupunkien aineistojen osalta ELY-keskus mainitsee luokittelun olevan erilainen ja yleisesti ottaen sen, että luokittelu vaihtelee. Hakukoneissa sama aineisto voi olla eri luokissa ja on tapauksia, joissa on hyvin tulkinnanvaraista mihin luokkaan aineisto kuuluu. Lisäksi mainittiin, että luokittelussa hyödyllistä olisi aineiston laatuun liittyvä tieto: tarkkuus, päivitysvyvyys tieto, milloin päivitetty. Erikseen mainittiin, että hyvä olisi olla mahdollista hakea vain päivittyneet tiedot.

Tilastokeskuksella ylläpidetään kansallista luokittelua JHS-mukaisesti ja paikkatietoaineistot saavat luokituksensa tämän kautta. Aineistojen luokittelua hyödynnetään Ympäristöministeriön mukaan esimerkiksi

siten, että Asemakaavayhdistelmä-aineisto kuuluu paikkatietohakemistosta ”maankäyttö” luokkaan. Mainitaan, että luokittelu käytännössä perustuu paikkatietohakemiston luokitteluun.

Asiasanastoon liittyen Sitowiseltä todettiin, että riippuu pitkälti aineistosta, joiden osalta sanasto on merkittävää, mutta todettiin, että jonkinlainen vakiintunut sanasto olisi tarpeellinen. Suomen ympäristökeskukselta todettiin, että hyödylliset asiasanat on haettu Gemet asiasanastosta, mutta tämän suhteen on joskus samoja ongelmia kuin Inspire standardin osalta, sillä asiasanat eivät täysin sovellu joihinkin aineistoihin. Lisäksi Suomen ympäristökeskuksella on käytetty paikkatietohakemiston avainsanoja. Tällä hetkellä heillä on mietinnässä, että mitä sanastoja pitäisi käyttää ja pitäisivät hyödyllisenä sitä, että tällainen sanasto olisi olemassa muualla ja sen voisi integroida Suomen ympäristökeskuksen omiin järjestelmiin. Suomen ympäristökeskus mainitsee myös, että metatiedot ovat suomeksi ja osa niistä on käännetty myös englanniksi, mutta tällöin siitä on tuotettu oma metatietokuvauksensa. Ympäristöministeriöstä todetaan, että asiasanastot ovat maankäytön osalta levällään, mutta parhaillaan kuitenkin tehdään sanastojen vakiointia yhteen toimivuuden parantamiseksi. Lisäksi mainitaan esimerkkinä ajantasa-asemakaavan, kaavayhdistelmän sekä kaavaindeksin, joiden osalta puhutaan samasta asiasta, mutta osittain myös eri asiasta. Ympäristöministeriöstä mainitaan myös, että on aloitettu kaavatyyppien koodiluettelointi. Ongelman juurisyynä epäillään osittain olevan se, että kunnilla on vapaus tämän (kaavoituksen) suhteen ja se johtaa eroavaisuuksiin. Lopuksi Ympäristöministeriöstä kerrotaan, että tavoitteena olisi, että puhuttaisiin samoista asioista samoilla nimillä.

Niin Suomen ympäristökeskuksessa, Uudenmaan ELY-keskuksessa kuin Tilastokeskuksessa mainitaan ontologiaan liittyvä haaste. Haussa tällainen esiintyy muun muassa, kun haetaan avainsanalla ”joki”, niin odotettu hakutulos osaisi tuoda myös ”vesistöt” avainsanaan liittyvät aineistot. Vastaava esimerkki haasteesta on avainsanoille ”johto”, ”putki” ja ”viemäri”, jotka aihealueeltaan liittyvät toisiinsa.

Kysymykseen tukeeko käytössä olevat työvälineen avainsanojen käyttöä relevanttien aineistojen löytämiseksi, todetaan Rambollilta, että tukee niissä tapauksissa kun metatiedoista voi hakea asiasanoilla. Suomen ympäristökeskukselta todetaan, että geokäyttöliittymä tukee Inspire-luokittelua sekä aineiston otsikosta hakua, mutta ontologia -haaste on siinä, että haku avainsanoilla ”Rakennus” ja ”Rakennukset” tuottavat eri tulokset. Lisäksi metatietopalvelu osaa hakea myös kuvauksista avainsanoja. Tilastokeskuksella ei tällä hetkellä ole käytössä paikkaa paikkatietoaineistojen metatiedolle, jolloin sisäisesti tukea avainsanojen käytölle ei ole. Tavoitteena on kuitenkin jatkossa ohjata prosesseja metatietojen avulla. Ympäristöministeriöstä mainitaan, että käytännöt vaihtelevat kunnittain ja maakunnittain, ja aiheuttaa siltä osin omat haasteensa.

Haastatellut organisaatiot toivat esille myös ongelmia ja kehittämiskohteita paikkatietoaineistojen luokitteluun ja asiasanastojen käyttöön. Muun muassa Rambollilta todettiin, että jos ei satu tietämään aineiston avainsanoja, niin tieto avainsanalista olisi hyvä, sillä ongelmia ilmenee aineiston nimen tai avainsanan tietämättömyyden takia. Rambollilta mainitaan myös tiedon tuottajan avulla ja sen kautta saattaa löytyä tarvittava aineisto. Sitowiseltä mainittiin, että ongelmia ja kehittämistarpeita ”Tulee valtavasti”, aihe olisi heidän mukaansa kokonaan omansa. Lisäksi he pohtivat, että kansiorakenteesta tulisi oppia pois ja siirtyä metatietoprofiilin käyttöön, sillä vaikka hierarkinen rakenne voisi olla hyvä, niin se ei välttämättä tue aina periytymistä tai verkko-muotoista assosiaatorakennetta. Sen lisäksi Sitowisellä mainittiin, että jos nimeämisestä ei ole sovittu tarkemmin, niin sen johdosta aineistojen löytäminen voi olla vaikeaa. Suositeltiin, että avain- ja asiasanojen suhteen sovittaisiin käytännöistä, jotta se tukee parhaiten

löydettävyyttä. Suomen ympäristökeskus korostaa erityisesti ontologiaan liittyvää ongelmaa sekä sitä, että tulisi olla sovitut avainsanat tai sanastot.

5.5 MUUT HUOMIOT

Haastatelluilta organisaatiolta kysyttiin myös odotuksia ja epäilyksiä paikkatietoalustan toteuttamisen suhteen. Rambollilta mainittiin, että vaikka WMS, WFS rajapintojen toteutus olisi INSPIRE standardin mukaista, niin siitä huolimatta taso vaihtelee ja pohdittiin, että laadullinen arviointi lisäksi (tähtiluokitus) olisi askel parempaan suuntaan.

Sitowiseltä puolestaan todetaan paikkatietoaineistojen olevan Suomessa hajallaan ja toivoo paikkatietoalustan kokoavan nämä yhteen. Lisäksi on epäselvää mikä paikkatietoalustan aineiston laajuus kaiken kaikkiaan on. Ehdottiin, että tulisi olla selkeä sisältö mitä (ainakin) on ja että ei tulisi ahnehtia liikaa aineiston suhteen. Metatietojen osalta jäätiin epäilemään jaksetaanko niitä noudattaa. Toivottiin, että MML katsoo, että aineisto kattavaa ja laadukasta. Lisäksi Sitowiseltä toivottiin ainakin keskeisiä aineistoja kattavasti.

Suomen ympäristökeskukselta toivottiin, että alustan olisi hyvä tarjota jonkinlaista katselukäyttöliittymää sekä toivottiin saada rajapinnat yhteen paikkaan siten, että niitä voisi hyödyntää suoraan muualla. Uudenmaan ELY-keskukselta todettiin, että jos Ympäristö-puoli (osasto) saisi tätä kautta toimimaan kaavat (myös pienten kuntien), niin se luultavasti hyödyntäisi huomattavasti julkishallintoa.

Tilastokeskus huomautettiin, että mikropalveluita (kuten geokoodaus) ei ole järkevää tuottaa kaikkien itse, vaan esimerkiksi keskitetysti siten, että kyseinen mikropalvelu olisi yleisesti muidenkin käytössä. Esimerkkinä mikropalvelusta mainittiin osoitteiden käsittelypalvelu sekä osoitteen validius -tarkastelu, etäisyyslaskenta, ja muita erilaisia spatiaalisia analyysejä.

Ympäristöministeriön osalta todetaan, että verrattuna nykyiseen paikkatietohakemistoon, niin uudelta alustalta odotetaan pääsyä lähemmäs itse aineistoa, sillä vaikka aineisto voi olla avointa, niin se voi silti vaatia rekisteröinnin ja sähköpostin antamisen tai muuta. Tätä asiaa hän toivoi yksiselitteisemmäksi. Ympäristöministeriö näkee hankkeen tarpeelliseksi aineistojen löydettävyyden näkökulmasta ja mainitaan, että aineistot ovat edelleen vajavaisia ja, että vaikka perusmetatiedot olisikin, niin itse aineisto on hajanaista ja sen tulkinta hyvin manuaalista ja tapauskohtaista.

Haastateltavilta tiedusteltiin myös metatietojen ylläpitämisen ja sen tarkentumisen myötä mahdollisesti koituvaa entistä tarkempaa huomioon ottamista, siitä aiheutuvaa lisätyötä ja työvälinevaikutuksia.

Rambollilta todettiin, että tärkeä, että metatiedoissa on standardimuotoista koneluotettavaa tietoa, sillä metatietoprofiilin vapaamuotoiset tiedot eivät ole helposti hyödynnettävissä. Todettiin, että tulisi olla vähän soveltamisen varaa siihen, miten aineisto on tuotettu. Kenties tähtiluokitus olisi hyödyllinen, josta selviäisi miten laadulliset tiedot täyttyvät.

Suomen ympäristökeskukselta todettiin, että kun joka tapauksessa heidän täytyy tuottaa INSPIRE metatietokuvaus (myös teemajako paikkatietohakemistoon), niin metatietokuvauksen tulisi tukea tätä, jottei luoda jälleen uutta tai toisenlaista kuvausta. Inspiren todettiin Suomen ympäristökeskuksen osalta olevan hyvä, mutta kaipaamaan jäätien attribuuttien selitteitä.

Tilastokeskukselta todettiin, että asia riippuu siitä, miten tämä vaikuttaisi nykyhetkeen, sillä nykyäänkin metatietoa viedään käsin. Tilastokeskuksen toiveena on tilastometatietomallin (GSIM, jossa YK on vastuukehittäjä) tulisi ymmärtää paikkatiedon metatietoa ja päinvastoin paikkatiedon metatiedon tulee ymmärtää tilastometatietoa (OGC Statistical Domain Working Group). Tilastokeskukselta mainittiin, että INSPIRE aineistojen tuottaminen on tällä hetkellä erittäin vaikeaa, koska se on niin monimutkainen ja, että todennäköisesti ilman konsultteja aineiston tuottamisesta ei selvitä.

Ympäristöministeriöstä huomautettiin, että kun metatietoprofiilin täyttööä vaaditaan, on odotettavissa myös kysymyksiä siihen miten profiili tulisi täyttää ja tähän tulisi tarjota kenties tukea ja kykyä vastata kysymyksiin kuten ”mitä tietoa halutaan mihinkäkin metatietoprofiiliin tietoon?”. Liittyen juuri metatietoprofiiliin Ympäristöministeriöstä toivotaan ainakin seuraavia tietoja:

- Aineiston tuottaja
- Luontipäivämäärä (tämän haastateltava toteaa olevan monimutkainen kaavojen osalta sillä tulisiko päivämäärän olla se, jolloin kaava on hyväksytty, milloin muokattu vai mikä)
- Yhteystiedot
- Bounding box (alue, jota aineisto koskee)

6 JOHTOPÄÄTÖKSET

Metatietoprofiililla on paikkatietoaineistojen suhteen kaksi toisestaan poikkeavaa käyttötarkoitusta. Ensiksi metatietoprofiililla on tärkeässä roolissa etsimiseen ja löytämiseen liittyvissä käyttötapauksissa, jolloin metatietoprofiilissa arvioitavia tietoja arvioidaan objektiivisesti. Käyttäjän tulisi voida arvioida ”mitä on tarjolla?”. Toiseksi sen avulla tulisi voida arvioida löydetyn paikkatietoaineiston soveltuvuutta subjektiivisesti omaan käyttötarkoitukseen. Alla on kuvattu tarkemmin paikkatietoaineistoon ja sen metatietoprofiiliin liittyvät johtopäätökset näiden kahden käyttötarkoituksen kautta.

6.1 PAIKKATietoaineiston etsimiseen ja löytämiseen liittyvät johtopäätökset

Ohjelmistojen ja automaation asettamat vaatimukset metatietoprofiilille

Haastatteluista nousi esille vahva digitalisoitumisen myötä nousseet uudet vaatimukset niin paikkatietoaineistolle kuin sen metatietoprofiilille, sillä paikkatietoaineistojen käsittely tapahtuu käytännössä automaattisesti, koneavusteisesti sekä ohjelmistojen avulla. Yhä useammin aineistoa käytetään rajapintojen avulla ja esimerkiksi rajapinnat nähtiin erittäin hyödyllisenä. Vastaavasti taas sellainen aineisto, jota ei voida varsinaiseksi paikkatietoaineistoksi kutsua (esim. ei asetu koordinaatistoon) tai käsitellä yhdessä toisen paikkatietoaineiston takia, koettiin hyvin hankalaksi käsitellä. Aineistojen yhdistelyyn ja käsittelyyn liittyy myös kaikilla haastatteluihin osallistuneilla organisaatioilla manuaalista työtä, jonka automatisoimisessa joko osittain tai kokonaan olisi huomattavaa hyötyä niin ajan säästön, kustannusten kuin lisäarvontuotannon näkökulmasta. Aineiston käsittelyyn ja analysointiin liittyy myös toimenpiteitä, joiden tekemisessä voisi hyödyntää tekoälyä.

Aineiston tuottajaan liittyvät vaatimukset metatietoprofiilille

Haastatelluilla organisaatioilla tuntui myös olevan melko hyvä tilanne tarvittavien paikkatietoaineistojen suhteen. Heillä on käsitys siitä mistä heidän toiminnassaan tarvittava aineisto löytyy ja harvoin tulee eteen tilanteita, että uutta aineistoa pitäisi alkaa etsiä. Haastateltavat myös tuntevat jo tiettyjä aineiston tuottajia,

joiden aineiston tiedetään olevan laadukasta, kattavaa ja standardimuotoista kuten esimerkiksi Museoviraston ja Suomen ympäristökeskuksen tuottama paikkatietoaineisto. Ajan myötä on siis selvinnyt ketkä tuottavat hyvää aineistoa perustuen siihen miten laadukasta ja standardimuotoista se on.

Saatavuuteen liittyvät vaatimukset metatietoprofiilille

Kuitenkin silloin kun tarve tulee löytää uutta paikkatietoaineisto, niin joko haastatelluilla itsellään tai heidän kollegoilla on usein tieto mistä aineistoa kannattaa alkaa etsiä. Paikkatietoalustan rooli tässä tilanteessa voisi olla hyvinkin oleellinen ja kasvaa keskeiseksi paikkatietoaineiston kanavaksi. Kaiken kaikkiaan paikkatietoalustan tarjoaman aineiston laajuutta ei selkeästi hahmotettu, joten esitystapoja mitä aineistoa alustasta ainakin löytyy ja miten aineistoon pääsee käsiksi, tulisi kehittää. Haastateltavilla tuntuu olevan hieman epäselvyyksiä siitä, mitä paikkatietoalusta sisältää ja mitä kautta saatavaan aineistoon pääsee käsiksi ja tarvitaanko aineistoon erillisiä käyttöoikeuksia. Kattavien kuvausten ja selkeiden esimerkkien avulla paikkatietoalustan tarjoamaa voisi tuoda paremmin esille. Vaikka kyseessä olisi niin sanotusti avoin data, ei varsinaiseen aineistoon kuitenkaan välttämättä ole suoraan avoimesti saatavilla. Aineiston saanti voi edellyttää rekisteröitymistä tai muita tunnuksia, joka hankaloittaa kuluttajaorganisaation päässä aineiston saatavuutta varsinkin silloin kun organisaatiossa ollaan tarkkoja käyttöoikeuksien suhteen.

Ontologiaan liittyvät vaatimukset metatietoprofiilille

Ontologiaan liittyvät haasteet nousivat esille useammassa haastattelussa organisaatiossa. Haasteet liittyvät aineiston löytämiseen tai löytämättä jäämiseen. Haettaessa avainsanalla ”joki”, odotetaan hakutuloksen pitävän sisällään myös ”vesistö” ja muut avainsanaan liittyvät aineistot. Vastaava esimerkki haasteesta on avainsanoille ”johto”, ”putki” ja ”viemäri”, jotka aihealueeltaan liittyvät toisiinsa. Ontologioiden hyödyntäminen mahdollistaisi esimerkiksi avainsanojen linkittämisen toisiinsa. Linkitysten avulla eri aineistojen metatietoja voitaisiin yhdistellä tiedonhaussa. Toisaalta ontologia voisi olla myös käsittehierarkiaan perustuvaa eli käsitteiden välisestä geneerisistä (laji-alalaji) ja partitiivisista (osa-kokonaisuus) tai assosiativisuuden suhteita ilmentävä. Avainsanoilla ”Rakennus” ja ”Rakennukset” tehty haku ei tulisi tuottaa eri tuloksia. Käyttäjää tulisi myös ohjata avainsanojen käytössä, sillä haastatteluiden perusteella tuli esille, että välillä voi olla vaikea tietää millä avainsanoilla ylipäätään on järkevä etsiä aineistoa. Ontologia-haastetta varmasti voidaan osaltaan ratkaista metatietoprofiilin avulla, mutta ontologia-logiikka voidaan myös erottaa omaksi älykkäämmäksi osaksi. Paljon on myös kiinni siitä, miten aineiston metatiedot ovat tuottajan tai muun tahon toimesta tuotettu. Kuten sanonta kuuluu - tyhjästä on paha nyhjäistä.

6.2 PAIKKATietoaineiston soveltuvuuteen liittyvät johtopäätökset

Kun puhutaan aineiston soveltuvuudesta, on tätä ennen käyttäjäpolussa pitänyt tapahtua aineiston löytyminen, jonka jälkeen käyttäjä pohtii, soveltuuko aineisto tarpeisiin.

Yksityiskohtaisuuden tasoon liittyvät vaatimukset metatietoprofiilille

Paikkatietoaineiston metatietojen tulisi kertoa siis objektiivisesti aineistosta. Täysin objektiiviseen näkemykseen voi kuitenkin olla vaikea päästä, sillä metatiedoissa voidaan kuvata vapaasti sanallisesti tiettyjä tietoja. Toisaalta juuri tietokoneohjelmistoilla tapahtuvan aineiston käsittelyn myötä on noussut tarve tunnistaa, miten aineisto soveltuu millekin ohjelmistolle. Lisäksi kun paikkatietoaineistoissa ollaan siirrytty vektorimuotoiseen esitystapaan, yksityiskohtaisuuden tasolla on vaikutusta siihen, miten hyvin aineisto soveltuu tarpeisiin. Erityisesti yksityiskohtaisuudella (tai mittakaavalla) on vaikutusta silloin kun erilaisia aineistoja tulisi yhdistellä.

Aineistojen yhdistelyyn liittyvät vaatimukset metatietoprofiilille

Aineistojen yhdistelyssä oleellista on siis niiden yhteensopivuus keskenään. Tätä yhteensopivuuden tutkimista voisi kenties tehdä metatietovertailulla tai muulla koneellisella vertailulla. Yhdistelyssä on myös oleellista tietää mitä aineiston sisältämät tiedot tarkoittavat. Tähän liittyen mainittiin esimerkki rakennus ja huoneistorekisterin ja maastotietokannan eroista rakennusluokkien suhteen, jossa esimerkiksi 'sauna' toisessa aineistossa tarkoittaa 'lomarakennusta'. Jotta paikkatietoaineistosta johdettavat analyysit olisivat yhteismitallisia, edellyttää se niiden käyttäjiltä aineiston ymmärtämistä. Tätä voisi helpottaa selkeillä ja kenties yhteiskäyttöisillä koodistoviittauksilla.

Asiasanastoihin ja koodistoihin liittyvät vaatimukset metatietoprofiilille

Niin koodistojen kuin metatietoprofiilin hyödyntämisen suhteen toivotaan yhtenäisiä käytäntöjä. Haastateltavat kertoivat kuitenkin, että tämä voi ja onkin osoittautua hankalaksi sillä niin koodistot, luokittelutavat, asiasanastot kuin metatietoprofiili on vaikea luoda sellaiseksi, että soveltuisi kaiken tyyppisen paikkatietoaineiston käyttöön. Paikkatietoaineistot luonnollisesti vaihtelevat ja joillekin aineistolle tietyt metatietoprofiilin osat eivät yksinkertaisesti ole relevantteja. Sama pätee asiasanastoihin, kuten Gemet:n asiasanastot, jossa asiasanat eivät välttämättä täysin sovellu joihinkin aineistoihin.

Aineistoesimerkkeihin ja sen katseluun liittyvät vaatimukset metatietoprofiilille

Haastatteluista nousi esille se, että aineiston soveltuvuus voidaan usein varmistaa vasta jälkikäteen ja kerrottiin, jopa se, että soveltuvuus voi olla sattumaa. Soveltuvuuden arviointi varsinainen aineisto lataamalla ja tutkimalla on kuitenkin aikaa vievää, joten mahdollisimman aikaisessa vaiheessa tapahtuva soveltuvuuden arviointi olisi hyödyllistä. Arviointi voisi perustua niin metatietoihin, aineistoesimerkkiin tai jopa varsinaiseen aineistoon katseluun kenties selainkäyttöliittymän avulla.

Tuotantohistoriaan liittyvät vaatimukset metatietoprofiilille

Paikkatietoaineiston soveltuvuudessa laatu ja aineiston ajantasaisuus ovat oleellista ja käyttäjät kiinnittävät huomioita aineiston homogeenisuuteen. Epähomogeeninen aineisto on koostettu eri paloista, eri alueilta, eri ajoilta ja niin edelleen. Suomen ympäristökeskukselta todettiin, että ennen hankintaa tapahtuvan paikkatietoaineistojen laadun arvioinnissa tulisi voida perehtyä siihen, miten aineisto tehty. Käytännössä tällä tarkoitetaan aineiston tuotantohistoriaa, jota INSPIRE-standardissa kuvataan kohdassa "lineage". Ahonen-Rainion mukaan tämän kentän voisi jakaminen pienempiin strukturoituihin osiin voisi olla hyödyllistä.

Aineiston lähteisiin ja ajantasaisuuteen liittyvät vaatimukset metatietoprofiilille

Lisäksi paikkatietoaineiston metatiedoista tulisi ilmetä aineiston lähteenä käytetty materiaali sekä aineiston ajantasaisuus tieto, eikä sitä tule sisällyttää aineiston selosteeseen (kuten aineiston "Asemakaavayhdistelmä kuntien rajapinnoista"-aineistossa tällä hetkellä), josta se ei erotu.

LÄHTEET

Ahonen-Rainio, Paula. 2005. Visualization of Geospatial Metadata for Selecting Geographic Datasets. Helsinki University of Technology

JHS-suositukset. JHS 158. <http://www.jhs-suositukset.fi/suomi/jhs158>

Maanmittauslaitos. Paikkatietoalusta-hanke. <http://www.paikkatietoalusta.fi>

Maanmittauslaitos. Paikkatietoalusta Beta-versio. <https://beta.paikkatietoalusta.fi>

Liikennevirasto. NAP-liikkumispalvelukatalogi. <https://finap.fi>

Suomen Ympäristökeskus. Metatietopalvelu.
<http://metatieto.ymparisto.fi:8080/geoportal/catalog/main/home.page>