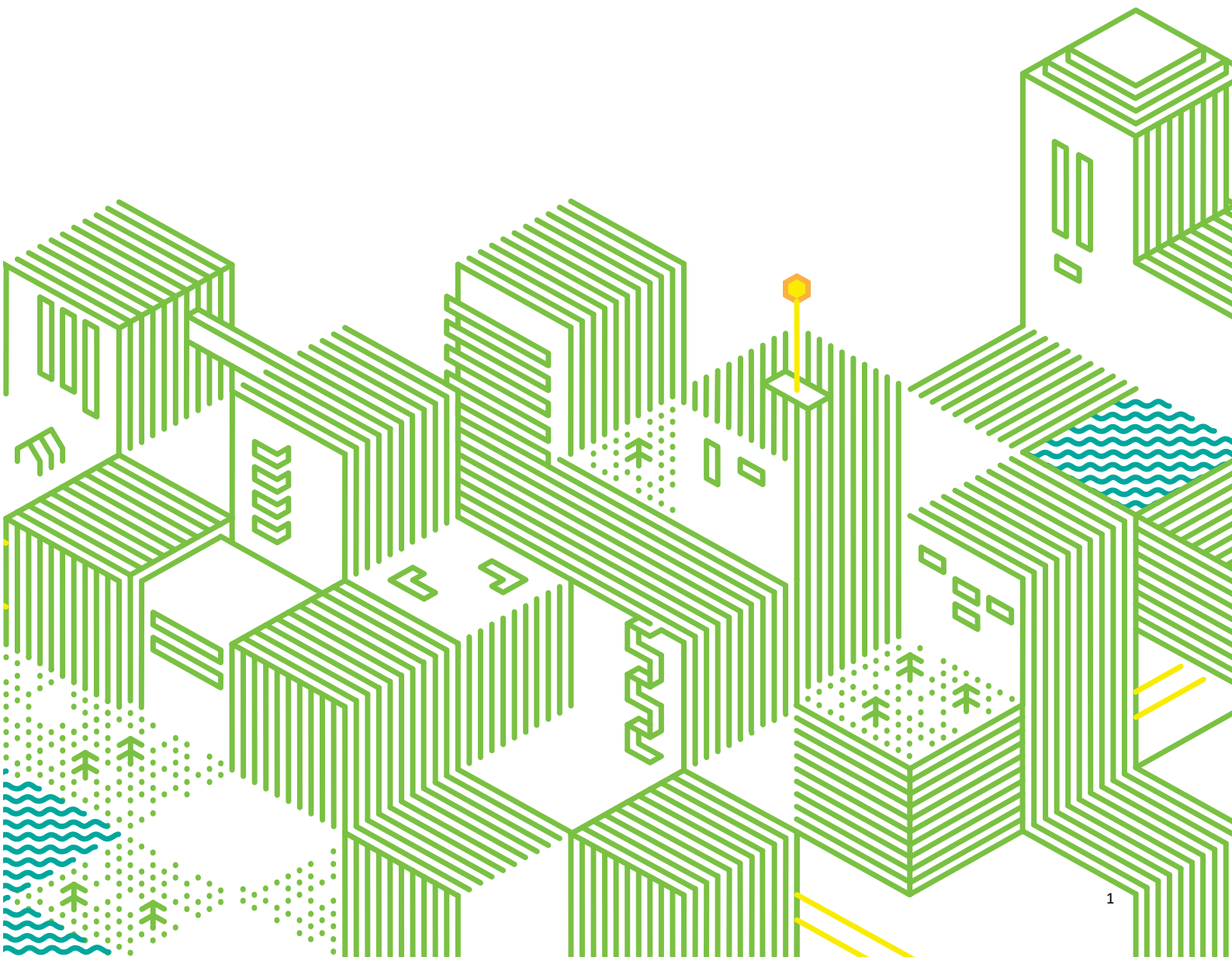


## INDEKSOINTIPALVELUN KÄYTTÖTARVESELVITYS

### Loppuraportti

Tiina Ketolainen, Niklas Lehmuskoski, Paula Salmela

27.3.2019



# Sisällys

1	INDEKSOINTIPALVELUN KÄYTTÖTARVESELVITYKSEN TAVOITTEET.....	3
2	SELVITYKSEN LÄHTÖKOHDAT .....	3
2.1	Indeksointipalvelu.....	3
2.2	Tilastokeskuksen sijaintitiedon viitearkkitehtuuryö.....	3
2.3	Haastatellut tahot .....	4
3	SELVITYKSEN TULOKSET.....	5
3.1	Palvelun tarve .....	5
3.2	Esiin nousseet minimivaatimukset.....	5
3.3	Tunnistetut käyttötapaukset .....	6
3.4	SWOT-analyysi .....	8
4	TULOKSET KÄYTTÄJÄRYHMITÄIN.....	9
4.1	Tilastokeskus .....	9
4.2	Maakuntien tilannekuva .....	10
4.3	YVA, ympäristölupa, YLVA-järjestelmä .....	11
4.4	SOTE-sektori.....	11
5	Liitteet.....	12

# 1 INDEKSOINTIPALVELUN KÄYTTÖTARVESELVITYKSEN TAVOITTEET

Indeksointipalvelun käyttötarveselvityksen tavoitteena on selvittää laajamittaisesti eri toimijoiden keskuudessa se, onko paikkaan sidottuja kohteita erilaisiin aluejakoihin sijoittavalle – indeksoivalle – palvelulle tarvetta. Indeksoinnilla tässä yhteydessä tarkoitetaan tiettyjen paikkatietokohteiden (esim. rakennukset tai osoitepisteet) sijoittamista ennalta määriteltyihin aluejakoihin, jotka voivat olla joko hallinnollisia (kunnanosa, kunta, maakunta), postinumeroalueita, tilastoruutuja jne. Tavoitteena on tuottaa luotettavaa tietoa eri aluejaoissa silloinkin, kun aluejaoissa tapahtuu muutoksia ripeässäkin tahdissa.

Selvitys on osa MML:n hallinnoimaa PTA-hanketta ja se kohdistuu maakuntien osahankkeen alaisuuteen. Selvityksen ei kuitenkaan ole tarkoitus ottaa kantaa siihen, kuka palvelun mahdollisesti toteuttaisi, eivätkä selvityksen tulokset sido MML:ää ottamaan vastuuta palvelun edelleen kehittämisestä ja tuottamisesta.

## 2 SELVITYKSEN LÄHTÖKOHDAT

Selvityksen taustalla vaikuttaa kansainvälinen GSGF -viitekehysmalli, jonka on YK:n EG ISGI (United Nations Expert Group On Integration of Statistics and Geospatial Information). Edelleen tästä dokumentista on tuotettu GSGF Europe Geostat 3 -projektissa. Kumpaakaan mallia ei ole vielä lopullisesti hyväksytty. Lisäksi Tilastokeskuksessa on käynnissä sijaintitiedon viitearkkitehtuuryö, joka tahollaan ratkoo samankaltaisia alueiden hierarkkiseen luokitteluun ja tilastotietojen paikkaan sitomiseen liittyviä kysymyksiä.

### 2.1 Indeksointipalvelu

Indeksointipalvelulla tarkoitetaan käytännössä sellaisia toiminnallisuuksia, joiden avulla olemassa olevat (tai muodostettavat) maastokohteet, esimerkiksi rakennukset tai osoitepisteet, liitetään osaksi jotakin olemassa olevaa tai erikseen luotavaa aluejakoa. Näiden annettujen aluejakojen avulla kohteet voidaan yhdistää muiden samaan aluejakoon kuuluvien kohteiden kanssa ja niitä voidaan tarkastella yhteisenä kohdejoukkona. Indeksointipalvelun toteuttamisen kannalta onkin keskeistä tunnistaa ne jo olemassa olevat aluejaot sekä mahdolliset tulevaisuudessa tuotettavat aluejaot, joita palvelun tulisi tukea. Aluejaot voivat olla joko hallinnollisia, ruutumaisia tai johonkin muuhun luokitteluun (esim. ihmisten käyttäytymiseen kunnan asiointiin) perustuvia. Indeksointipalvelun ensisijainen tarkoitus on paikantaa määritellyt kohteet tiettyihin aluejakoihin kuuluviksi objekteiksi, ei tuottaa muuta tilastollista tietoa näiden kohteiden osalta. Oletuksena on, että indeksointipalvelun käyttäjillä on pääsy kohteisiin sidottuun tilasto- tai muuhun tietoon saatavillaan jotain muuta kautta.

### 2.2 Tilastokeskuksen sijaintitiedon viitearkkitehtuuryö

Mahdollisesti toteutettava indeksointipalvelun näkökulmasta oleellista on huomioida se, että erilaisia aluejakoja on olemassa huomattava määrä. Olisikin tärkeää määrittellä yhdessä valtakunnallisesti, millaisia aluetasojakoja ja -luokitteluja käytetään. Jo vuonna 2014 tehdyssä JulkiCTlabin Paikkatiedon ekosysteemin kehittämisen -selvityksessä tunnistettiin huomattava määrä kuntatasoa tarkempia alueyksiköitä (esim. kuntien osa-alueet, taajama-alueet ja postinumeroalueet), joihin erilaisia tilastollisia muuttujia tulisi sijoittaa. Lisäksi Tilastokeskuksella olisi tilastoruutujen tarvetta myös koko valtakunnan kattavalle pienaluejaoille. Edelleen em. selvityksessä sekä Tilastokeskuksen kanssa käydyssä keskustelussa nousivat esille olemassa olevien tilastoruutujen yksiköt (250m x 250m, 1km x 1km sekä 5km x 5km). Tilastoinnin tarpeisiin 250m x 250m ruutukoko on yleensä riittävä, mutta ongelmaa aiheuttaa kohteiden epätarkka sijainti, jolloin tiedot saattavat sijoittua väärään ruutuun.

## 2.3 Haastatellut tahot

Indeksointipalvelun tarpeiden kartoittamista varten haastateltiin seuraavia tahoja neljässä ryhmähaastattelussa:

### **Tilastokeskus (13.2.2019):**

Rina Tammisto, IT-kehittämispäällikkö

Antti Santaharju, tekninen arkkitehti

Tapio Kytö, paikkatietoasiantuntija

### **Maakuntien liitot: Pirkanmaa, Uusimaa ja Varsinais-Suomi (15.2.2019):**

Henri Jutila, erityisasiantuntija, Uudenmaan liitto

Miina Vainio, paikkatietosuunnittelija, Pirkanmaan liitto

Laura Virtanen, suunnittelija, Varsinais-Suomen liitto

Lasse Nurmi, vanhempi suunnittelija, Varsinais-Suomen liitto

Antti Saarinen, suunnittelija, Varsinais-Suomen liitto

### **YVA, ympäristölupa, YLVA-järjestelmä (18.2.2019):**

Juha Lahtela, neuvotteleva virkamies, Ympäristöministeriö

Mikko Kantokari, kehittämispäällikkö, Uudenmaan ELY-keskus

Teija Tohmo, ylitarkastaja, Uudenmaan ELY-keskus

Eliina Kerko, ylitarkastaja, Uudenmaan ELY-keskus

Riitta Tornivaara-Ruikka, ylitarkastaja, Uudenmaan ELY-keskus

Annukka Engström, ylitarkastaja Uudenmaan ELY-keskus

Antti Alho, IT-erityisasiantuntija, KEHA-keskus

Ilona Joensuu, ympäristöylitarkastaja, Etelä-Suomen AVI

### **SOTE-sektori (19.2.2019):**

Ada Pentinmikko, suunnittelija, Vantaan kaupunki

Katja Turunen, erikoisasiantuntija, Valvira

Timo Hiekka, järjestelmäasiantuntija, Valvira

Harri Antikainen, tutkimustohtori, Oulun yliopisto, Impro-hanke

Jarmo Rusanen, professori, Oulun yliopisto, Impro-hanke

Mikko Alin, investointipäällikkö, Maakuntien tilakeskus

Jouni Pääkkönen, projektipäällikkö, Maakuntien tilakeskus

Heli Kuittinen, kehittämisasiantuntija, Maakuntien tilakeskus

Kimmo Parhiala, projektipäällikkö, THL

## 3 SELVITYKSEN TULOKSET

### 3.1 Palvelun tarve

Selvitykseen osallistuneiden käyttäjäryhmien haastatteluiden perusteella voidaan todeta, että indeksointipalvelulle on olemassa tarvetta, mikäli seuraavat reunaehdot toteutuvat:

- Käyttäjän on mahdollista tuoda palveluun omia aluemaisia kohteitaan, joiden perusteella tietoa (esim. rakennuksia tai osoitepisteitä) voidaan indeksoida tiettyyn aluejako-koon kuuluvaksi. Tällä tarkoitetaan, että käyttäjän on mahdollista tuoda indeksointi-palveluun oma shapefile-taso, mikä sisältää aluejaon (esim. vaikutusalueet) ja tämän jälkeen olemassa olevien rakennusten tms. ominaisuustiedot voidaan yhdistää alue-jakoon.
- Aluejakojen muokkaaminen siten, että salattujen tilastoruutujen -tai alueiden muodostuminen voitaisiin välttää (esim. eri kokoisten tilastoruutujen hyödyntäminen taajama- ja haja-asutusalueella). Tämä tarkoittaa käytännössä sitä, että jos käyttäjä haluaa käyttää esimerkiksi tilastoruutuja hyödyksi, voi hän valita palvelussa tilastoruudun koon siten, ettei ruutujen tietoja muutu salatuksi halutulla alueella. Palvelussa voi olla esimerkiksi ominaisuus, jossa palvelu valitsee automaattisesti pienim-män mahdollisen ruutukoon niin, ettei no data -ruutuja tule.

Indeksointipalvelun kehityksessä tulee huomioida erityisesti Tilastokeskuksessa käynnissä oleva sijaintitiedon viitearkkitehtuuryö, jossa pohditaan mm. surrogaattitietojen luomista, aluetunnusten johtamista sekä aluejaoissa tapahtuvien muutosten analyysiin liittyviä palveluita. Selvityksessä ei ilmennyt syitä, joiden vuoksi indeksointipalvelu tulisi nimenomaan toteuttaa osana MML:n PTA-hanketta vaan korostettiin sitä, että palvelun tulisi olla helposti integroitavissa jo olemassa oleviin järjestelmiin ja palveluihin.

Selvityksen yhteydessä ilmeni, että mahdollisten tulevien käyttäjämäärien arvioiminen indeksointipalvelun osalta on huomattavan vaikeaa, koska vielä ei ole selvillä se, tulisiko indeksointipalvelu olemaan oma itsenäinen palvelunsa vai integroituna johonkin muihin palveluihin. Kuitenkin oletettavaa on, että indeksointipalvelun avulla paikkaan sidottujen tietojen analysoiminen helpottuisi ja tietojen hyödyntäjinä voisivat toimia muutkin kuin paikkatietojen parissa pääasiallisesti työskentelevät henkilöt.

### 3.2 Esiin nousseet minimivaatimukset

Käyttötapausten perusteella on nostettu esiin indeksointipalvelulle tulevia minimivaatimuksia. Nämä minimivaatimukset ovat sellaisia, joiden tulisi palvelussa toteutua, jotta se olisi käytettävyydeltään hyvä ja siitä olisi käyttäjille hyötyä. Lähtökohtana indeksointipalvelulle on se, että **tiedon sijainti voitaisiin tallentaa ainoastaan kerran sen primäärilähteessä** ja muut järjestelmät voisivat tätä sijaintitietoa omissa eri aluetasojenkin tarkasteluissaan hyödyntää. Aluejakoihin sijoitetuilla kohteilla tulee siis olla ominaisuutena myös sijaintikoordinaatit, jotta alueellista tarkastelua voidaan suorittaa, millä tahansa tarkkuustasolla. Indeksointipalvelun toimivuus vaatii sen, että **sijaintitieto tallennetaan mahdollisimman tarkasti ja hierarkkisen rakenteen tarkimmalla tasolla**. Selvityksessä tuli ilmi, että osoite ja kiinteistötunnus eivät ole riittävän tarkkoja sijaintitietoja, vaan **indeksoitavan sijainnin tallennuksessa tulisi pyrkiä koordinaattitietoon**. Esimerkiksi ampumaradan oikea sijainti voi olla kaukana katuosoitteesta. Kuitenkin esim. rakennusten osalta olisi mahdollista indeksoida myös

rakennusten ääriiviivoja kuvaavia kohteita, mikäli pistemäinen koordinaattitieto olisi yhdistettävissä VRK:n ylläpitämään pysyvään rakennustunnukseen.

Selvityksen aikana nousi esiin **tarve muodostaa aluejakoja omien olemassa olevien aluerajausten** (esim. vanhusten kotihoito tai palveluverkkoon perustuvan reittianalyysin) **perusteella**. Tällainen toiminnallisuus tulee olla osa indeksointipalvelua, jotta siitä on käyttäjille hyötyä. Palvelussa olevat **aluejaot eivät voi olla staattisia**, vaan niitä tulee voida muuttaa käyttötapausten mukaan.

Palvelun toiminnan kannalta on keskeistä pistemäisten kohteiden (osoitteet, rakennukset, kiinteistöjen keskipiste) olemassa oleva tieto. **Lisäksi on huomioitava myös viivamaiset sekä aluemaaiset kohteet**, joiden indeksoinnin tulee myös olla mahdollista. **Tietojen luokittelun ja rakenteen yhdenmukaisuus** on kansallisesti tavoiteltava lopputulos, jota voidaan hyödyntää myös osana kansainvälistä yhteistyötä. Esimerkkinä yhtenäisestä luokittelusta ja sen suunnittelusta voidaan käyttää mm. JHS 206 Maakuntien palveluluokituksen mukaista jaottelua eri sosiaali- ja terveyspalveluiden luokittelusta.

### 3.3 Tunnistetut käyttötapaukset

#### *Kohteiden indeksointi omiin annettuihin aluejakoihin*

Maakuntien tilannekuvan osalta tarpeet käyttötapauksille tulevat maakuntien toiminnasta ja toiminta-arkkitehtuurista. Keskeisen tärkeä toiminto maakuntien toiminnassa – samoin kuin sosiaali- ja terveyspalveluiden suunnittelussakin – on **palveluverkon suunnittelu**, joka nivoo muuttoliikkeen, maankäytön suunnittelun sekä liikennesuunnittelun yhteen. Palveluverkon suunnittelun käyttöön indeksointipalvelun käyttötapauksena tuli haastattelussa esiin omien aluerajausten lisääminen palveluun. Esimerkiksi **.shp-muotoisten polygonien lataamisen** tulisi olla mahdollista, jotta suunnittelijoiden on mahdollista kiinnittää indeksoitavat kohteet juuri siihen tarkasteltavaan alueeseen, jolle käynnissä oleva palveluiden suunnittelu kohdistuu. Tuotetun indeksoidun aineiston avulla olisi mahdollista **tuottaa käyttäjän antamien parametrien mukaiset tiedot esim. saatavilla olevasta rakennuksiin sidotusta väestötiedosta**. Esimerkkinä tällaisesta käyttötapauksesta voidaan mainita kotihoidon alueen sisäinen väestönjakautuminen ja palveluiden suunnittelu.

Omiin aluejakoihin (esim. tietty terveydenhuollon perusyksikkö) perustuu myös SOTE-sektorilla monenlaista toimintaa. Päivittäisessä tilannekuvan arvioinnissa voidaan arvioida sitä, miten asiakaskäyntien määrä jakautuu eri toimipisteiden välillä ja onko tarvetta esimerkiksi siirtää henkilöstöä eri toimipisteiden välillä tai ohjata asiakkaita eri toimipisteisiin nopeamman asioimisen mahdollistamiseksi.

#### *Reittiaineistoihin perustuvien aluejakojen muodostaminen*

Indeksointipalvelun tulisi mahdollistaa aluerajauksen muodostaminen ”lennosta”, esimerkiksi tietyillä muuttujilla (esimerkiksi etäisyys ajallisesti tai matkallisesti terveysasemasta/palvelukeskuksen yksiköstä) **muodostettuun reittiaineistoon perustuen seuraavalla tavalla:**

- 1) Annetaan tarkasteltava alue
- 2) Haetaan alueen palvelupisteet ja asuinrakennukset
- 3) Haetaan rakennuksiin sidottu tieto ulkopuolisista tietovarannoista
- 4) Ja lasketaan jokaiselle asuinrakennukselle lyhyimmät etäisyydet palvelupisteeseen valitulla parametrilla, joka todennäköisesti indikoi matkaa tai

matka-aikaa, ehkä myös rakennusten asukasmääriä ja -jakaumia iän tms. perusteella

Kohta 4 edellyttää erillistä reititysovellusta ja -dataa, joka integroidaan indeksointipalvelua hyödyntävään palvelusovellukseen aluerajausten muodostamiseksi.

Toiminnallisuus mahdollistaisi uusien palvelualueiden muodostamisen ja erilaisilla parametreilla tuotettujen palvelualueiden tilastollisten muuttujien joustavan vertailemisen, vaikka varsinainen rakennuksiin sidottu tieto haettaisiinkin indeksointipalvelun ulkopuolisesti tietovarannosta. **Tästä toiminnallisuudesta olisi hyötyä erityisesti suunnittelussa, joka kohdistuu palvelu-/asiointialueiden reuna-alueille**, koska tällöin saataisiin todelliset liikkumis- ja asiointisuunnat paremmin näkyviksi. Lisäksi arjen palvelujen osalta liikennemallien luominen helpottuisi.

#### *Tietojen tarkastelu muuttuneiden aluejakojen osalta*

Jo nykyisellään maakuntien suunnitteluorganisaatioissa on käytössään erilaisia hallinnollisiin aluejakoihin ja tilastollisiin ruutujakoihin perustuvia aineistoja. Tarvetta olisi myös muuttuneiden aluejakojen huomioimisella historiallisia aineistoja tarkasteltaessa. Esiin nousi tarve sille, että muuttuneita tietoja voitaisiin tarkastella sekä nykyisessä aluejaossa että historiallisesti voimassa olleiden aluejakojen mukaisissa tarkasteluyksiköissä, jolloin käyttäjä voisi itse valita sen, indeksoitaisiinko tiedot olemassa oleviin aluejakoihin vai jo poistuneisiin aluejakoihin. Historiatiedot tulee muodostaa ulkopuolisen lähteen avulla. Historiatietojen yhdistäminen suoraan nykyiseen aluejakoon tekisi eri ajanjaksojen tilanteiden vertailun helpommaksi, eikä mahdolliset inhimilliset erehdykset tarkasteltavien historiallisten aluejakojen valinnassa vaikuttaisi vertailun lopputuloksiin.

#### *Mahdollisuus tilastoruutujen koon säätelyyn*

Lisäksi maakuntien tilannekuvaa selvittäneessä haastattelussa esille nousi erilaisten ruutukojen tarve, sillä erityisesti harvaan asutuilla seuduilla ja yleensäkin taajamien ulkopuolella tietojen vähäisyys aiheuttaa helposti sen, että käytettäessä 250m x 250m ruutuaineistoa olemassa olevat tiedot joudutaan salaamaan. Vastaavasti tiheään asutuilla kaupunkiseuduilla suuremman ruutukoon käyttäminen tekee käsiteltävästä tiedosta turhan karkeaa. Ns. salatujen ruutujen muodostumiseen voitaisiin vaikuttaa siten, että käytettävissä olevan ruudun kokoa voitaisiin säätää tarpeen mukaan. Näin käyttäjä voi säätää sitä, kuinka monia rakennuksia ruutuun osuu. Mikäli käyttäjällä on tiedossa rakennuksiin sidottu väestötieto, on ruudun kokoa mahdollista säätää myös väestömäärän mukaan.

#### *Vesialuejakoihin perustuvat aluejaot*

Ympäristölupahakemusten käsittelyssä ja valvonnassa indeksointipalvelun yhtenä käyttötapauksena nähtiin vesiasioihin liittyvät aluejaot. Vesiasioissa käytetään jo tällä hetkellä erilaisia osittain hierakkisia aluejakoja. Näitä ovat esimerkiksi vesipolitiikan puitedirektiivin mukaiset vesimuodostumat, päävesistöalueet, vesistön osa-alueet ja valuma-alueet. Viranomaiselle näistä tärkeimpiä ovat vesipolitiikan puitedirektiivin mukaiset vesimuodostumat. Lupa-  
viranomaisen näkökulmasta indeksointipalvelun yhtenä käyttötarpeena nähdään myös maankäyttömuotojen kuormitustietojen laskeminen ja tarkastelu.

Ympäristövaikutusten arvioinnin osalta aluejakoja ei käytetä, koska tarkastelut keskittyvät yleensä yhteen kohteeseen, joka voi olla pistemäisen lisäksi aluomainen (suuri teollisuuslaitos) tai viivamainen (liikenteen tai energiahuollon kohde). Tarkasteltava kohde voi sijoittua yhden tai useamman kunnan tai joissain tapauksissa maakunnan alueelle, joten aluejakojen suhteen on haastavaa määritellä tarpeellisia käyttötapauksia.

### 3.4 SWOT-analyysi

Työssä toteutettiin haastatteluiden yhteydessä SWOT-analyysi indeksointipalvelun vahvuuksista, mahdollisuuksista, haasteista ja epäilyksistä (kuva 1). SWOT-analyysin perusteella indeksointipalvelun vahvuuksina voidaan nähdä se, että kaikilla selvityksessä mukana olevilla sektoreilla on olemassa paljon (paikka)tietoa, jota voisi yhdistää aluejakoihin. Lisäksi **useilla sektoreilla on jo totuttu hyödyntämään ja yhdistelemään aluemaista tietoa, joten palvelu toisi tähän uuden työkalun.** Indeksointipalvelun haasteena voidaan nähdä eri toimijoiden erilaiset tarpeet aluejakojen suhteen. Tämän vuoksi edellä mainitun omien aluejakojen lisääminen palveluun tulisi olla mahdollista. Lisäksi aluejakojen tulisi olla päivittyviä, kuitenkin niin, että vanhat aluejaot säilyisivät historiatietona. Haasteena nähdään myös tiedon harmonisuus sekä tiedon luokitteluperusteet, joihin tiedon tuottajien tulisi kiinnittää huomiota jatkossakin.

Indeksointipalvelun tärkeimpinä mahdollisuuksina voidaan pitää kokonaisuuksien parempaa hahmottamista, yhteismitallisuuden ja helppokäyttöisyyden tuomia hyötyjä tutkimukselle sekä eri aineistojen parempaa yhdisteltävyyttä. Uhkina ja epäilyksinä indeksointipalvelulle nähdään tietosuoja-asiat, tiedon avoimuus, tekninen toimivuus, aluejakojen ajan tasalla pitäminen ja oikean vastuutahon löytäminen palvelulle. Indeksointipalvelu voidaan toisaalta nähdä keinona tietosuojan lisäämiselle, sillä aluejakojen avulla voidaan tarkkoja sijaintitietoja häivyttää. Palveluun tulee myös määritellä tarkasti käyttäjäryhmät ja käyttäjäryhmien saataville saatettavat tiedot. Palvelulle tulee määritellä ministeriötason vastuutaho, jotta palvelu ei siiloudu vain jonkin tietyn sektorin käyttöön. Vastuutahon tulee indeksointipalvelun hankinnan yhteydessä määritellä palvelutaso, jotta tekninen toimivuus sekä aluejakojen ylläpito varmistetaan.



### VAHVUUDET

Kaikilla sektoreilla, jotka mukana selvityksessä, on olemassa paljon tietoa, jota voisi yhdistää aluejakoihin.

Useilla sektoreilla on totuttu hyödyntämään ja yhdistelemään aluemaista tietoa eri lähteistä.

Maakuntien tilakeskus on tuottamassa yksityiskohtaista ja yhteismitallista tietoa hallinnassaan olevista tiloista, erityisesti SOTEn julkiset palvelut.

### MAHDOLLISUUDET

Syntyy uusia näkökulmia ja tarkastelutapoja.

Aineistojen yhteiskäytöllä saadaan parempia päätöksiä, yhteiskäyttöisyys lisää päätöksen avointa valmistelua, ainakin viranomaisten välillä.

Datan yhteistuotannon mahdollisuudet voivat kirkastua.

Tietojen yhdenmukaisuuden parantuminen helpottaa eri aineistojen yhdistelyä.

Paikallistuntemuksen vahvistuminen, kun eri lähteistä saatuja aineistoja helpompi yhdistää ja vertailla.

Kokonaisuuksien paremmat hahmottamismahdollisuudet

Myös asiakkaiden (toiminnanharjoittajat, kansalaiset) palvelu paranee.

Helppokäyttöisyys ja yhteismitallisuus voivat tuoda merkittäviä hyötyjä tutkimukselle.

### HAASTEET

Eri toimijoiden erilaiset tarpeet erilaisten aluejakojen suhteen

Olemassa olevan tai tuotettavan tiedon harmonisuus:

- eri päivityssyklit
- eri paikoissa olevat tiedot
- eri tasoiset tiedot (sisältö, tarkkuus)
- vanhat numeeristettavat tiedot

Tietosuojakysymykset ja käyttöoikeuksien hallinta

Tiedon luokitteluperusteet

Olemassa olevan tiedon laatu voi olla heikkoa ja puutteellista; sekä eri tavoin luokiteltua eri alueilla/aluetasoilla

### EPÄILYKSET

Mahdollistaako tietosuojalainsäädäntö aineistojen viemisen indeksointipalvelun ylipäättään?

Onnistutaanko muuttamaan myös toimintatapoja?

Onko kaikki tieto avoimesti käytettävissä?

Pystytäänkö käyttötarpeita ennakoimaan etukäteen?

Miten tekniseen toimivuuteen ja toteutukseen pystytään varautumaan?

Mikä on palvelun vastuutaho?

Kuka pitää aluejaot ajan tasalla?

Kuka vastaa palvelun ylläpidon rahoituksesta? → tarvitaan ministeriötason vastuutaho

Kuva 1. SWOT-analyysin yhteenveto.

## 4 TULOKSET KÄYTTÄJÄRYHMITTÄIN

Tässä kappaleessa on avattu tarkemmin kussakin haastattelussa esiin nousseita havaintoja, joita ei välttämättä ole kuvattu indeksointipalvelun kannalta keskeisinä huomioina. Ne ovat kuitenkin merkittäviä aineistoon ja aluejakoihin liittyviä tekijöitä kullakin yksittäisellä sektorilla.

### 4.1 Tilastokeskus

Tilastokeskuksessa on parhaillaan käynnissä sijaintitiedon viitearkkitehtuurityö, jonka yhteydessä muodostetaan sijaintitietovarannon käsitelmä sekä tavoitetilakuvaus sijaintitiedon käsittelyä varten tarjottavista palveluista. Sijaintitietovarannon käsitelmällin lähtökohtana on sijaintitiedon ja tilastotiedon yhdistäminen, jolloin lopputuloksena syntyy paikkatietoa.

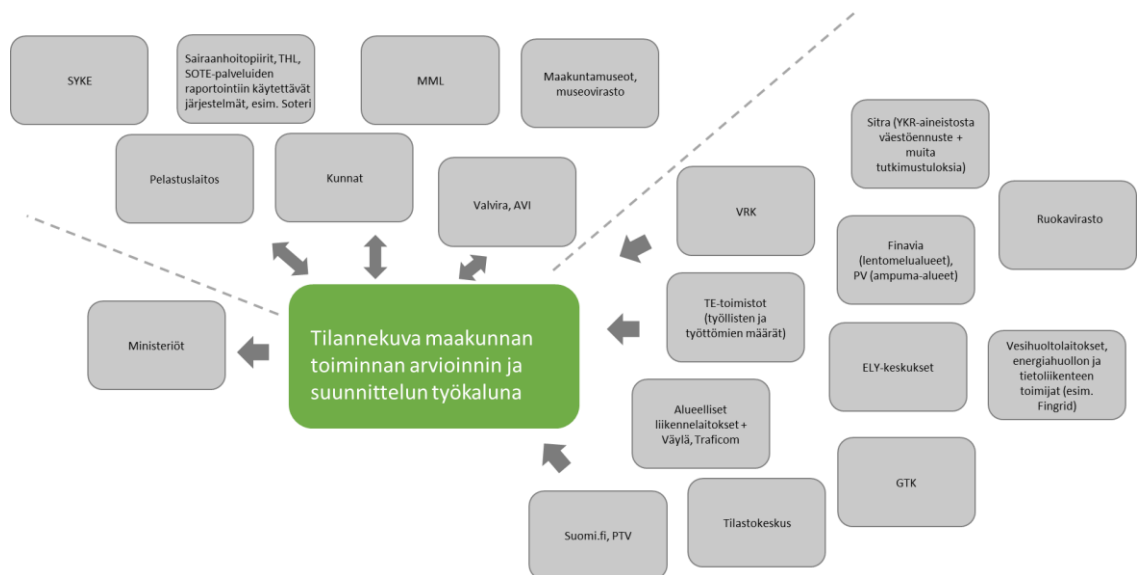
Tavoitetilakuvaus ei edellytä, että kaikki kuvatut palvelut olisi toteutettava Tilastokeskuksen omina palveluina vaan synergiaa muiden sijainti- ja aluetietoja tarjoavien palveluiden ja niiden tuottajien kanssa pyritään aktiivisesti etsimään. Indeksointipalvelun näkökulmasta keskeiseksi Tilastokeskuksen palvelutarjoaman osista on tekemässämme selvityksessä tunnistettu seuraavat palvelukomponentit: **aluerajojen muutosten analysointipalvelu** (osana sijaintitietojen tulosten tulkintaa ja raportointia), **surrogaattitunnusten luominen sijaintiin linkittyvälle tiedolle** (osana johdettujen tietojen muodostamisen palveluita), **aluetunnusten johtaminen** (osana johdettujen tietojen muodostamisen palveluita). Aluerajojen muutoksen

analysointipalvelun osana tulisi huomioida myös se, miten **tilastoidut aikasarjat sijoitetaan aluejakoihin**.

Tilastokeskuksessa on lisäksi tunnistettu paljon tarpeita, joita ratkotaan samanaikaisesti Maanmittauslaitoksen PTA-hankkeen yhteydessä. Näidenkään tarpeiden osalta Tilastokeskuksessa ei nähdä välttämättömäksi tarpeiden ratkomista omien ratkaisujen avulla vaan mielellään hyödynnetään PTA-hankkeen yhteydessä toteutettavia ratkaisuja, mikäli niiden toiminnallisuudet vastaavat Tilastokeskuksen tarpeisiin. Tällaisia tunnistettuja tarpeita ovat mm. sijaintitietojen laadun validointi, kohteiden (esim. rakennukset tai osoitteet) geokoodaaminen sekä aineistoille mahdollisesti tehtävät koordinaattimuutokset.

## 4.2 Maakuntien tilannekuva

Maakuntien suunnitteluorganisaatioissa käytössä olevat aineistolähteet ovat varsin moninaiset ja tietoa kootaan useiden eri toimijoiden tuottamista aineistoista. Lisäksi aineistoista koostettua tietoa palautetaan jonkin verran takaisin aineistotuottajien sekä ministeriöiden suuntaan. Kuvassa 1 on esitetty tärkeimmät aineistolähteet maakuntien suunnitteluorganisaatioiden näkökulmasta.



Kuva 1. Maakuntien suunnitteluorganisaatioiden käyttämien aineistojen tuottajat.

Nykyisten maakuntien yleisimmin käyttämä aluejako on ruutuaineisto, jonka haasteina nähdään tiedon salaaminen (se, että käyttäjä ei ole raakadatassa kiinni) sekä ruutujen rajautuminen erilaisia aluerajauksia käytettäessä. Indeksointipalvelu voisi tuoda tähän muutosta siinä tapauksessa, että hän voisi **itse asettaa haluamansa alueen rajat**, eikä olisi enää riippuvainen ruutujaon keinotekoisista suorakulmaisista rajauksista. Lisäksi erittäin hyödyllistä olisi se, että **käyttäjällä voisi rajata tarkemmin niitä kohteita, joiden osalta tietoa suodatettaisiin esim. väestötietojen osalta**. Näin ollen käyttäjällä olisi mahdollista vaikuttaa tietojen salaamiseen esimerkiksi muuttamalla rajaamaansa aluetta hieman, jolloin riittävästi kohteita osuisi rajatulle alueelle, eikä tietojen salaamista olisi tarpeen tehdä.

Omien aluejakojen lisääminen indeksointipalveluun on maakuntien tilannekuvan kannalta oleellinen toiminnallisuus, esimerkiksi erilaisten pienaluejakojen lisääminen palveluun nähtiin todella tarpeellisena toiminnallisuutena. Nykyisissä maakunnissa nähdään, että indeksointipalvelussa tulisi olla toiminnot sekä julkiseen käyttöön että viranomaiskäyttöön. On olemassa paljon sellaista aineistoa, jota ei kannata julkistaa ja joka on vain viranomaisten käytössä.

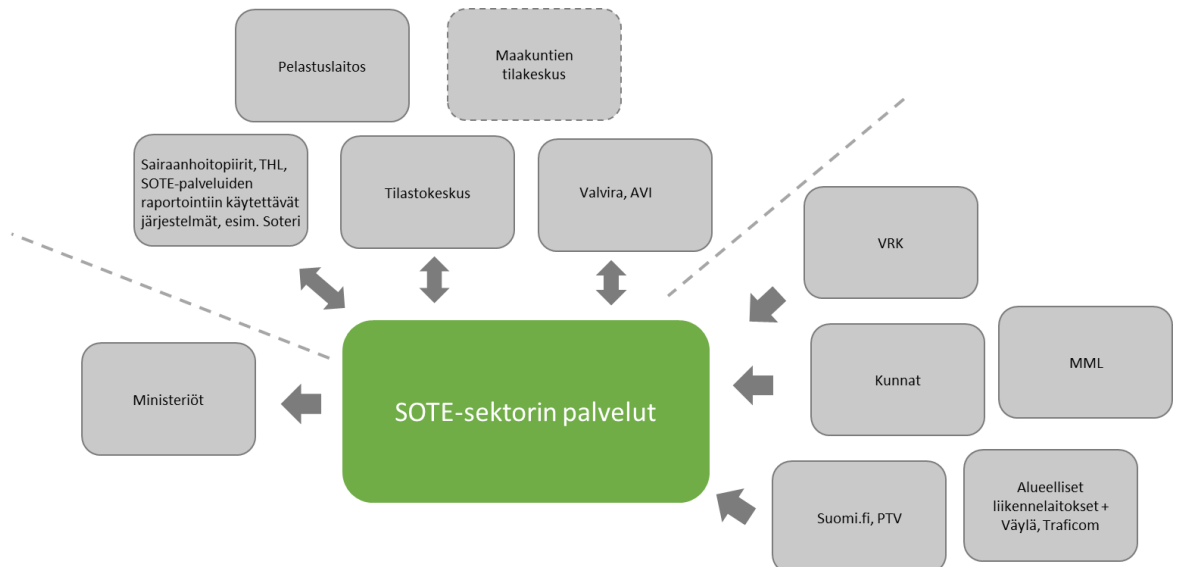
Nykyisissä maakunnissa paikkatietoaineistoja käyttävät pitkälti vain paikkatietoalan ammattilaiset, mutta indeksointipalvelun nähtiin mahdollistavan paikkatietojen laajemman käytön, jolloin käyttäjämäärä tulisi moninkertaistumaan, sillä analyysijä voisivat tehdä muutkin kuin ammattilaiset. Palvelun käytön tehostamiseksi sovellukseen tarvitaan TJS-integraatio käyttäjäorganisaatiolla olevien aineistojen liittämiseksi.

#### 4.3 YVA, ympäristölupa, YLVA-järjestelmä

YVA-kohteiden paikkatieto on olemassa vain kyseessä olevan kohteen alueen ELY-keskuskella. YVA-kohteissa paikkatieto on usein pistemäisen sijaan viivamaista tai jopa aluemaista. Erilaisten aluejakojen käyttö on ympäristövaikutusten arvioinnissa tällä hetkellä vähäistä. Ympäristöasioissa pistemäiset kohteet eivät ole riittäviä, koska kohteiden vaikutus leviää laajalle alueelle, joten kohteet ovat enemmänkin aluemaista. Monessa kohteessa osoite ei välttämättä kuvaa kohteen todellista sijaintia, sillä katusoitte voi olla kaukana todellisesta sijainnista tai osoitteen takana voi olla useita kohteita, jotka vaativat lupaa tai valvontaa. Näistä syistä sijainnin tallentaminen koordinaatteina on tärkeää. Haasteena indeksointipalvelulle pidetään sitä, että nykyisten ELY-keskusten rajat vaihtelevat E-, L- ja Y-vastuualueiden välillä, jolloin sama kohde voi sijoittua kahden ELY-keskuksen alueelle riippuen vastuualueesta. Ympäristöasioissa on paljon kohteita, joiden tarkka sijainti ei ole julkista tietoa (esimerkiksi pohjavedenotannot). Tämän vuoksi indeksointipalvelun käyttäjien rooleihin tulee kiinnittää erityistä huomiota. Indeksointipalvelun hyöty ympäristövaikutusten arvioinnin näkökulmasta voisi olla erityisen merkittävä silloin, kun arvioidaan erilaisten laitosten toiminnasta mahdollisesti aiheutuvien riskien vaikutusta alueen väestöön. Erilaisilla aluerajauksilla voitaisiin tarkastella sitä, kuinka suurta väestömäärää vaikutukset voisivat koskettaa.

#### 4.4 SOTE-sektori

Sote-sektorin haastattelussa kävi ilmi, että tällä hetkellä Suomessa ei ole luotettavaa (sijainniltaan ja ominaisuustiedoiltaan) paikkatietokantaa, jonka olemassaolo olisi tärkeää sote-sektorin kannalta. Jatkossa sote-sektorin paikkatiedot tulisi tallentaa luotettavasti tietokantaan(-kantoihin) tarkimmilla mahdollisella tasolla, mielellään koordinaatteihin sidottuna. Kun tieto olisi mahdollisimman tarkassa muodossa, olisi sen indeksointi eri aluejakoihin mahdollista ja hyödyllistä. Myös sote-sektorilla tiedontuottajat ja hyödyntäjät ovat moninainen joukko, jota on pyritty kuvaamaan kuvassa 2.



Kuva 2. Sote-sektorin aineiston tuottajat ja hyödyntäjät

Sote-sektorin kannalta tärkeimmäksi mahdollisen indeksointipalvelun aluejakoiksi nousi erilaiset palveluverkot, joiden mukaan tarkimmalla mahdollisella tasolla tallennetut rekisteritiedot pitäisi pystyä indeksoimaan. Palveluverkot ovat hyvin yksityiskohtaisia aluejakoja ja palveluverkkoja voi olla useita päällekkäisiäkin eri käyttötapauksista riippuen. Tämän vuoksi indeksointipalveluun tulisi pystyä tuomaan **omia palveluverkkoaluejakoja esimerkiksi shapefile-muodossa** tai indeksointipalvelussa voisi olla **jopa oma työkalunsa palveluverkkojen tuottamiseen**. Palveluverkkojen lisäksi **asiointialueet** koettiin tarpeellisiksi aineistoiksi indeksointipalveluun. Haastattelussa nousi esiin myös, että tulevan maakuntauudistuksen myötä kuntarajat hämärtyvät ja maakuntarajoista tulee uudet käytettävät hallinnolliset rajat.

Sote-sektorin haastattelussa nousi esiin myös historiatiedon tarpeellisuus. Esimerkiksi valvonnan näkökulmasta historiatietoa on seurattava viisi vuotta taaksepäin ja yliopistotasaisen tutkimuksen kannalta luotettavan historiatiedon saatavuus on tärkeää. Lisäksi käyttötapauksista haastattelussa nousi esiin tilastointi ja raportointi, jota maakunnissa on tehtävä kerran vuodessa. Käyttäjiksi nostettiin esiin kaikki suomalaiset, jos valinnanvapauslaki tulee joskus voimaan.

Tilastokeskuksen tuottama ruututietokanta nähtiin monessa tapauksessa erittäin hyödylliseksi "aluejakoiksi", sillä se mahdollistaa alueellisen tarkastelun perinteisistä aluejajoista (esim. kuntarajat tai kuntien kaupunginosarajat) riippumatta. Ongelmaksi nousee usein kuitenkin tilastoruutujen pieni koko, mikä aiheuttaa tietojen salausta, sillä jos tilastoruudun alueella asuu alle kymmenen henkilöä, ei tietoja voi näyttää. Tämä ongelma voisi olla ratkaistavissa palvelulla, jossa tilastoruutujen kokoa voisi muuttaa "lennosta" sopivaksi, niin että tarkasteltavalle alueelle ei muodostuisi no data -ruutuja.

## 5 Liitteet

### Liite 1. Haastattelu ja kysymykset.

Haastattelut toteutettiin puolistrukturoituna teemahaastatteluina asiantuntijaryhmille ja kysymysrunkoa käytettiin apuna yksittäisen teeman (tässä aineistokokonaisuus) monipuolisessa käsittelyssä. Haastatteluissa käytiin läpi seuraavat kysymykset vapaamuotoisesti keskustellen. Kysymykset on jaoteltu teemoittain (teemat tummennettu listassa):

#### **Niiden toimijoiden ja käyttötapauksen tunnistaminen, joilla on intressejä erilaisen paikkaan sidotun tiedon hyödyntämisessä ja tiedon skaalaamisessa erilaisiin tilastoyksiköihin (aluejakoisiin)**

- Mitä käyttötapauksia toimialallasi on, jossa tarvitaan erilaisia tilastointiyksiköitä? Esim. raportointi?
- Mitä (rekisteri)aineistoja tuotate, joiden indeksointi jatkokäyttöä varten olisi tarpeellista? Mitkä tahot ovat aineistojenne tärkeimmät hyödyntäjät?
- Mitä (rekisteri)aineistoja käytettä, joille indeksointia tulisi tehdä? Mitkä tahot ovat aineistojenne tärkeimmät lähteet?
- Keskeinen tietosisältö aineistojen osalta?
- Millaisia palveluita tuotate itse sisäiseen/ulkoiseen käyttöön?

#### **Olemassa olevan tiedon sijaintitarkkuus ja yksilöitävyys**

- Onko tiedolla täsmällinen sijaintitieto (koordinaatit yms., vähintään osoite)
- Voiko tieto olla ainoastaan aluejakoon sidottua?

- Millaista yksilöivää tietoa kohteilla on (osoite, kiinteistötunnus, pysyvä rakennustunnus jne.)?

#### **Palvelussa tarpeellisia (alue)jakoja, joille tietoa tuotetaan/ei tällä hetkellä tuoteta**

- Millaisia aluejakoja on käytössä? Onko tarpeet erilaisia eri aineistojen osalta?
- Onko tarpeen kyetä valitsemaan alueita, joissa hyödynnetään erilaisten aluejakojen yhdistelmiä? Esim. kuntaraja + postinumeroalue?
- Onko tarvetta määrittellä omia, erilaisia alueita? Esim. neliö, polygoni, säde, bufferi, karttanäkymä, zoomausikkuna

#### **Indeksointiin ja aineiston alueelliseen skaalautumiseen liittyviä haasteita ja kehitystarpeita**

- Tiedon elinkaari eri aluetasoilla: kuinka pitkälle historiaan on tarvetta nähdä tietojen jakautuminen alueellisesti? Onko tarpeen nähdä vanhan aluejaon mukainen tilanne vai voidaanko tiedot esittää nykyisen aluejaon mukaisesti?
- Aiheuttavatko pienet tilastoyksiköt ongelmia? Esim. tutkittavien kohteiden vähäisyys + salauksen tarve?
- Onko liian tarkan tiedon kohdalla salaukseen yms. liittyviä tarpeita? Esim. SOTE-puolella ja YVA-arvioinneissa mahdollisesti erilaisten tarkastusten tulosten sijoittaminen?
- Miten mahdollista salausta vaativien tietojen käyttöönottoa voitaisiin skaalautuvuutta muuttamalla helpottaa, jos mitenkään?
- Onko tunnistettu, että muilla käyttäjätahoilla on tarpeita samankaltaisten aluejakojen hyödyntämiseen?

#### **Mahdollisten käyttäjämäärien arviointi (karkealla tasolla)**

- Onko nimeämänne käyttötapaukset mahdollista järjestää yleisyyden perusteella?
- Erottuuko jokin käyttötapaus selkeästi muita tapauksia yleisemmäksi käyttäjämäärissä tai tietoon liittyviä kyselyjä mitattaessa?
- Voiko käyttötapauksia mitata esim. tunti/päivä/viikko/kuukausi/vuositasolla? Olisiko tietyn käyttötapauksen perusteella tarvetta indeksointipalvelulle vuosittain/kk/vko/päivätasolla?