

Metsätiedon palvelualueen konseptitestaus

Metsätehon tuloskalvosarja 7/2019

Jarmo Hämäläinen, Tapio Räsänen, Kirsi Rieki & Juha-Antti Sorsa, Metsäteho Oy

Risto Ritala, Tampereen yliopisto

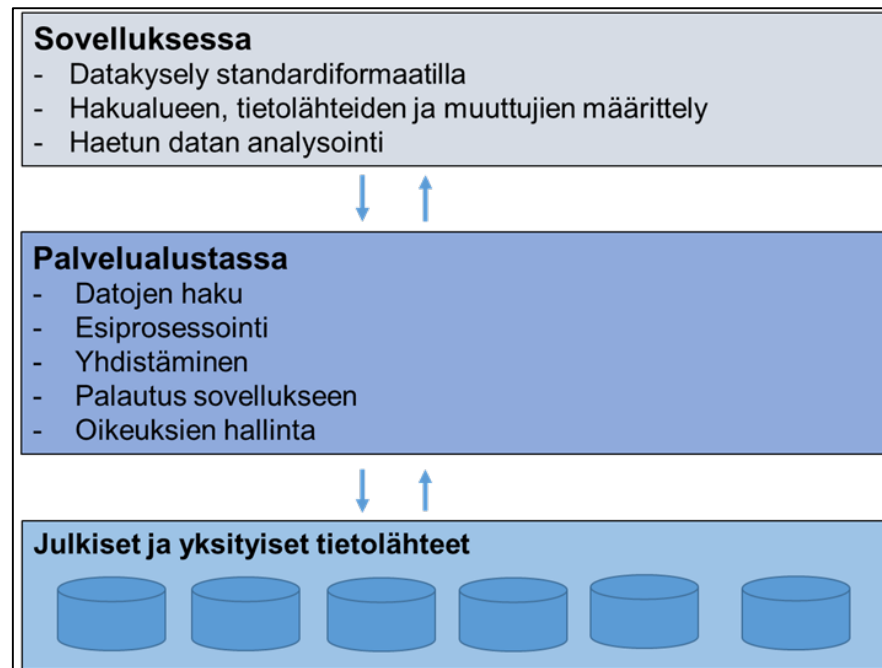
Tiivistelmä

- Hankkeessa rakennettiin ja testattiin metsätiedon palvelualustakonsepti, joka on tarkoitettu tehostamaan metsätiedon hallintaa ja hyödyntämistä.
- Testialusta rakennettiin Microsoft Azure -pilvialustan perustalle ja siihen yhdistettiin 14 eri tyyppistä tietoaainestoa. Datakyselyjä varten kehitettiin ForestJSON-kyselykieli.
- Konsepti osoittautui toimivaksi. Se tukee erityisesti data-analyysiin perustuvien kehittyneiden päätöstukisovellusten laadintaa, hilatasoisen datan käsittelyä ja sen ajantasaistusta sekä uusien julkisten ja yksityisten tietolähteiden käyttöönottoa.
- Alusta ehdotetaan tuotteistettavaksi Suomen metsäkeskuksen metsävaratietojärjestelmän yhteyteen, jolloin metsäkeskus olisi palvelun omistaja. Palvelualustan toiminnallisuus tehostaa ajantasaisen metsävaratiedon tuottamista, päivitystä ja jakelua.
- Tulevaisuudessa alusta voisi olla myös datojen, mallien, sovellusten ja palvelujen markkinapaikka.
- Eri tietolähteiden rajapintojen parantaminen on tärkeää. Lisäksi datojen luotettavuusestimaatit korostuvat jatkossa.
- Hilatiedon merkitys on kasvamassa ja sen tietosisällön parantaminen on olennaista. Tärkeitä kehittämiskohteita ovat etenkin puulajeittaiset puustotunnukset, maaperätiedot sekä ympäristöasioiden hallintaa tukevat tietokokonaisuudet.



Tarkoituksena tehostaa eri tietolähteiden käyttöä sovelluksissa.

- Metsätiedon palvelualusta tarjoaa palveluja datojen fuusiointiin, jakeluun ja käyttöoikeuksien hallintaan.
- Madaltaa kynnyistä kehittää uusia, metsätietoa entistä monipuolisemmin hyödyntäviä sovelluksia ja palveluita.
- Tavoitteena on sovellus- ja palvelukehityksen kustannustehokkuuden paraneminen sekä joustavuuden lisääminen, kun uusia tietotarpeita ja -lähteitä ilmaantuu.
- Erityisesti pyritään edistämään datalähtöisten päätöstukijärjestelmien kehittämistä eri toimijatasoille.



Palvelualustan periaate

Datakysely sovelluksessa

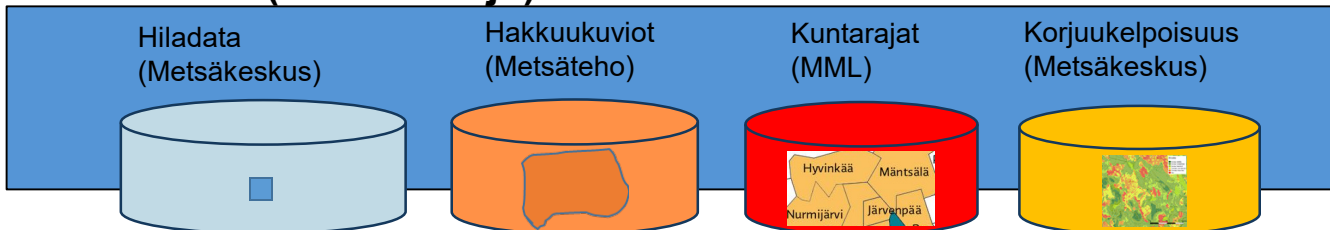
Datan palautus sovellukselle

Datojen haku
ja
yhdistäminen

GridId	Gridcell_volume	Hakkuutapa	KUNTA.NAMEFIN	KKL
685520393	256.11	6	Nurmijärvi	2
...

Palvelualustan
spatiaalinen käsittely-
yksikkö on hilapiste

Tietolähteet (esimerkkejä)

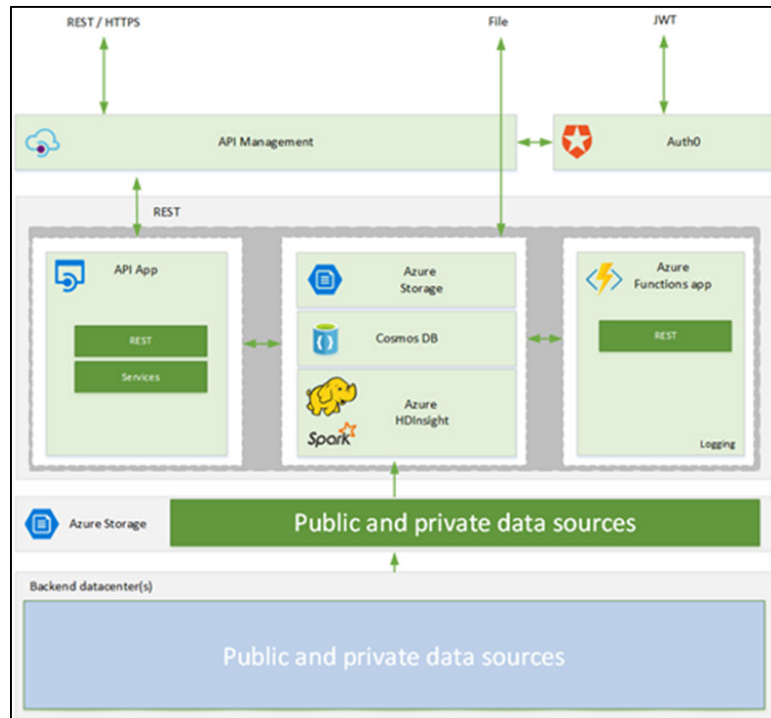


Kokeilujärjestelmän tekninen rakenne Microsoft Azure -alustalla

Käyttäjien
autentikointi

Työn prosessointi
laskentaklusterissa

Aineistoja
tietokannassa



Kuva: CGI Suomi Oy



Alustan testiaineistot

Aineisto	Tuottaja	Luonne
Hilamuotoinen metsävaratieto	Metsäkeskus	Julkinen
Hilamuotoinen metsävaratieto (MVMI)	Luke	Julkinen
Metsävarakuviot	Metsäkeskus	Julkinen
Kemera-aineistot	Metsäkeskus	Julkinen
Metsänkäyttöilmoitukset	Metsäkeskus	Julkinen
Korjuukelpoisuusluokitus	Arbonaut & Metsäkeskus	Julkinen
Maastotietokanta	Maanmittauslaitos	Julkinen
Suomen kuntajako 2017	Maanmittauslaitos	Julkinen
Säähavainnot (rajapintapalvelusta)	Ilmatieteen laitos	Julkinen (ilmainen palvelu)
Sääennusteet (rajapintapalvelusta)	Ilmatieteen laitos	Julkinen (ilmainen palvelu)
Metsäpaloindeksi (rajapintapalvelusta)	Ilmatieteen laitos	Julkinen (maksullinen palvelu)
Maanpinnan kosteusennuste (rajapintapalvelusta)	Ilmatieteen laitos	Julkinen (koeaineisto, maksullinen palvelu)
Hakkuukuviot (hakkuukonedata)	Metsäteho	Yksityinen (koeaineisto)
Muutostulkinta-alueet (Sentinel 2 -satelliittidata)	Terramonitor	Yksityinen (koeaineisto)



Aineistojen käsittely alustassa

- Suurin osa alustaan liitetyistä julkisista tietolähteistä ei ole nykyisellään saatavilla rajapinnan kautta, vaan erillisesti ladattavina tiedostopaketteina. Osa näistä aineistoista päivittyy vain harvoin.
- Suurten aineistojen lataaminen rajapinnan kautta datakyselyn yhteydessä heikentää olennaisesti alustan suorituskykyä ja vasteaikaa.
 - Suuret aineistot päädyttiin näistä syistä esiprosessoimaan valmiiksi palvelualustalle.
- Esiprosessoinnissa aineisto muutetaan valmiiksi alustan käyttämään hilamuotoon. Tällöin kyselyvastausten koostaminen on tehokasta.
- Tässä pilottihankkeessa rajapinnan kautta saatiin ainoastaan säähän liittyviä, dynaamisia tietolähteitä, joilla tiedonsiirtomäärät olivat pieniä. Rajapinnoista saatavien aineistojen tiedot muutetaan kyselyn yhteydessä hilalle.



ForestJSON-kyselykieli datojen hakuun

- Palvelualustan käyttö perustuu dataa käyttävään sovellukseen ohjelmitavaan kyselyyn, joka lähetetään alustalle.
- Projektissa määriteltiin ForestJSON-kyselykieli, joka pohjautuu paikkatietoalalla käytettävään GeoJSON-formaattiin.
- Kyselyssä määritellään datahaun sisältö:
 - Hakualue
 - Tietolähteet
 - Tietolähteiden yhdistämistapa
 - Tietolähteiden kentät (muuttujat)
 - Haun ehdot (suodatus)
 - Mahdolliset laskennalliset kentät
- Kyselyssä kutsutaan alustan laskentafunktioita:
 - Kasvunlaskenta
 - Tyhjien arvojen korvaaminen



ForestJSON Generator

A screenshot of the ForestJSON Generator web application. The interface is in Finnish and shows a 'Tietolähteen valinta' (Source selection) section with a dropdown menu set to 'Metsäkeskuksen hilasmuotoisen metsävarallisuuden aineisto'. Below this is a 'Lisävalinnat' (Additional selections) section with a text input field containing 'SMKGRID'. There are also sections for 'KORJUUKELPUOK_MAAKAS2019' and another 'Lisävalinnat' section with checkboxes for 'valitse kaikki', 'Gridi', 'GridX', 'GridY', and 'KML', where 'KML' is checked. A 'Poista' (Remove) button is at the bottom.

Valmis

UUID

54dd61b8-56ef-4ca0-85a6-84dc5aace7c5

Tuloslinkit

[result-20190522_085215-54dd61b8-56ef-4ca0-85a6-84dc5aace7c5-L414.csv](#)

[result-20190522_085215-54dd61b8-56ef-4ca0-85a6-84dc5aace7c5-L423.csv](#)

Kopioi kaikki linkit

Karttakäyttöliittymään ja alavetovalikoihin perustuva ForestJSON Generator laadittiin testauksen ja erillisten datakyselyjen avuksi sekä helpottamaan kyselykoodin kirjoittamista. Alustan perusajatuksena on, että datakyselyt ohjelmoidaan käyttösovelluksiin.



Alustan testitapauksia

Palvelualustaa kokeiltiin erityisesti seuraaviin käyttötarkoituksiin:

- Hilatietojen päivitys hakkuutiedoilla
- Hilan puustotietojen päivitys Motti-kasvumalleilla
- Kuviotietojen aggregointi hilatiedoista
- Hakkuumahdollisuuksien analysointi eri tietolähteistä

Lisäksi alustan toiminnallisuuksia testattiin laajasti eri tietoaineistojen avulla.



Käyttäjätestaukset (1/2)

- Testialusta ja hakkuumahdollisuussovellus olivat testattavana 20.12.2018–31.3.2019. Testaajina oli 12 henkilöä viidestä eri organisaatiosta.
- Käyttäjäpalautteen keskiarvo oli 3,5 asteikolla 1–5 (5 paras).
- Palvelualustalla nähtiin monenlaisia hyödyntämismahdollisuuksia organisaatioissa, mutta käyttöalueiden arviointi oli vielä alustavaa. Käyttömahdollisuuksina mainittiin mm. seuraavia:
 - leimikon tai korjuulohkon tarkennetun tiedon muodostus hakemalla hilatietoja ja yhdistämällä niitä muihin tietolähteisiin sekä tiedot ajantasaistamalla
 - organisaatioiden omien metsävara-aineistojen ajantasaisuuden hallinta ja vertailut
 - luonnonsuojelualueiden metsävaratietojen ajantasaistus
 - hakkuumahdollisuuksien haku eri kriteereillä
 - korjuuolosuhteiden reaaliaikainen ennustaminen
 - laajat tietohaut yleisesti erilaisilla suodatuksilla
 - tietojen yhdistäminen ja jalostaminen uudeksi korkeamman tason tiedoksi, esim. mahdollisimman tarkan tiedon tuottaminen dynaamisesti
 - integrointityön vähentäminen sovelluskehityksessä ja sovellusten ylläpidossa hankkimalla kaikki tarvittava data yhden rajapinnan kautta.



Käyttäjätestaukset (2/2)

- Tietolähteinä palvelualueella katsottiin tarvittavan lähtökohtaisesti niitä julkisin varoin tuotettavia paikkatietoaineistoja, joita voidaan käyttää metsävaratietojen tuottamiseen, ajantasaistukseen, hakkuumahdollisuuksien arviointiin (mm. metsän käytön rajoitteet ja kaavatiedot) sekä erilaisia olosuhteita kuvaavia aineistoja.
- Aineistojen saatavuus on tärkeää – palvelualueelta nähtiin kanavana aineistoihin, joihin yksittäisillä organisaatioilla ei ole pääsyä tai se on hankalaa.
- ForestJSON-kyselykieleen ei tarvittu laajennuksia tai uusia ominaisuuksia, ja dokumentaatio oli riittävä. Käytettävistä funktioista olisi tarvittu enemmän tietoa.
- Testauskäyttö oli lähinnä kyselyjen muodostamista käyttöliittymän avulla, jolloin mahdolliset sovelluskehittäjien erityistarpeet eivät liene tulleet testauksessa kattavasti esille.

Käyttäjätestauksesta saadut kehittämissuhteukset

- Kuviomuotoisen palautetiedon muodostaminen ja käsittely (toteutettiin hankkeen loppuvaiheessa)
- Vektorimuotoisten aineistojen käsittely alustassa (ei konvertointia hilamuotoon)
- Puustotietojen kasvunlaskennan nopeuttaminen tai aineistojen päivitys alustalla valmiiksi nykyhetkeen
- Runkolukusarjojen laskenta hilatietoihin mukaan
- Mahdollisuus hakea vain edellisen haun jälkeen muuttuneet tiedot (tietolähteiden tietosisältömuutosten hallinta)
- Parannuksia ForestJSON Generator -sovelluksen käyttöliittymään



Testialustan suorituskyky ja kustannukset eri laajuisilla datakyselyillä.

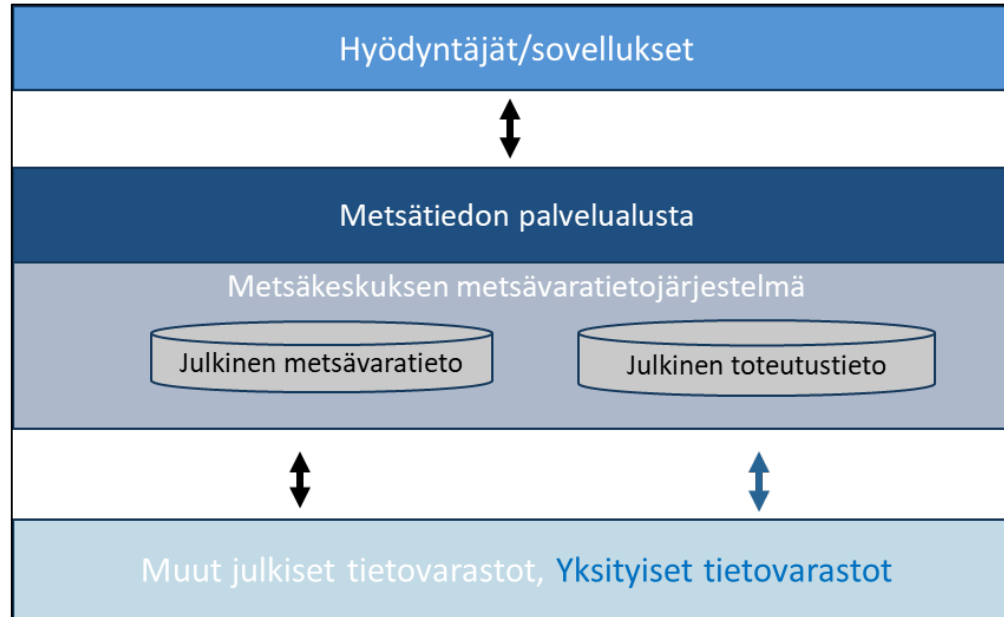
Alue	Pinta-ala, ha	Hiloja, kpl	Datat	Vastausaika	Kustannukset, €/kysely
"Työmaa"	4,3	167	hila, mvmi, kunta	1 min 30 sek	0,03
"Kunta"	93 400	3,6 milj.	hila, mvmi, kunta	6 min	0,12
"Maakunta"	542 850	21,2 milj.	hila, mvmi, kunta	17 min	0,36
"Kunta"	93 400	3,6 milj.	hila, mvmi, kunta, kklk, htapa	24 min	0,51
"Maakunta"	542 850	21,2 milj.	hila, mvmi, kunta, kklk, htapa	1 h 13 min	1,53
"Suuralue"	13,5 milj.	527 milj.	hila, mvmi, kunta	3 h 50 min	4,83

"Kunta" = Vihti + Nurmijärvi
"Maakunta" = Espoo, Hausjärvi, Vantaa, Hyvinkää, Järvenpää, Kerava, Kirkkonummi, Kärkölä, Mäntsälä, Nurmijärvi, Riihimäki, Sipoo, Tuusula, Vihti
"Suuralue" = Noin 1/3 Suomen pinta-alasta, Etelä-Suomi

Tuloksista saa suuntaa antavan käsityksen alustan suorituskyvystä. Testikyselyt on tehty siten, että vain yksi kysely on ollut kerrallaan prosessoitavana. Todellinen tuoteistetun palvelun tehokkuus riippuu pilvialustan kokoonpanosta ja yhtäaikaisten kyselyjen määrästä.



Ehdotus palvelualustan roolista seuraavan sukupolven metsätietoekosysteemissä.



Metsätiedon palvelualusta ehdotetaan liitettäväksi Metsäkeskuksen metsävaratietojärjestelmän yhteyteen. Ne muodostavat yhdessä "metsätietoalustan", joka varastoi, yhdistää ja välittää dataa eri tietolähteistä loppukäyttäjien sovelluksille.

Johtopäätökset

- Konseptitestausta osoitti, että palvelualustaa voidaan menestyksellä käyttää metsätietoaineistojen laajamittaiseen hakuun ja yhdistämiseen.
- Pilvialusta tarjoaa tehokkaan ja skaalautuvan teknisen ympäristön palvelulle.
- Metsäkeskuksen metsävaratietojärjestelmän yhteydessä alusta on omiaan tehostamaan ajantasaisen metsävaratiedon tuottamista, päivytystä ja jakelua. Palvelualustasta on hyötyä kaikille toimijoille, jotka hyödyntävät tai jatkojalostavat metsätietoa.
- Alustakonseptissa on metsäalan innovaatioita ja uutta liiketoimintaa vauhdittavaa potentiaalia. Tulevaisuudessa se voisi olla jakelualustaa laaja-alaisempi datojen, mallien, sovellusten ja palvelujen markkina-alue.
- Eri tietolähteiden rajapintojen toimivuuden parantaminen on tärkeää.
- Hilatiedon merkitys päätöstukijärjestelmien perusaineistona on kasvamassa ja sen tietosisällön parantaminen on olennaista. Tärkeitä kehittämiskohteita ovat etenkin puulajeittaiset puustotunnukset kuten runkolukusarjat ja laatutiedot, maaperätiedot sekä ympäristöasioiden hallintaa tukevat tietokokonaisuudet.
- Alusta tukee myös bayesilaisten yhdistettyjen estimaattien laskentaa, jos sama suure on saatavissa useasta tietolähteestä epävarmuusarvioineen. Se antaa mahdollisuuksia tuottaa yhä täsmällisempää tietoa metsän käsittelyn tueksi. Tällöin datojen luotettavuusestimaatit ja esimerkiksi joukkoistetusti kerätyn tiedon laadunhallinta korostuvat jatkossa.
- Jatkotutkimusaiheet liittyvät mm. datojen luotettavuusarvioiden tuottamiseen ja hyödyntämiseen sekä hilojen päivityksessä tehokkaiden ja luotettavien kasvumallien kehittämiseen.



Kirjallisuutta

Hämäläinen, J., Räsänen, T., Riekki, K. Sorsa, J-A. & Ritala, R. 2019. Metsätiedon palvelualustan konseptitestaus. Metsätehon raportti 252. Saatavissa: <http://www.metsateho.fi/metsatiedon-palvelualustan-konseptitestaus-raportti>. [Viitattu 24.6.2019].

Hämäläinen, J., Räsänen, T., Ritala, R., Häme, T. & Tergujeff, R. 2017. Seuraavan sukupolven palvelualustan kehittäminen – esiselvitys. Metsätehon raportti 241. Saatavissa: <http://www.metsateho.fi/seuraavan-sukupolven-palvelualustan-kehittaminen-esiselvitys/>. [Viitattu 24.6.2019].

Mäkelä, J., Koljonen, M., Vorne, H., Riekki, K. & Räsänen, T. 2019. ForestJSON-ohje. Metsätiedon palvelualustan konseptitestaus. Metsätehon raportti 252. LIITE. Saatavissa: <http://www.metsateho.fi/metsatiedon-palvelualustan-konseptitestaus-raportti>. [Viitattu 24.6.2019].

Rajala, M. & Ritala, R. 2016. Data platform promoting forest data utilization through uniform access to heterogeneous data. Metsätehon raportti 240. Saatavissa: <http://www.metsateho.fi/data-platform-promoting-forest-data-utilization/>. [Viitattu 24.6.2019].

Riekki, K. & Räsänen, T. 2019. Tietolähteiden tekninen kuvaus. Metsätiedon palvelualustan konseptitestaus. Metsätehon raportti 252. LIITE. Saatavissa: <http://www.metsateho.fi/metsatiedon-palvelualustan-konseptitestaus-raportti>. [Viitattu 24.6.2019].

Ritala, R. & Juntunen, T. 2019. Uusien tietolähteiden sovitukset metsätiedon palvelualustan hilaresoluutioon. Metsätiedon palvelualustan konseptitestaus. Metsätehon raportti 252. LIITE. Saatavissa: <http://www.metsateho.fi/metsatiedon-palvelualustan-konseptitestaus-raportti>. [Viitattu 24.6.2019].

