

Viljelijän datavaraston liittäminen maatalouden data-avaruuteen

Samu Vesiluoma¹, Moona Vesiluoma¹ ja Hannu Haapala²

¹Jyväskylän ammattikorkeakoulu, IT-instituutti, Piippukatu 2, Jyväskylä

²Jyväskylän ammattikorkeakoulu, Biotalousinstituutti, Piippukatu 2, Jyväskylä

e-mail: samu.vesiluoma@jamk.fi

Jyväskylän ammattikorkeakoulun Biotalouskampuksen Älymaatila -kehittämishankkeessa (2021–2023) kehitettiin tutkimuskäyttöön Viljelijän datavarasto, joka keskittää ja tallentaa eri lähteistä kerättyä maatalouden dataa. Data kerättiin Tarvaalasta hankkeen käytössä olleelta tutkimuspelloilta. Viljelijän datavarasto mahdollistaa datan hallinnan, analysoinnin sekä visualisoinnin. Tarkoituksena oli tutkia Viljelijän datavaraston liittämistä yhtenäiseen maatalouden data-avaruuteen ja tutkimustuloksena saatiin onnistunut tiedonsiirtokanava kahden eri datalähteen välille. Maatalouden data-avaruus avaa uusia mahdollisuuksia maataloudelle, kuten datan arvonluonnin, innovaatiot, päätöksenteon, optimoinnin ja kaupankäynnin.

Avainsanat: Data-avaruus, älymaatila

Johdanto

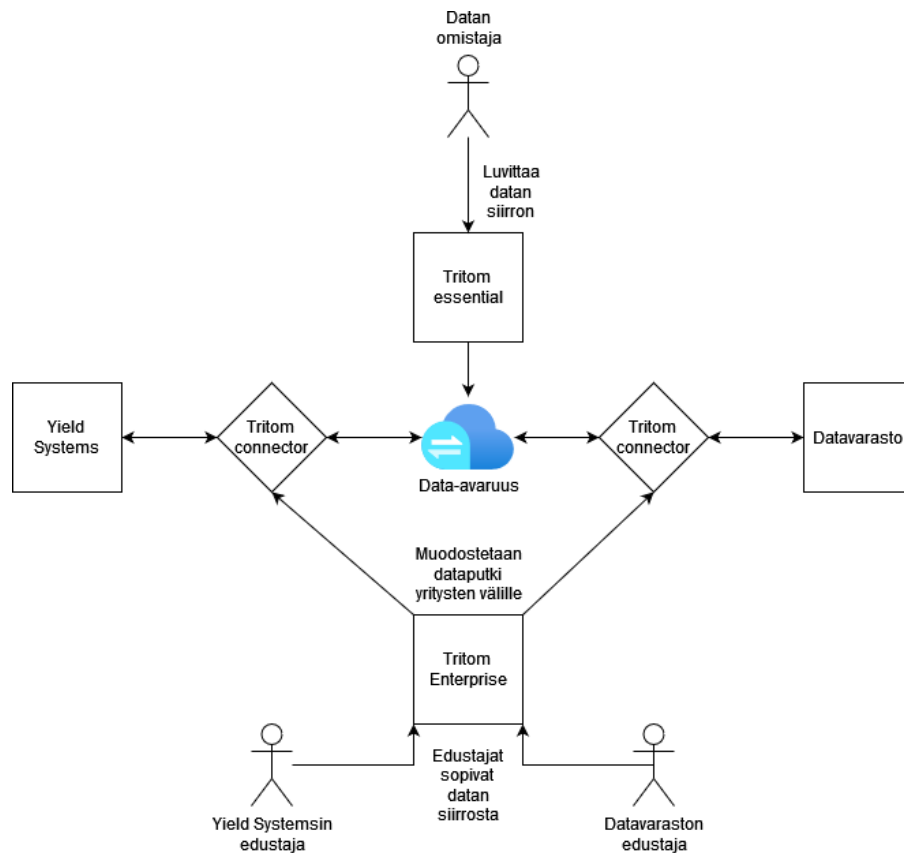
Maatalous on yksi tärkeimmistä elinkeinoista Suomessa, mutta se kohtaa myös monia haasteita. Haasteita ovat muun muassa ilmastonmuutos, kilpailu, tuottavuus ja ympäristövaikutukset. Maataloudessa tuotetaan ja hyödynnetään suuria määriä dataa, joka liittyy esimerkiksi maaperään, kasveihin, eläimiin, sääolosuhteisiin, tuotantoprosesseihin ja markkinoihin. Nykyään data on monimuotoista ja se kattaa laajan kirjon erilaisia tietoja. Se voi olla esimerkiksi numeerista tietoa, kuten lämpötila tai sademäärä, tai kuvallista tietoa, kuten satelliitti- tai dronekuvia pelloista. Dataa kerätään monin eri tavoin, kuten sensoreilla, satelliiteilla, dronella tai jopa maanviljelijöiden omilla havainnoilla. Kerätty data kerätään tyypillisesti laitevalmistajien pilvipalveluihin. Ongelmaksi nousee kerättyjen tietojen integrointi ja hyödyntäminen helposti ja tehokkaasti niin että se palvelee kaikkia osapuolia. Datan hyödyntäminen ei myöskään ole niin tehokasta kuin se mahdollisesti voisi ideaali tilanteessa olla.

Datan tehokas hallinta, analysointi ja visualisointi ovat avainasemassa maatalouden kehittämisessä, sillä ne mahdollistavat paremman ymmärryksen ja päätöksenteon. Riittävän datamäärän avulla on mahdollista kehittää monenlaisia analyyseja tai jopa kouluttaa tekoäly- ja koneoppimismalleja, jotka auttaisivat syvemmän ymmärryksen luomisessa. Tässä tutkimuksessa esitellään Viljelijän datavaraston liittämistä maatalouden data-avaruuteen Tritom-palvelun näkökulmasta, joka oli yksi Biotalouskampuksen Älymaatila -kehittämishankkeen tuloksista. Data-varasto sisälsi dataa eri lähteistä, kuten maaperäsensoreista, drone-kuvista, peltolohkokartasta ja satoennusteesta.

Materiaalit ja menetelmät

Jamkissa kehitetty Viljelijän datavarasto liitettiin osaksi maatalouden yhtenäistä data-avaruutta Tritom-palvelun avulla. Data-avaruuteen yhdistämisessä auttoi kumppaniyritys Yield Systems, joka tarjosi tutkimuksessa käytetyn pellon satoennustedataa viljelijän datavarastoon. Toisena kumppanina oli DataSpace Europe, jonka Tritom-palvelun avulla tietojen välittäminen ja luvittaminen tapahtuivat. Datan luvittaminen ja siirtäminen osoitettiin vuonna 2023 Okra-maatalousnäyttelyssä sekä Jamkin Biotalouskampuksen Pellonpiennarpäivillä järjestetyissä demoissa.

Demoa varten asennettiin Tritom-palvelu ja Viljelijän datavarasto testipalvelimelle. Molemmat palvelut asennettiin käyttäen docker-kontteja niiden helppokäyttöisyyden takia. Demon kolmasosapuoli Yield Systems asensi Tritom-palvelun myös omaan ympäristöönsä datan lähettämistä varten. Kuvassa 1 havainnollistetaan demon toimintakaavio ja kuinka Tritom liittyy data-avaruuteen liittymiseen. Docker-konttien lisäksi molemmat osapuolet muodostavat ja hyväksyvät datakanavat osapuolten välille sovitulla datalla Tritom Enterprise -palvelussa. Datakanavan luonnin jälkeen pitää datan siirtoa varten käyttäjän tai datanomistajan luvittaa sen siirtäminen osapuolten välillä Tritom Essentialissa. Ilman luvitusta data ei siirry kohteesta toiseen.



Kuva 1. Demon toimintakaavio, jossa data-avaruuteen liitytään datan välitys palvelun avulla.

Demossa haluttiin visuaalisesti esittää datan siirtyminen data-avaruudessa, joten Yield Systemsin dataa visualisoitiin sen siirtymisen jälkeen datavaraston tarjoamalla karttapohjalla. Visualisointia varten data tuli mallintaa datavarastoon sopivaksi ennen varsinaista tiedonsiirtoa. Tritom-palvelu tarjoaa mahdollisuuden kuvailla dataa toisille käyttäjille ennen data yhteyden muodostamista. Datakuvaus ennen itse siirtämistä on tärkeää, koska Tritom on vain datanvälityspalvelu ja se kattaa vain yhden osan ETL-tietojen käsittelyprosessista.

ETL on tietojen käsittelyn prosessi, joka koostuu kolmesta vaiheesta: Extract (poiminta), Transform (muunnos) ja Load (lataus).

Extract: Tässä vaiheessa data poimitaan eri lähteistä. Tritom-palvelu mahdollistaa tämän vaiheen.

Transform: Tässä vaiheessa raakadata käsitellään ja muunnetaan sen lopulliseen tallennusmuotoon tai -rakenteeseen. Tämä voi sisältää datan puhdistamista, validointia ja yhdistämistä. Tämä vaihe suoritetaan datavarastossa.

Load: Tässä lopullisessa vaiheessa käsitelty data ladataan datavarastoon.

ETL-prosessi on perinteinen tapa yhdistää dataa useista järjestelmistä yhteen tietokantaan, tietovarastoon tai datalakeen. Se mahdollistaa datan analysoinnin ja liiketoimintapäätösten tekemisen. ETL-prosessia käytetään usein tietovarastoinnissa (Klepman 2017).

Näiden vaiheiden jälkeen Yield Systemsin data oli valmiina visualisoitavaksi karttapohjalle. Demon aikana tiedot poistettiin ja siirrettiin useaan kertaan toiminnan esittämiseksi.

Tulokset

Vuonna 2023 esitetyt demot todistivat, kuinka maatalouden eri toimijat voidaan yhdistää yhtenäiseen data-avaruuteen. Data-avaruus avaa uusia mahdollisuuksia maataloudelle. Se mahdollistaa entistä paremman datan arvonluonnin, uusia dataja yhdistäviä innovaatiota, entistä tilannetietoisen päätöksenteon ja optimoinnin (Pesonen, Haapala, Hyväluoma, Kallio, Karjalainen, Linna ja Ruponen N.d.). Se myös mahdollistaa dataan perustuvan kaupankäynnin, eli viljelijät ja yritykset voisivat myydä ja ostaa heitä kiinnostavaa dataa (Ala-Kleemola 2022). Data-avaruuteen liittyminen tarjoaa datan arvonluonnin, jonka avulla on mahdollista laskea esimerkiksi tuotteen hiilijalanjälki. Uusien innovaatioiden luominen yhteisessä data-avaruudessa on huomattavasti helpompaa, jos dataa on mahdollista saada selvien pelisääntöjen ja protokollien mukaisesti.

Tulosten tarkastelu

Maataloudessa on suurta potentiaalia olla ensimmäisten toimialojen joukossa luomassa yhtenäistä data-avaruutta, jonka avulla datan omistajat voisivat luoda datamarkkinoita, joista he oikeasti hyötyisivät. Data-avaruuteen liittyminen Tritomin avulla ei ollut erityisen työlästä, mutta suurin työ mitä yritys joutuu tekemään, tapahtuu datan siirtämisen jälkeen. Datan sovittaminen yritykselle sopivaan käyttöön on data-avaruuteen liittymisen työläin osuus. Tämän kaltaista demoa ei tiedettävästi maatalouden data-avaruuden puolella ole kokeiltu, joten nämä tulokset ovat ainakin Suomen mittakaavassa ensimmäisiä.

Johtopäätökset

Datan jakaminen ja yhdistäminen eri toimijoiden sekä alustojen kanssa lisää datan arvoa ja mahdollistaa uusia yhteistyömuotoja. Maatalouden yhtenäinen data-avaruus avaa uusia mahdollisuuksia maataloudelle, kuten datan arvonluonti, innovaatiot, päätöksenteko, optimointi ja kaupankäynti. Tämän kaltaisesta datataloudesta hyötyvät kaikki maatalouden osapuolet: viljelijät, yritykset, tutkijat, neuvontapalvelut, viranomaiset ja kuluttajat.

Maatalouden data-avaruuden hyödyntämistä varten on erityisen tärkeää, että maatalousalan yritykset liittyisivät Tritom-palvelun kaltaisiin datan välityspalveluihin. Kunnollista data-avaruutta ei voi rakentaa, jos yritykset ja maanviljelijät eivät halua liittyä datan välityspalveluihin. Ilman datan välityspalvelua datan omistajat eli maanviljelijät jäävät mahdollisesti ulkopuolisiksi ja integraatiot tapahtuisivat vain yritysten välillä. On myös mahdollista, että yritykset eivät halua liittyä data-avaruuteen, koska eivät koe saavansa siitä hyötyä itselleen tai he eivät mahdollisesti osaa hyödyntää käytettävissä olevaa dataa välityspalvelun kautta.

Kiitokset

Tämä tutkimus ja demo on toteutettu ”Biotalouskampuksen Älymaatila -hankkeessa”, jonka päärahoitus on saatu Euroopan aluekehitysrahastosta (EAKR) Keski-Suomen liiton kautta. Kiitos myös erityisesti yhteistyö yrityksille Dataspace Europe oy ja Yield Systems oy demon mahdollistamisessa.

Kirjallisuusviitteet

Ala-Kleemola, K. 2022. Data-avaruus mullistaa viljelijöiden mahdollisuudet hyötyä tilalla tuotetusta datasta. Artikkelit data-avaruus verkkosivulla. Julkaistu 31.8.2022. Viitattu 15.12.2023. <https://www.data-avaruus.fi/fi/artikkelit/data-avaruus-mullistaa-viljelijoiden-mahdollisuudet-hyotya-tilalla-tuotetusta-datasta>

Haapala, H. 2023. Datan avulla lisää euroja maatalousalalle – miten? Blogiteksti Jamkin verkkosivulla. Julkaistu 16.1.2023. Viitattu 15.12.2023. <https://blogit.jamk.fi/tarvaalantarinoita/2023/01/16/datan-avulla-lisaa-euroja-maatalousalalle-miten/>

Klepman, M. 2017. Designing Data-Intensive Applications. Sebastopol, Kalifornia, Yhdysvallat: O'Reilly Media, Inc. s. 91–92.

Pesonen, L., Haapala, H., Hyväluoma, J., Kallio, K., Karjalainen, S., Linna, P. & Ruponen, O. N.d. Älymaatalous 2030 tiekartta. Asian-
tuntijaraportti-pdf maaseutuverkoston verkkosivuilla. Viitattu 15.12.2023. <https://projects.luke.fi/agrihubi/wp-content/uploads/sites/52/2022/02/Alymaatalous-2030-tiekartta-.pdf>