

Massadatan ja mydatan strategiset linjaukset ja toimet

Lausuntopyyntö LVM/1996/13/2015

1. Johdanto

Liikenne- ja viestintäministeriö on pyytänyt lausuntoa massadatastrategiastaan. Data- ja siihen tiiviisti kytkeytyvä palvelutalous tarjoavat suomalaisille yrityksille erinomaisen mahdollisuuden tehdä kestävä kilpailukykyloikka tuottamalla lisäarvoa tiedolla ja palveluajattelulla. Julkinen sektori voi puolestaan tehostaa oleellisesti toimintaansa parantamalla päätöksentekoa tietojohdamisella ja siirtymällä älykkääseen toimintamalliin - enemmän vähemmällä.

Avainasemassa massadatan hyödyntämisessä ovat yritykset ja julkisen sektorin toimijat mutta myös kansalaiset, jotka ovat paitsi toiminnan kohteita niin myös datatalouden toimijoita. Massadatan täysimittainen hyödyntäminen edellyttää viranomaisilta toimia, jotka luovat hyödyntämiselle edellytyksiä ja mahdollisuuksia sekä poistavat esteitä. Lisäksi tarvitaan näkyviä ja kuuluvia esimerkkejä, jotka rohkaisevat yrityksiä investoimaan ja julkista sektoria uudistamaan toimintaansa.

LVMn strategia edistää tärkeää asiaa ja strategialuonnos sisältää oikeita asioita. COSS tuo lausunnonaan keinoja strategian toteuttamiseksi sekä uusia, mielestämme relevantteja näkökulmia.

2. Kommentteja kappaleisiin 'Massadata- ja mydatastrategian tavoitteet' ja 'Taustaa'

Massadatastrategian tavoite tulisi olla kilpailukykyinen, datatalouden hyvinvointi-Suomi. Strategia tulee jalkauttaa kaikille toimialoille. Erityisesti tulisi keskittyä uudistumisen, kasvun ja hyvinvoinnin kannalta kriittisiin toimialoihin.

Tarvitsemme esimerkkejä, tiennäyttäjiä matkalla datatalouteen. Jo nyt löytyy erinomaisia esimerkkejä näyttämään mallia datan ja palveluajattelun hyödyntämisestä ja niitä tulisi nostaa strategian taustaosuuteen. On monia julkisen sektorin toimijoita, jotka ovat avanneet dataa ja olleet luomassa uusia palveluja esimerkkinä HSL & Reittiopas, <http://www.reittiopas.fi/>. Seuraavassa on esimerkkejä yrityksistä, jotka ovat rakentaneet datasta menestyksellistä liiketoimintaa:

- **Vainu** löytää asiakkaita asiakkailleen analysoimalla avointa dataa internetistä – kasvanut kahdessa vuodessa yli 30 työntekijän yritykseksi ja se on perustamassa toimipisteitä ulkomaille: <https://vainu.io/>
- **Suomen Asiakastieto** perustaa luottotietoliiketoimintansa PRHn avoimiin tietokantoihin, joita se täydentää ja jalostaa. Suomen Asiakastieto on kannattava yritys tuottaen yli 40% voiton liikevaihdosta: <http://www.asiakastieto.fi/web/fi>
- **Kemppi** kerää dataa hitsauslaitteidensa avulla ja tarjoaa asiakkailleen Welding

- Management -palvelua pilvestä: <http://www.kemppi.com/fi>
- **Vaisala** tekee säähavaintoasemia mutta myy säätietoa palveluna eri toimialojen asiakkailleen: <http://www.vaisala.fi/fi/Pages/default.aspx>
- **MariaDB Corporation**: kehittää yhtä maailman suosituimmista tietokantaohjelmistoista ja tarjoaa siihen liittyen palveluita yrityksille, jotka keräävät ja käsittelevät dataa <https://mariadb.com/>

3. Yleisiä kommentteja kappaleeseen 'Valtioneuvoston linjaukset ja toimenpiteet massadatan hyödyntämiseksi liiketoiminnassa'

Strategialuonnoksesta puuttuu vielä kuvaus strategian toimeenpanemiseksi. Tarvitaan tavoitteet ja niitä mittaavat mittarit sekä konkreettisia toimenpiteitä tavoitteiden toteuttamiseksi. Ohjauksen ja mittaamisen avulla tiedetään miten strategia on toteutunut, mitä on saatu aikaan. Ohjausta varten tulee nimittää ja resursoida hankehallinnan organisaatio, joka ohjaa ja seuraa strategian toteutusta.

Periaatelinjauksien tulee johtaa toimeenpanoon, konkreettisiin hyötyihin. Valmistuttuun strategia ja sen periaatelinjaukset tulee sisältää vain sellaisia päätöksiä ja toimia, jotka on vastuutettu ja joiden toteuttamiselle on osoitettu resurssit. Muussa tapauksessa linjaukset uhkaavat jäädä hyviksi toiveiksi.

4. Kommentteja linjauksiin

4.1 Sääntely ja datapolitiikka

Linjaus 1.1 Tietoaineiston käytön edistäminen

Data-analytiikan avulla voidaan tehostaa toimintaa ja parantaa kansalaisten elämänlaatua mutta hyödyntämisessä on esteitä ja hidasteita, joita tulisi poistaa. Data-analytiikan avulla on esimerkiksi mallinnettu ikäihmisten riskiä joutua laitoshoitoon. Mallin avulla riskiryhmän seniorit voidaan tunnistaa ajoissa ja tehotoimien avulla he voivat elää pitempään kotona tutussa ympäristössä ja yhteiskunnan varoja säästyy. Samalla tavalla voitaisiin auttaa nuoria ja vaikka maahanmuuttajia. Yhdistämällä anonymisoituja sosiaali-, terveys- ja koulutietoja olisi mahdollista mallintaa pudokkuutta ja estää siten lapsia ja nuoria putoamasta koulusta ja yhteiskunnasta. Tietojen yhdistämistä eri tietolähteistä tulisi helpottaa siten, että viranomaiset saavat käyttöönsä tehtävässään tarvittava oleellinen tieto - yksityisyyden suojaa vaarantamatta. Datan hyödyntämiseen liittyvää suostumuksen hallintaa tulisi kehittää. Terveyden tutkimusaineiston käyttöön liittyy perustellusta syystä tarkka suostumuslupamenettely, jolla taataan ihmisten oikeudet omiin tietoihin ja tietosuojaa. On törmätty tilanteisiin, joissa tutkimustiedon käyttö uusissa käyttötilanteissa vaatii niin työläitä lupaprosesseja, että tieto jää hyödyntämättä. Yksi syy työläyteen on se, että lupia ei osattu aikanaan kysyä siten, että tietoa voitaisiin hyödyntää uusiin käyttötarkoituksiin. Voitaisiinko lupamenettelyjä yksinkertaistaa siten, että arvokas data olisi hyödynnettävissä monipuolisesti vaarantamatta tietosuojaa? ('tutkimustulosten toissijainen käyttö').

Linjaus 1.2 Kyvykkyyksien kehittämisellä kasvua

Strategialuonnoksen ehdotukset osaajien kouluttamiseksi ovat oikeita, datatalous tarvitsee lisää dataosaajia. Dataosaaminen ei kuitenkaan riitä vaan tarvitaan myös kykyä hyödyntää datasta löytyviä oivalluksia ja mahdollisuuksia osana toiminnan ja palvelujen kehittämistä. Tarvitaan toiminnan kehittämisen ja palveluajattelun osaajia uudistamaan prosesseja ja toimintamalleja, rakentamaan uudenlaisia palveluja yhdessä asiakkaiden, asiakkaiden asiakkaiden ja muiden toimijoiden kanssa. Toiminnan kehittämisen taitoja ja erityisesti palveluajattelua tulee opettaa korkeakouluissa osana datatalouden opiskelua.

4.2 Datan palveluinfrastruktuurin ja käytänteiden kehittäminen

Linjaus 2.2 Datan jakaminen vauhtiin

Datan jakamisen ja uudelleenkäytön potentiaali on valtava. Potentiaali piilee erityisesti kun avointa ja yritysten omaa dataa yhdistellään uusiksi oivalluksiksi. Potentiaali valuu hukkaan mikäli tiedot eivät ole yhteensopivia.

Tulee panostaa siihen, että datan uudelleenkäyttö ei esty tarpeettomasti sen takia, että data ei ole käytettävissä yhteensopivassa muodossa. Tietojärjestelmät tulee suunnitella yhteensopivuutta silmälläpitäen. Ostajan tulee edellyttää ostamaltaan järjestelmästä dokumentoituja, hyvin määriteltyjä rajapintoja yhteensopivuuden varmistamiseksi. Muuten voi käydä, että datan omistaja jää toimittajan vangiksi loukkuun eikä voi hyödyntää dataansa ja kehittää toimintaa minkään muun kuin alkuperäisen toimittajan kanssa monopolin ehdoilla. VM:n JHS 166 -suositus sisältää mekanismeja, joilla voidaan sopimuksellisesti varmistaa, että tilaaja ei joudu dataloukkuun. Teettäjien ja ostajien koulutuksen lisäksi tarvitaan toimittajien asennemuutos. Ammattitaitoinen toimittaja huolehtii, että data on saatavissa järjestelmästä vaivattomasti uusiokäyttöön.

Tarvitaan toimialakohtaisia metadata- ja ontologiakirjastoja, jotka edistävät yhteensopivuutta ja ohjaavat kehitystä kohti toimialastandardeja. Toimialaorganisaatiot ja -järjestöt olisivat luonnollisia tahoja kokomaan standardeja, esimerkkinä sote-alueella STM osana Kaste -toimintaa, opetustoimi / OKM ja Opetushallitus, tele /Ficom, kauppa /Kaupan liitto, pankki /Finanssialan keskusliitto, jne.

Linjaus 2.3 Valtion johdolla synergiahyötyjä

Hallitus on tehnyt päätöksiä tyyliin 'pätetään ensin ja katsotaan seuraukset jälkeenpäin' (Juha Sipilä <http://www.hs.fi/kotimaa/a1433995519929>). Valistunut päätöksenteko perustaa päätökset ymmärrykseen ja keskusteluun päätöksen seurauksista. Datalla johtaminen hyödyntää päätöksenteossa analytiikkaa ja ymmärtää ilmiöiden syy-seuraussuhteita. Julkisella sektorilla datalla johtamisen tärkein alue on siiloutunut sote-sektori. On

esimerkkejä, joissa säästöt proaktiivisessa sosiaalitoimessa maksetaan moninkertaisesti reaktiivisessa sairaudenhoidossa. Esimerkiksi yhdistämällä sosiaali- ja terveystoimen tiedot työllisyyden ja koulutoimien tiedoilla saataisiin rakennettua malli, jonka avulla voidaan ymmärtää paremmin päätösten ja toimenpiteiden syy- seuraussuhteita sekä säästöjen ja investointien kokonaiskustannuksia. Keskustelu eläkerahastojen riittävydestä ja eläkkeiden sitomisesta palkkaindeksiin käy kuumana. Jos eläkerahastojen simulointimallit, niiden parameterit ja data avattaisiin julkiselle keskustelulle, ymmärrys lisääntyisi ja päästäisiin nopeammin toimeenpantavissa oleviin valistuneisiin päätöksiin.

4.3 Dataa hyödyntävä analytiikka-, palvelu- ja sovelluskehityksen edistäminen

Linjaus 3.1 Datan soveltaminen ja yhteiskäyttö, toiminnan skaalaaminen

Strategiassa tulee hyödyntää yhteistyötä muiden kehittämisohjelmien ja tahojen kanssa. Yksi luonteva yhteistyötaho on kuuden suuren kaupungin 6aika -ohjelma, joka rakentaa kestävä ja älykästä Suomea: <http://6aika.fi/>. Esimerkiksi COSS:n (Center of Open Source Solutions) johtaman konsortion hanke-ehdotuksessa 'Datasta oivalluksia ja bisnestä' rakennetaan data-analytiikan ja palvelumuotoilun keinoin oivalluksia ja liiketoimintaa. Hankkeessa on kolme projektia: 'Aviapolis, palveluja matkaajille' -projektissa analysoidaan Finnairin ja Finavian dataa ja ideoidaan palveluja Helsinki-Vantaan lentokentälle. Oulussa toimiva Hyvä pitkä elämä - projektissa etsitään tutkimusdatasta analytiikan keinoin selityksiä hyvään ja pitkään elämään ja ideoidaan palvelumuotoilun avulla Mydata- sovelluksia ja muita palveluja ikäihmisille. Projekti tukeutuu Oulun yliopiston koordinoimaan ja Tekesin rahoittamaan Digital Health Revolution -ohjelmaan <http://cht oulu.fi/news/digital-health-revolution-was-launched-in-oulu>. Hankkeen kolmannessa projektissa 'Datasta oivalluksia ja bisnestä -valmennus' päivitetään yritysten kilpailuetu datatalouteen. Projektissa valmennetaan 20 ICT -yritystä ottamaan analytiikka ja palvelumuotoilu osaksi strategiaansa. Valmennusohjelma tuoteistetaan ja julkaistaan COSS:n palveluna avoimen lähdekoodin analytiikkatyökaluineen kaikkien korkeakoulujen ja halukkaiden käyttöön. Katso video: bit.ly/datastaoivalluksia

Strategialinjauksessa on ehdotettu innovaatiohub:n perustamista edistämään dataliiketoimintaa. Sellaista oltiin rakentamisesta vuosina 2013-2014 TEM:n, Tekesin, VTT:n, Aallon, IBM:n ja nykyisen EIT Digitalin kesken. Tuolloin ei kuitenkaan ollut olemassa rahoitusinstrumenttia, jolla rahoittaa innovaatioalustaa. Sellainen instrumentti tulee luoda ja perustaa massadataa, analytiikkaa ja palvelumuotoilua hyödyntävä innovaatiokeskittymä edistämään dataliiketoimintaa. 6Aika- ohjelman hanke Datasta oivalluksia ja bisnestä tuottaa innovaatiokeskittymään tarvittavia työkaluja ja menetelmät ja toiminta sopisi osaksi Espoossa toimivan EIT Digitalin Business Development Accelerator & Business Community-konseptia <https://www.eitdigital.eu/innovation-entrepreneurship/business-development-accelerator/>. Globaalina toimijana EIT Digital toimii erinomaisesti sillaksi ajatuksille ja yrityksille myös maailmalle.

Linjaus 3.2 Kansainvälistyminen

Suomen avoimien tietojärjestelmien
keskus – COSS ry
PL 35
33101 Tampere
info@cooss.fi

Lausunto
2016-02-12

5(5)

Suomella on hyvät mahdollisuudet julistautua datatalouden veturiksi myös kansainvälisille toimijoille. Tavoitteen toteuttamiseksi tulee varmistaa lainsäädännön keinoin, että Suomi on jatkossakin datataloudelle otollinen maa järjestämällä tietoturva ja -suoja korkealle tasolle ilman lakia kohdistamattomasta massavalvonnasta.

Suomen avoimien tietojärjestelmien keskus – COSS ry



Timo Väliharju
toiminnanjohtaja

