

Tämä on rinnakkaistallenne.  
Rinnakkaistallenteen sivuasettelut ja typografiset yksityiskohdat  
*saattavat poiketa* alkuperäisestä julkaisusta.

Julkaisun tekijä(t): Blomster, Miikka

Julkaisun nimi: Datan tehokas hyödyntäminen vaatii useiden haasteiden ratkaisemista

Julkaisuvuosi: 2022

Versio: Kustantajan versio

Käytä viittauksessa alkuperäistä lähdettä:

Blomster, M. (7.12.2022). Datan tehokas hyödyntäminen vaatii useiden haasteiden ratkaisemista. Digiliike: tietoa digitaalisesta markkinoinnista - blogi.

Haettu 8.12.2022 osoitteesta

<https://blogi.oamk.fi/2022/12/07/datan-tehokas-hyodyntaminen-vaatii-useiden-haasteiden-ratkaisemista/>

# Datan tehokas hyödyntäminen vaatii useiden haasteiden ratkaisemista

[MIIKKA BLOMSTER](#)

JULKAISTU 7.12.2022

Digitaalisen datan monipuolinen ja tehokas hyödyntäminen vaatii yrityksiltä monenlaisia kyvykkyyksiä datan ymmärtämisestä liiketoiminnan luonteen kokonaisvaltaiseen kartoittamiseen. Juuri julkaistun selvitystyömme yhteydessä paneuduimme näihin syvällisemmin, joten tässä käymme läpi vain oleellisimpia haasteita digitaalisen datan hyödyntämiseksi.

Datakyvykkyys dataprosesseissa rakentuu organisaatiossa olevan kokonaisvaltaisen data- ja liiketoimintaymmärryksen varaan. Tässä näkemyksessä yrityksen johdon ja henkilöstön kollektiivinen osaaminen ja ymmärrys liiketoiminnan luonteesta ja sen eri toimintojen datasta muodostavat pohjan datakyvykkyydelle toimivan datastrategian muodossa (1). Datastrategian onnistuminen pohjustaa organisaation toimintoja mukailevien dataprosessien onnistumista (2) ja niiden tehokas toteuttaminen vaatii ymmärrystä tarkoituksenmukaisesta data-arkkitehtuurista ja yrityksen tarpeisiin soveltuvista järjestelmistä ja niiden rakentamisesta ja integraatiosta (3).

Varsinainen kyvykkyys dataprosessin sisällä puolestaan muodostuu organisaatiosta kumpuavan dataosaamisen ympärille datan tarkoituksenmukaisen keräämisen (4), varastoinnin, hallinnan ja jalostamisen (5) ja tulkinnan ja johtopäätösten raportoinnin (6) kautta. Nämä toiminnot ovat osa-alueita, joita yritykset voivat selkeästi pyrkiä johtamaan ja kehittämään, ja puolestaan niissä tapahtuva kehitys parantaa myös yritysten kykyä kehittää organisaation datakyvykkyyden perustuksia. Malli kuvastaa hyvin dataprosessia, mutta se tulee aina linkittää yhteen kulloinkin tarkasteltavan liiketoimintaprosessin kanssa. Siksi onkin olennaista, että dataymmärrystä tukee myös liiketoiminnallinen ymmärrys, ja myös päinvastoin.

4. Datalähteet  
ja datan  
kerääminen

5. Datan  
varastointi,  
hallinta ja  
jalostaminen

6. Datan  
tulkinta ja  
raportointi

3. Data-arkkitehtuuri ja järjestelmät

2. Dataprosessit ja organisaation toiminnot

1. Datastrategia

*Datakyvykkyyksien rakentuminen organisaatiossa*

Tehokkaat *datan hyödyntämisen johtamis- ja toimintamallit* muodostuvat organisaatiosta kumpuavan dataosaamisen ympärille datan tarkoituksenmukaisen keräämisen, varastoinnin, hallinnan ja jalostamisen ja tulkinnan ja johtopäätösten raportoinnin kautta. Nämä toiminnot ovat osa-alueita, joita yritykset voivat selkeästi pyrkiä johtamaan ja kehittämään, ja puolestaan niissä tapahtuva kehitys parantaa myös yritysten kykyä kehittää organisaation datakyvykkyyden perustuksia.

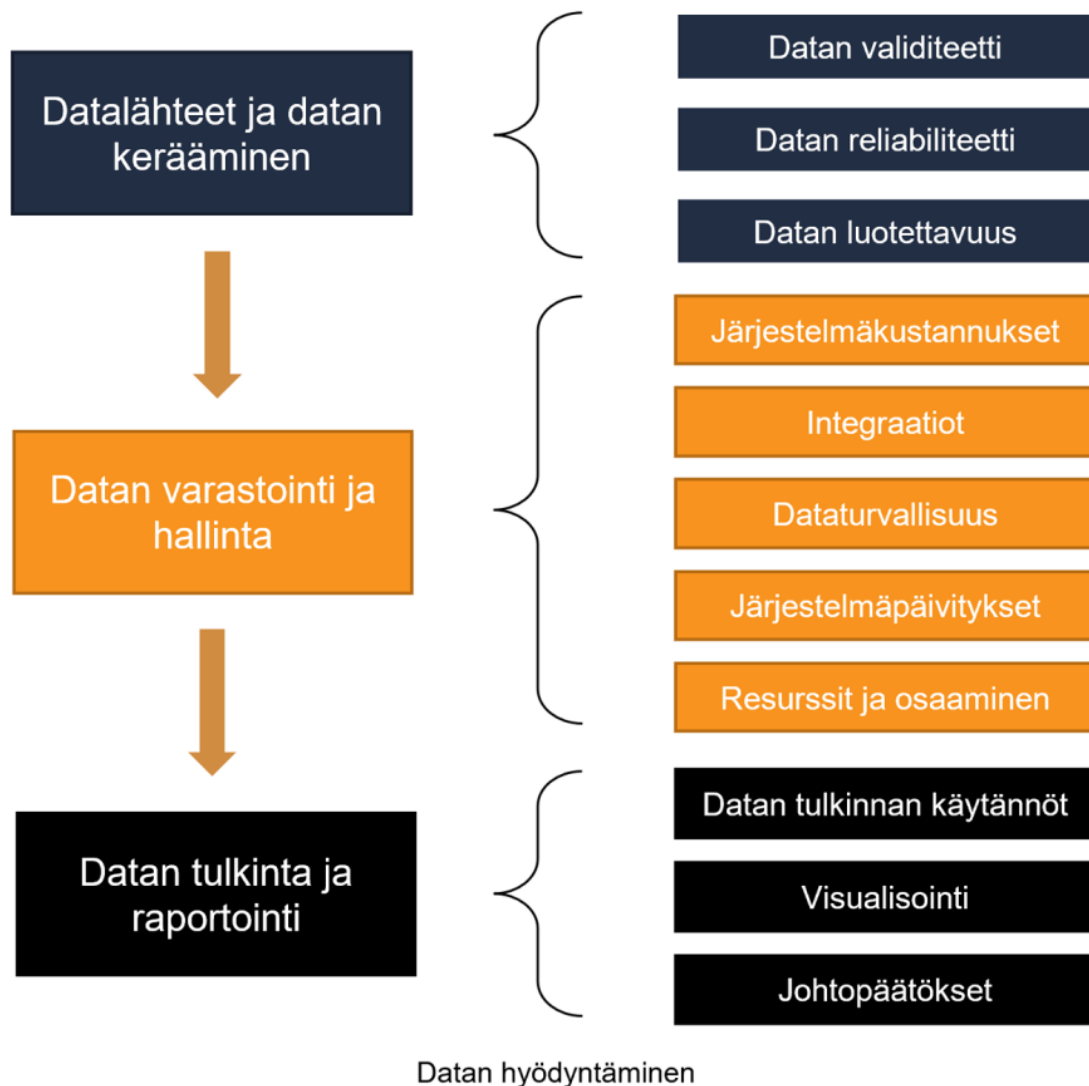
Parhaiten *datalähteiden tunnistamisessa ja datan keräämisessä* suoriutuneet yritykset olivat myös aktiivisesti kehittäneet toimintojaan dataohjautuvaksi ja datan rooli niissä oli varsin hyvin johdettu liiketoiminnallisen tarpeen perusteella. Toisessa päässä olivat yritykset, joilla oli puutteita liiketoimintojen ymmärryksen tehostamisen kanssa, eli ne eivät ihan suoraan ymmärtäneet millainen datan rooli toiminnoissa on, ja miten sitä voidaan hyödyntää liiketoiminnan kehittämisessä. Jotta datalähteet saataisiin laadukkaasti tuotua liiketoimintaprosesseihin lisäarvoa tuoviksi, olisi ensimmäinen askel vastuiden määrittäminen *datan omistajuudesta*, tai kuten data-mesh-ajattelutavassa, yhdistettäisiin datan tuottajat, käsittelijät ja kuluttajat suunnittelemaan datatuotteita ja pohtimaan miten uutta dataa minnekin viedään, varastoidaan, hallitaan ja jalostetaan. Tämä vaatii (yrityksen

koosta ja organisaatiotavasta riippuen) eri toimintojen henkilöiden keskinäistä kontribuutiota, datavirtojen avoimempaa integraatiota, ja toimintojen tehokasta koordinoitua. Toisin sanoen, *yritysten tulee johtaa datatyötä ja -prosessia systemaattisesti*, jotta datasta päästään laadukkaaseen tulkintaan ja johtopäätöksiin ja sitä kautta datan tehokkaaseen hyödyntämiseen.

Datan keräämisen ja datalähteiden tunnistamisen suhteen oleelliset havainnot liittyvät *oleellisten datalähteiden ja niiden keskinäisten suhteiden tunnistamiseen*. Kun datamassat kasvavat, myös niiden validointiin ja luotettavuuteen joudutaan kiinnittämään enemmän huomiota. Jos data on ihmisten syöttämää (esim. B2B-myyntidataan liittyviä tietoja) niin määrän kasvaessa myös virheherkkyys kasvaa. Datajoukkojen yhdistäminen ja eri datalähteiden keskinäisten suhteiden ymmärtäminen ja integroiminen tapahtumadatan validoimiseksi voi tulla kysymykseen

Selvitystyön perusteella varsinainen *datan varastointi ja hallinta* vaatii yritykseltä päämäärätietoisuutta datan laajamittaista hyödyntämistä kohtaan. Sillä tulisi olla aikomuksenaan keskittyä liiketoimintojensa tehostamiseen datan avulla ja valmiutta investoida laajamittaisempaan data-arkkitehtuuriin. Se että tuleeko data-arkkitehtuurista keskitetty vai hallitaanko sitä data-mesh-periaatteen mukaisesti soveltuvin käytäntein, on yrityksen strateginen päätös, joka tosin riippuu toimialasta, yrityksen organisointirakenteesta, toimintojen tuottamasta datasta yms. seikoista. Varastoinnista ja hallinnasta päätettäessä tavoitteena tulisi olla, että data saadaan johdettua aukottomasti kohti sen laajamittaista hyödyntämistä ja parempaa päätöksentekoa. Varastoinnin ja hallinnan tulee mahdollistaa datan saaminen eri toiminnoista integroidusti päätöksentekoon, jotta eri toiminnoille voidaan määrittää *oleelliset mittarit ja raportoinnin keinot*.

Johdon päätöksenteko strategiasta ja mitä dataa, mistä ja miten kerätään



*Dataprosessi yrityksessä*

Selvitystyön yhteenvedona voidaan nostaa muutamia selkeimpiä Oulun alueen yrityksissä havaittuja haasteita datan laajamittaista hyödyntämistä ajatellen. Näitä olivat:

1. *Organisaatioiden strategisen ymmärryksen puutteellisuus datan hyödyntämisen edistämiseksi*

- Organisaation strategia, toimintamallit ja data eivät useinkaan näyttäytyneet toisiaan tukevinä, datalle ei osattu hahmottaa selkeää strategista roolia ja se koettiin enemmän operatiivisia toimintoja ohjaavana työkaluna. Tämän tulkittiin johtuvan datatyön strategisen johdon puutteesta. Osaltaan tätä tuki myös havaittu osaamiskuilu datatyön tekijöiden ja yritysten ylimmän johdon välillä.
- Myös yritysten ylimmän johdon ymmärrys erilaisten data-arkkitehtuurien rakenteesta, sekä yrityksen omista tarpeista

sen suhteen, oli ainakin osittain puutteellinen. Useassa yrityksessä pääasiallinen järjestelmäinfrastruktuuri on keskittynyt eri toimintojen sisälle, joille ei ole luotu integroitua datan hyödyntämisen mallia, ja siten strateginen datan hyödyntäminen eri toimintojen kesken ja liiketoiminnan kokonaisvaltaisen johtamisen kannalta jää puutteelliseksi. Tämä luonnollisesti korostui maturiteettitasoltaan heikoiten sijoituvissa yrityksissä.

## 2. *Kyky mallintaa liiketoimintaprosesseja ja toimintoja*

- Yritysten liiketoimintaprosessien mallintaminen, datapisteiden havaitseminen ja dataprosessin luonti on haastava kokonaisuus, jolle tulee pystyä määrittämään niin soveltuvat järjestelmät kuin yrityksen toimintakulttuuriin soveltuvat toimintamallit. Vastuiden määrittäminen, datan omistajuus, data tuotteena jne. ovat käsitteitä, jotka tulisi pystyä tuomaan organisaatiossa henkilöstölle käsitteellisiksi konsepteiksi ja määrittämään niille rakenteet. Liiketoimintojen ja niissä liikkuvan datan tarkoituksenmukainen hyödyntäminen kumpuaa varsin vahvasti toimintojen ja liiketoimintaprosessien oikeaoppisesta mallinnuksesta.

## 3. *Data-arkkitehtuurin rakentaminen ja hallinta*

- Markkinoilla on paljon erilaisia järjestelmiä yritysten hyödynnettäväksi datatyössään ja niiden läpikäyminen on suuri urakka ja oikean vaihtoehdon valitseminen vaikeaa. Yrityksille tämä nähdään resurssikysymyksenä, sillä selvitystyön aineiston perusteella muodostettiin johtopäätös tietovajeesta, joka rajoittaa yritysten ymmärrystä dataprosessin hallinnasta ja siten soveltuvista järjestelmistä liiketoiminnan tarpeisiin.

## 4. *Integraatiot*

- Järjestelmien integraatiovaikeuksia olivat kokeneet useat yritykset. Niiden huomattiin vaikeuttavan paitsi toimintojen ohjausta ja liiketoiminnan tehostamista, myös monen muun datan validointia (kuin integrointivaikeuksia kokevan datan). Esimerkiksi, jos historiallisesti asiakasdata on kerätty taloushallinnon alle ja taloushallinnon dataa ei saada integroitua myynnin ja markkinoinnin tapahtumadatan kanssa, syntyy vaikeuksia tunnistaa markkinoinnin ja myynnin toimenpiteiden vaikutuksia toteutuneisiin myynteihin ja koko myyntisuppilon määrittäminen vaikeutuu, joka johtaa liiketoiminnan johtamisen vaikeuksiin.

## 5. *Datatyön toimintamallien puuttuminen*

- Dataputkien rakentaminen, prosessien luonti ja työnkulkujen määrittäminen on erittäin haastavaa työtä ja niiden saaminen

toimiviksi vie aikaa ja vaatii resursseja. Niukoista resursseista huolimatta yritysten täytyisi kyetä tunnistamaan itselleen tärkeät datalähteet ja luomaan niiden hyödyntämiseksi toimivat prosessit. Tämä edellyttää selkeiden tavoitteiden asettamista eri toiminoille ja niissä oleville prosesseille. Tavoitteiden mukaan mitatut työnkulut ja niiden säännölliset mittaamissyklit, eli aikatauluttaminen, on oleellinen osa datatyön ja liiketoimintaprosessien yhteensovittamista ja siten muovaa myös datastrategiaa.

#### 6. *Raportointi- ja seuraamiskäytänteiden tarkoituksenmukaisuus*

- Selvitysaineistosta ilmeni, että Oulun alueen yritysten kyky luoda selkeät ja johdonmukaiset johtamismallit dataa hyödyntäen olivat varsin vaihtelevat. Osa yrityksistä koki olevansa hyvin perillä datan hyödyntämisestä, mutta dataan perustuvien johtamiskäytänteiden seuraamisjaksot saattoivat olla useiden kuukausien mittaisia, tai mittarit ja hyödynnetty data ei ollut liiketoimintojen kannalta oleellisinta dataa. Parhaimmissa organisaatioissa oli selkeästi tunnistettu oleellisimmat datat, johtamisessa oli käytössä päivittäinen seuranta ja viikkotasoinen palaverointi sekä toimenpidesuunnittelu.

#### 7. *Datan hallinta, validointi ja luotettavuus*

- Datan hallinnoinnin käytänteet ja validointi vaikuttavat oleellisesti datan luotettavuuteen. Siihen vaikuttavat myös sen ajantasaisuus, datajärjestelmät, datan eheys, osaaminen, tiedon siiloutuminen ja organisaation datan hyödyntämisen kulttuuri. Epäluotettavan datan, tai datan, joka ei kuvasta sitä mitä sen luullaan kuvastavan, käyttämisestä päätöksenteon perustana ei voi suositella kenellekään. Siksi olisi oleellista, että yrityksissä olisi tietotaitoa tunnistaa oleelliset datan lähteet, kyky hallita niitä luotettavasti, validoida käytettävä data ja kasvattaa datan luotettavuutta eri keinoin. Haastattelusta yrityksistä monet eivät kantaneet kovinkaan suurta huolta datan luotettavuudesta.

#### 8. *Kommunikaatio ja datan ymmärtäminen organisaation eri tasojen välillä*

- Yhteinen kieli on aina oleellinen tekijä menestyvän tiimityön kannalta. Useissa haastatteluissa tuotiin esille puutteita yhteisymmärryksessä yritysten johdon (ylin- ja keskijohto) ja datatyöntekijöiden (data-engineerit, analyytikot ja yms.) välillä. Ihmisten erilainen koulutustausta tuottaa aina haasteita yrityksille ja niitä on turha piilotella. On vaarallista jos kumpikin osapuoli, liiketoiminta ja datatyö, ovat vain ymmärtävinään toistensa työtä. Avoin kommunikaatiokulttuuri ja yhteisen sävelen löytäminen, jossa saa ja pitää uskaltaa

kysyä ilman negatiivista kommentointia, tulisi olla jokaisen nykyaikaisen yrityksen tavoitteena. Se mahdollistaa dynaamisten kyvykkyyksien kehittymisen ja myös datan paremman hyödyntämisen.

9. *Datatyöntekijöiden osaamisen tunnistaminen, toimenkuvan määrittäminen ja työn organisointi*

- Dataprosessit ovat hyvin moninaisia ja yritysten resurssit, kulttuuri ja toimintalogiikat vaikuttavat niihin. Samoin datatöiden parissa toimii paljon erilaisilla työnimikkeillä toimivia asiantuntijoita, jotka voivat tehdä paljolti samantyyllisiä töitä vaikka tittelit eroavat. Pääsääntöisesti *data arkkitehdit* suunnittelevat datakokonaisuuudet, *data-engineerit* rakentavat järjestelmien väliset integraatiot ja tiedon hallinnan kokonaisuuudet ja *data scientistit* ja – *analyttikot* toimivat dataprosessin validoinnin, analysoinnin ja raportoinnin parissa, mutta organisaatiosta ja toimintakulttuurista riippuen toimenkuvat voivat olla laajempia tai suppeampia. Yritysten haaste on tunnistaa eri tehtäviin tarvittavaa osaamista ja organisoida tehtäviä olemassa olevilla resursseillaan.

10. *Osaamisvaje datatyön ymmärtäjistä, tekijöistä, analyttikoista ja "välikästä" datatiimin ja johdon välillä.*

- Resurssipula on tämänkin raportin eri kohdissa nostettu varsin useasti esille. Koska data on niin kokonaisvaltainen käsite ja sitä tulee kaikista yritysten toiminnoista tuhansin eri tavoin, yrityksillä on havaittavissa selkeä haaste datan hyödyntämisen laadukkaaseen toteuttamiseen, varsinkin osaamisresurssin näkökulmasta. Tämä on tietysti osittain yhteiskunnallinen haaste, johon voidaan alueellisesti vastata koulutuksella, mutta myös organisaatioiden sisäinen haaste, johon ne voivat itse vastata myös koulutuksella, rekrytoinnin kehittämällä, ulkoisia resursseja paremmin hyödyntämällä ja olemassa olevia resurssejaan paremmin tunnistamalla. On mahdollista, että organisaatioissa on hyödyntämätöntä resurssia asiaan liittyen, mutta tällaista resurssia ei ole kyetty tunnistamaan.

11. *Johtamisprosesseissa toivotaan selkeyttä:*

- Erityisesti datatyön tekijät nostivat esille omaa työtään hankaloittavina seikkoina:
  - tavoitteiden puuttumisen,
  - työnkulkujen hajanaisuuden,
  - työkalujen hyödyntämisen puutteet (järjestelmät ja ohjelmistot),
  - työajan riittämättömyyden,
  - aikataulujen puutteen,

- osaavien työntekijöiden puutteen, ja
- laaja-alaisen osaamisen hyödyntämisen.

Näiden voidaan tulkita edustavan organisaatioita koskevaa resurssipulaa, joka näkyy hajanaisina johtamiskäytänteinä, puutteellisena kommunikaationa ja toimintaprosessien ja työn organisoinnin hajanaisuutena.

Kaiken kaikkiaan Oulun alueen yritysten datakyvykkyyksien kehittäminen on varsin kokonaisvaltainen tehtävä, jossa tulee fokuksittua paitsi järjestelmäarkkitehtuurien ja dataratkaisuiden kehittämiseen ja integrointiin, myös henkilöstöresurssien kasvattamiseen ja datatyön tekijöiden kouluttamiseen. Selvitystyön otanta toteutettiin Oulun alueen yrityksissä, joissa maturiteetin oletettiin olevan keskiarvoa korkeammalla tasolla, joten on olemassa myös suuri joukko yrityksiä, jotka joiden datakyvykyys ei ole selvitystyön yritysten tasolla. Näissä yrityksissä, yleensä mikro- ja pienyrityksissä, haasteena ovat riittävän syvän osaamisen ja tarvittavien resurssien kehittäminen, joten hankkeen puitteissa, ja sen jälkeenkin, myös näiden yritysten tarpeet tulee huomioida ja niille tulee kehittää ja suunnata koulutustarjontaa. Raportin lopussa on suunnitelma selvitystyön pohjalta suunnitelluista koulutuksista erikokoisille yrityksille ja organisaation eri tasoille ja rooleille. Nämä koulutukset alkavat tammikuussa 2023 ja lisää koulutuksia aiheen tiimoilta tullaan järjestämään keväällä 2023.

Selvitystyön loppuraportin voi ladata [täältä](#).

**Miikka Blomster**, lehtori, digitaalinen markkinointi  
Oulun ammattikorkeakoulu