

Datanhallinnan ja laskennan kehittämisohjelma

datajalaskenta2021

Tutkimushallinnon tietovirrat ja tutkimustietovaranto
Johtoryhmän kokous 30.10.2017

Anu Nuutinen
johtava tiedeasiantuntija
Suomen Akatemia
anu.nuutinen@aka.fi

Taustaa: Datanhallintaan ja laskentaan vaikuttavia muutostrendejä

Suuritehoisen data-analyysin merkitys kasvaa

- Mittausdatan ja datalähteiden määrä kasvaa
- Tutkimuksessa sekä palvelu- ja tuotekehityksessä hyödynnetään dataa entistä enemmän
- Yhä monimutkaisemmat tutkimuskysymykset asettavat suurempia vaatimuksia data-analyysiympäristölle

Data-analytiikasta lisäarvoa kaikilla tieteenaloilla

- Tekoälyn sovellukset kuten koneoppiminen ja robotiikka
- Laskennallisten menetelmien soveltaminen ihmistieteissä
- Uudet läpimurrot perinteistä data-analyysiä harjoittavilla tieteenaloilla

Kasvava datamäärä vaatii kehittyvän laiteympäristön ja uusia menetelmiä

- Tutkimus- ja simulaatiodatan tuotanto kiihtyy
- Sensitiivisen datan ja henkilökeskeisen tiedon hyödyntäminen tutkimuksessa lisääntyy nopeasti
- Tutkimusaineistojen käyttövaltuuksien ja pääsyoikeuksien hallinnan rooli korostuu

Data-analytiikkaympäristön käyttäjäystävällisyys

- Pilvipalvelut
- Graafiset käyttöliittymät
- Sähköinen asiointi ja automatisointi vähentävät rutiinityötä ja byrokratiaa

Lähde: CSC:n tuottama materiaali Datanhallinnan ja laskennan kehittämisohjelman ohjausryhmälle, elokuu 2017.

Kehittämisohjelman tavoitteet 2017–2021

- Investoida datanhallinnan ja laskennan infrastruktuureihin ja niihin liittyviin palveluihin
 - Päivittää infrastruktuurit kansainvälisen yhteistyön varmistamalle tasolle
 - Parantaa tutkimuksen ja koulutuksen palveluita (ml. valtion tutkimuslaitosten huomiointi)
 - Toteuttaa hallitusohjelman digitalisaatiotavoitetta
- Huomioida alan eurooppalainen tutkimusinfrastruktuuripolitiikkakehitys
- Osaamistason nostamisesta ja kyvykkyyksien lisäämisestä hyötyvät myös julkinen hallinto ja elinkeinoelämä.

Kehittämisohjelman rahoitus

- Puoliväliriihessä 2017 päätös panostuksista, yhteensä 33 miljoonaa euroa
 - lisäksi kehittämisohjelman käynnistämiseen (ml. koordinaatio) rahoitusta vuoden 2017 budjetista 2 milj. euroa
- OKM sopinut Suomen Akatemian kanssa kehittämisohjelman ohjaamiseen ja koordinointiin tarvittavasta työpanoksesta
- Alkuvuonna 2018 ohjausryhmä linjaa 33 milj. euron käytöstä kehittämisohjelman toteuttamiseen. Tämän jälkeen OKM tekee CSC:n kanssa erillisen sopimuksen.
- Kehittämisohjelman yllä mainittujen panostusten lisäksi suunnataan vuosittaisessa OKM:n korkeakoulu- ja tiedepolitiikan osaston ja CSC:n välisessä sopimuksessa ostettavia palveluita tukemaan kehittämisohjelman toteuttamista. Tämän panostuksen vuosittainen volyymi on noin 10,5 milj. euroa.

Ohjausryhmän kokoonpano

- tiedeasiantuntija Erja Heikkinen, **OKM**, pj.
- erityisasiantuntija Teemupekka Virtanen, **STM**
- kaupallinen neuvos Antti Eskola, **TEM**
- viestintäneuvos Anne Miettinen, **LVM**
- professori Sampsa Hautaniemi, HY (**yliopistot**)
- professori Paula Eerola, HY (**yliopistot**)
- vararehtori Kirsi Viskari, SAIMIA (**ammattikorkeakoulut**)
- ohjelmapäällikkö Sirpa Thessler, LUKE (**tutkimuslaitokset**)
- pääjohtaja Heikki Mannila, **Suomen Akatemia**
- kehityspäällikkö Tiina Tanninen-Ahonen, **Innovaatorahoituskeskus Tekes**
- tutkimusjohtaja Esa Virtanen, ABB Oy (**yrietykset**)
- opetusneuvos Petteri Kauppinen, **OKM**

Asiantuntijajäsenet:

- opetusneuvos Juha Haataja, **OKM**
- ylitarkastaja Sami Niinimäki, **OKM**

Asiantuntijas sihteeri:

- johtava tiedeasiantuntija Anu Nuutinen, **SA**

Ohjausryhmän tehtävät

Huolehtii

- datanhallinnan ja tieteellisen laskennan osalta tutkimusinfrastruktuurin ja palvelujen laajasti osallistavien **yhteistyörakenteiden toteutumisesta**
 - Miten eri hallinnonalojen omat datanhallintaan ja tutkimusinfrastruktuureihin liittyvät ratkaisut tukevat kehittämissohjelman tavoitteita?
- tutkimusinfrastruktuurin **arvioinnista** ja **priorisoinnista**
- **yhteishankintojen koordinoinnista**

Aktivoi ja kartoittaa

- datanhallintaa tarvitsevaa ja tieteellistä laskentaa hyödyntävää **tutkimusta ja koulutusta** eri tieteenaloilla
 - Aktivoi menetelmien ja palveluiden käyttöä myös uusilla/aloittavilla tieteenaloilla
 - Edistää ministeriöiden yhteistyötä tutkimustoimijoiden ohjauksessa ja seuraa toimivaltaisten tahojen edistymistä ohjelman tavoitteiden saavuttamisessa

Varmistaa

- alan tutkimukseen ja koulutukseen liittyvien tulosten **laajan vaikuttavuuden yhteiskunnassa**
 - Massadata, omadata, robotiikka, tekoäly ym.
 - Edistää ministeriöiden yhteistyötä tutkimustoimijoiden ohjauksessa ja seuraa toimivaltaisten tahojen edistymistä ohjelman tavoitteiden saavuttamisessa

Keskeisiä datanhallintaan ja laskentaan liittyviä kansallisia toimijoita

Tutkimusinfrastruktuurikomitea

- Tiekartan väliarviointi käynnissä, valmistuu alkuvuonna 2018

Tekoälyryhmä

- Elinkeinoministerin asettama ohjausryhmä valmistelee ehdotusta Suomen tekoälyohjelmaksi.
- Tekoälytyön välitulokset julkistettiin 23.10.2017, ks. tekoalyaika.fi
- ”Tekoälyohjelman tavoitteena on nostaa tekoäly ja robotiikka suomalaisyritysten menestystekijäksi. Suomen tavoitteena on olla maailman paras tekoälyn soveltaja.”
- Loppuraportti valmistuu huhtikuussa 2019.

Tekoälyohjelmakokonaisuus valtioneuvoston selvitys- ja tutkimussuunnitelmassa 2018

- Hakuaika on 2.10.2017–9.11.2017

Tieteellisen laskennan yhteistyöfoorumi

- OKM:n asettama yhteistyöfoorumi arvioi tieteellisen laskennan tulevaisuuden kehitystä suomalaisten korkeakoulujen ja suomalaisen tutkimuskentän kannalta.
- Toimintakausi jatkuu 31.12.2018 saakka.

Data-asiain kansalliskomitea

- Tiedeakatemiain neuvottelukunnan asettaman komitean tehtävänä on muun muassa olla Suomen yhteistyöorganisaatio kansainvälisen CODATAN suuntaan (CODATA = the International Council for Science: Committee on Data for Science and Technology).

Käynnissä olevia kansallisia hankkeita

ATT-hanke, ajankohtaista

- EU:n tietosuoja-asetus ja sen vaikutukset -seminaari 31.10. HY, Fabianinkatu 33
- Tietosuoja ja tutkittavan suostumus osallistua tutkimukseen -seminaari 1.11., HY, Fabianinkatu 33
- Avoimen tieteen foorumi Helsingin Paasitornissa 16.11.2017, teemana on tiedon avoin saatavuus.

Tutkimushallinnon tietovirrat ja tutkimustietovaranto -hanke

- Tutkimusta kuvaavat tiedot ovat sujuvasti hyödynnettävissä erilaisissa palveluissa (sis. tutkimusaineistojen metatiedot).
- Tiedot syötetään järjestelmiin vain kerran.
- Tietovirtojen sujuvoittaminen edellyttää avoimia rajapintoja, semanttisista yhteentoimivuutta, yhteisiä tietomäärittelyksiä sekä valmiuksia tiedonsiirron automatisointiin.
- > Valtakunnallinen tutkimustietovaranto.
- Yhteismitalliset tiedot mahdollistavat näkymän Suomessa tehtävään tutkimukseen, mikä parantaa edellytyksiä tutkimustiedon käytölle ja tutkimuksen yhteiskunnalliselle vaikuttavuudelle.
- Aikataulu: 2017-2020

Eurooppalainen tutkimusinfrastruktuuripolitiikkakehitys

European Open Science Cloud (EOSC)

- Komission laatima EOSC-deklaraatio tukee monia samoja asioita, joita edistetään Datanhallinnan ja laskennan kehittämissuunnitelman kautta (FAIR-periaatteet, DMP, osaaminen ja koulutus, PAS ja HPC).
- Deklaraation pohjalta valmistellaan parhaillaan tiekarttaa (EOSC Implementation Roadmap, aikataulu 12/2017), jossa mm. hallintomalli ja rahoitus käsitellään laajemmin
 - Governance structure, incl. mandate & selection procedure for Executive Board
 - Broad (federated) architecture, incl. EOSC core services
 - Financing model
- EOSC pilot käynnissä

Euro HPC

- European cooperation framework on High Performance Computing

Go Fair -aloite

- Alankomaiden aloite
- Verkostomainen toimintatapa mm. FAIR-periaatteiden käyttöönoton edistämiseksi

FAIR data principles: findable, accessible, interoperable, reusable

DMP: Data Management Plan

PAS: Pitkäaikaissäilytys

HPC: High Performance Computing

Horisontti 2020: 2. Call for Implementing the European Open Science Cloud EOSC (187,5 M€)

Topics	Opening	Deadline	Indicative budget (M€)
INFRAEOSC-04-2018 (RIA) Connecting ESFRI infras through cluster projects	05.12.2017	22.03.2018	95.0 6-24 M€/project
INFRAEOSC-01-2018 (RIA) Access to commercial services through EOSC hub	05.12.2017	22.03.2018	12.0
INFRAEOSC-05-2018-2019 (CSA) Support to EOSC governance	10.01.2018	19.04.2018	20.0
INFRAEOSC-05-2018-2019 (RIA) Support to Governance	26.07.2018	21.11.2018	15.0 + 15.0 2019
INFRAEOSC-02-2019 (RIA) Prototyping innovative services	16.10.2018	29.01.2019	28.5
INFRAEOSC-06-2019-2020 (RIA) Enhancing the EOSC portal and connecting thematic clouds	14.11.2018	20.03.2019	2.0

For 2020, the indicative budget is 84 M€

Aikataulutus 09/2017–

- Ohjausryhmän I kokous 18.9.
- Sidosryhmätapaamiset syksyn aikana käynnissä
- Käynnissä selvitys tutkimuslaitosten datanhallinnan ja laskennan tarpeista ohjausryhmän käyttöön (Rehtori Turo Kilpeläinen, KAMK).
- **Tutkijoiden ja tutkimusorganisaatioiden työpaja 13.11.**
 - Tavoitteena on koota mahdollisimman konkreettinen tilannekuva datanhallinnan ja laskennan tutkimusinfrastruktuuriin ja palveluihin kohdistuvista tarpeista.
- Ohjausryhmän II kokous 11.1.2018
 - Ohjausryhmä linjaa 33 milj. euron käytöstä kehittämisohjelman toteuttamiseen.
 - Käsitellään tarkemmin eurooppalaisen tutkimusinfrastruktuuripolitiikan kehitystä.
- Hankinnan valmistelu ja kilpailutusprosessi käynnistyvät.
- Ohjausryhmän III kokous 16.4.2018
- Ohjausryhmän IV kokous 27.9.2018
- Kehittämisohjelman vuosiseminaari marraskuussa 2018 (alustava tieto)