

minedu.fi/DL2021

Datanhallinnan ja laskennan kehittämishjelman ohjausryhmä

Aika: torstaina 27.9.2018 klo 13–15

Paikka: opetus- ja kulttuuriministeriö, kh. Kalevala, Meritullinkatu 10

Ohjausryhmän kokoonpano ja varajäsenet

tiedeasiantuntija *Erja Heikkinen*, opetus- ja kulttuuriministeriö, pj.
erityisasiantuntija *Teemupekka Virtanen*, sosiaali- ja terveysministeriö (klo 13.30 alkaen)
kaupallinen neuvos *Antti Eskola*, työ- ja elinkeinoministeriö
viestintäneuvos *Anne-Miettinen*, liikenne- ja viestintäministeriö (ylitarkastaja *Tuomas-Kaivola*)
professori *Sampsa Hautaniemi*, Helsingin yliopisto
vararehtori *Kirsi Viskari*, Saimaan ammattikorkeakoulu (kokoustan etäyhteyttä ei saatu toimimaan)
ohjelmapäällikkö *Sirpa Thessler*, Luonnonvarakeskus
pääjohtaja *Heikki Mannila*, Suomen Akatemia
kehityspäällikkö *Tiina Tanninen-Ahonen*, Innovaatorahoituskeskus Business Finland
tutkimusjohtaja *Esa Virtanen*, ABB Oy
professori *Paula Eerola*, Helsingin yliopisto (professori *Petri Myllymäki*)
opetusneuvos *Petteri Kauppinen*, opetus- ja kulttuuriministeriö

Asiantuntijajäsenet

Opetusneuvos *Juha Haataja*, opetus- ja kulttuuriministeriö (este)
Ylitarkastaja *Sami Niinimäki*, opetus- ja kulttuuriministeriö

Johtava tiedeasiantuntija *Anu Nuutinen*, Suomen Akatemia (asiantuntijas sihteeri)

Vierailijat

Tuotepäällikkö *Tommi Berg*, Valtori (Läsnä kohdat 1–4.)
Johtaja *Pekka Lehtovuori*, CSC
Erityisasiantuntija *Aino Sipari*, LVM (Anne Miettisen sijaisena)

minedu.fi/DL2021

Asialista

1. Kokouksen avaus ja asialistan hyväksyminen

Puheenjohtaja avasi kokouksen ja kokouksen etukäteen toimitettu asialista hyväksyttiin.

2. Edellisen kokouksen muistio

Kokousmuistio 16.4.2018 kokouksesta hyväksyttiin sellaisenaan.

3. Valtori ja CSC: DL2021-toimintaan liittyvä kolmikantayhteistyö tutkimuslaitosten kanssa

Tommi Berg (Kapasiteettipalveluiden tuotepäällikkö, Valtori) ja Pekka Lehtovuori (johtaja, CSC) esittelivät Valtorin, CSC:n ja tutkimuslaitosten yhteistyön kehittämistä. Vuoden 2018 aikana on tunnistettu palveluvalikoimaa, jota CSC voisi tarjota Valtorin kautta tutkimuslaitoksille. Tämän hetken selvimpiä haasteita ovat tietoliikenneyhteyksien muodostaminen sekä tunnistautuminen. VY-verkon uuden ratkaisun (VY NG) odotetaan ratkaisevan tietoliikennekaistan mahdollisia rajoitteita datamäärien kasvaessa. Esitetään yhteistyöprojektin perustamista ratkaisemaan tunnistettuja haasteita.

Tutkimuslaitoksille on tärkeää kuulla hyvissä ajoin CSC:n strategisista suunnitelmista, jotta niihin voi varautua proaktiivisesti. Todettiin, että DL2021-ohjausryhmä on valittu kansallinen foorumi tutkimuslaitoskokonaisuudesta keskusteluun liittyen datanhallinnan ja laskennan infrastruktuureihin, palveluihin ja osaamiseen.

Valtori ja CSC pyrkivät ratkaisemaan kriittisimmät haasteet jo vuoden 2018 aikana, erityisesti verkko-ongelmat. **Sovittiin, että ohjausryhmässä palataan esiteltävään Valtorin ja CSC:n yhteistyöprojektiin noin puolen vuoden kuluttua.**

4. Ministeriöiden tutkimuslaitoksia koskeva tilannekuva edistymisestä DL2021-kehittämisohjelman tavoitteiden saavuttamisessa

Tilannekuvan olivat toimittaneet LVM, MMM, STM, TEM, YM sekä PLM. Ohjausryhmän ministeriöjäsenet tai heidän sijaisensa esittelivät avainkohtia oman ministeriönsä vastauksesta sekä Anu Nuutinen esitteli vastaukset niistä ministeriöistä, jotka eivät ole ohjausryhmässä edustettuina.

LVM toi vastauksessaan mm. esille, että CSC:n lisäksi myös useat markkinatoimijat kehittävät ja tarjoavat laskentakapasiteettia ja siihen liittyviä palveluita mm. dataintensiivisen talouden, tekoälyn ja tutkimuksen käyttöön. Kehittämisohjelmalla tulisi varmistaa, ettei julkisesti rahoitettu laskentakapasiteettipalvelujen tarjonta aiheuta markkinahäiriötä.

OKM:n on selvittänyt asian. Valtion hankintalain mukaisesti CSC tuottaa in-house -palveluja valtiolle (OKM, yliopistot, ammattikorkeakoulut, tutkimuslaitokset) ja tämän johdosta CSC voi myydä vain hyvin rajoitetusti palveluita ulkopuolisille. Rajana on 5 % liikevaihdosta tai 500 000 euroa, eli määrä on vähäinen, joten markkinahäiriötä ei aiheuteta. Puheenjohtaja totesi lisäksi, että kokonaisuutena tällaista palvelua ei ole markkinoilta saatavilla, koska CSC tarjoaa DL-infrastruktuuriin liittyen myös käyttäjätuen.

minedu.fi/DL2021

Keskusteltiin substanssituesta eli mitä substanssiosaamista pitää olla tutkimusorganisaatioilla itsellään ja mitä CSC voi realistisesti tarjota esimerkiksi bioinformatiikkaan liittyen. Tutkimuslaitoskentällä tunnistettiin myös epäselvyyttä datanhallinnan palveluissa (FAIR DATA). Ilmatieteen laitoksessa ja Luonnonvarakeskuksessa on meneillään oma kapasiteettihankinta.

Sovittiin, että Tulanetissä tiivistetään tutkimuslaitosten tutkimusjohtajien kanssa vuorovaikutusta DL-ohjausryhmän kokousten asioista. Sovittiin, että kokousmateriaaleja saa jakaa eteenpäin omalle sidosryhmälle.

Tilannekuvatietopyynnön keskeisenä tavoitteena oli selvittää, miten ministeriöt ovat huomioineet tutkimuslaitosten kyvykkyyden hyödyntää DL2021-tutkimusinfrastruktuuria ja palveluita. Joidenkin ministeriöiden osalta todettiin, että toimet tutkimuslaitosten tukemiseksi eivät vielä näy tulosohjausprosessissa. **Sovittiin, että palataan tulevaisuuden kokouksissa vielä uudestaan siihen, millaisia kehittämistoimia ministeriöt ovat käynnistäneet tutkimuslaitosten tukemiseksi.**

5. CSC:n DL2021-tilannetasaus: hankintaprosessin eteneminen, pilotit ja kompetenssikehitys

Pekka Lehtovuori (johtaja, CSC) kävi läpi DL2021-ympäristön ja kompetenssin kehittämistoimia, tutkimuslaitospilottien tilannetta, tehokkaan käyttöönoton varmistamiseen liittyviä näkökohtia sekä DL2021-infrastruktuurihankinnan aikataulua.

Keskusteltiin siitä, että suomalaisen teollisuuden kannalta on tärkeää, että kapasiteettia ja osaamista rakennetaan. On tärkeää, että yliopistoilla on käytössään parhaat laitteet, jotta yliopistosta voidaan tarjota myös teollisuudelle palveluja. Korostettiin myös osaamisen keskeisyyttä ja nähtiin tärkeänä datan ympärille rakentuvat ekosysteemit.

Sovittiin, että tutkimuslaitoksille toimitetaan lista siitä, mitkä pilotit on nyt valittu ja kysytään vielä Tulanetin kautta, onko pilotteihin muita halukkaita osallistujia kuin ensimmäisessä vaiheessa ilmoittautuneet organisaatiot.

Keskusteltiin myös lyhyesti kehittämisohjelman vaikuttavuuden mittaamisesta sekä viestinnästä. Todettiin, että mittaristoa on tärkeää miettiä jo nyt. Viestinnässä on tarpeellista ottaa huomioon myös se, että DL2021-kehittämisohjelma on osa suomalaisen laskennallisen tieteen kokonaisuutta. Jotta kokonaiskuva hahmottuu, on tärkeää luoda kuva myös tutkimusorganisaatioiden omista ja rahoittajien toimista.

Viestintäsuunnitelmaan toivottiin lisättäväksi myös datanhallintapuolta. Tutkimuslaitokset myös toivoivat, että viestinnässä nostetaan esille laskennallisen tieteen saavutuksia tutkimuslaitoksista. Pekka Lehtovuori myös korosti, että poliittisille päättäjille on tärkeää viestiä DL-infrastruktuurin elinkaaren vuoksi uusiutuvista rahoitustarpeista.

Kehittämisohjelman vaikuttavuuden seurannan mittaristoon ja viestintään liittyvät toiveet ja kommentit pyydettiin toimittamaan ohjausryhmän asiantuntijasihteerille.

minedu.fi/DL2021

6. Tieteellisen laskennan yhteistyöfoorumin laatima ajankohtaiskatsaus datanhallinnan ja laskennan tutkimusinfrastruktuureiden, palveluiden ja osaamisen kehitykseen

Yhteistyöfoorumin puheenjohtaja, Petri Myllymäki (professori, HY) esitteli Tieteellisen laskennan yhteistyöfoorumin (YTF) kommentteja DL2021-kokonaisuuteen. CSC sai paljon positiivista palautetta.

Data-intensiivinen tutkimus ja tekoäly luovat uusia tarpeita, joihin CSC ei voi nykyresursseilla vastata vaarantamatta perinteistä laskentaa. Jollei korvamerkittyä lisärahoitusta keskitetysti hallinnoituun ratkaisuun esim. FIRIn kautta ole mahdollista saada, voidaan CSC:n palveluita täydentää yksiköiden omalla kapasiteetilla, mutta tällöin yksiköissä itsessään pitää investoida myös näihin asioihin, ja ne pitää tunnistaa strategisesti kriittisiksi resursseiksi. Tarvitaan myös kansallista yhteistyötä esim. hyvien käytäntöjen jakamiseksi (YTF:n rooli jatkossa?).

Yhä suurempi osa maailman datasta syntyy muualla yhteiskunnassa kuin tutkimusorganisaatioissa. Datan hallinnasta ja säilytyksestä aiheutuvat kustannukset pitää ratkaista.

Myllymäki painotti esityksensä lopuksi, että käyttäjäkunnan yksiköissä pitää olla tekninen perustaso (tutkimuksen lähituki) toteutettuna omasta takaa, erityisesti nyt kun dataintensiivinen tutkimus leviää yhä uusille tieteenaloille, joista moni ei perinteisesti ole ollut CSC:n asiakas. Tutkimuksen tuki ei voi olla ainoastaan CSC:n vastuulla, kun käyttäjäkunta moninkertaistuu.

Keskustelussa todettiin, että dataintensiivisessä tutkimuksessa on lopulta enemmän kysymys ihmisistä kuin rahoituksesta laitehankintoihin. Datan pitkäaikaissaavutettavuus on keskeistä, mikä on olennaista esim. tutkimuslaitosten näkökulmasta. Puheenjohtaja korosti, että DL-ympäristö ei ole datan pitkäaikaissäilytyksen ratkaisu.

Tärkeää on myös miettiä datan käytettävyyttä yhdessä yritysten kanssa. Korostettiin myös data-analyysin merkitystä osana kaikkien pääaineiden opintoja.

YTF:n nykyisen toimintakauden päättyessä pohdittiin, että kaikilla tutkimusorganisaatioilla pitäisi olla nimetty henkilö, jolla on vastuu data- ja laskentaintensiivisestä tutkimuksesta. Näiden tahojen on tärkeää verkostoitua ja CSC:n tulee olla tässä yhteistyössä tiiviisti mukana. Jokaiseen tutkimusorganisaatioon tarvittaisiin vähintään yksi tutkimuksen IT-lähituen erityisasiantuntija.

7. Muut asiat

- EURO-HPC:n ja EOSC:n valmistelun tilannekatsaus siirrettiin seuraavan kokoukseen.
- Vuosiseminaari järjestetään 22.11.2018 klo 12-15 Kansallismuseon auditoriossa Helsingissä, ohjelma on valmistunut.

8. Seuraava kokous 12.12. klo 10-12