

minedu.fi/DL2021

Datanhallinnan ja laskennan kehittämishjelman ohjausryhmä

Aika: keskiviikkona 12.12.2018 klo 10–12

Paikka: CSC, Life Science Center Keilaniemi, Keilaranta 14, Espoo

Ohjausryhmän kokoonpano ja varajäsenet

tiedeasiantuntija *Erja Heikkinen*, opetus- ja kulttuuriministeriö, pj.
erityisasiantuntija ~~*Teemupekka Virtanen*~~, sosiaali- ja terveysministeriö
kaupallinen neuvos *Antti Eskola*, työ- ja elinkeinoministeriö
viestintäneuvos *Anne Miettinen*, liikenne- ja viestintäministeriö
professori ~~*Sampsa Hautaniemi*~~, Helsingin yliopisto (dosentti ~~*Susanna Pirttikangas*~~, Oulun yliopisto)
vararehtori *Kirsi Viskari*, Saimaan ammattikorkeakoulu
ohjelmapäällikkö ~~*Sirpa Thessler*~~, Luonnonvarakeskus (tietohallintojohtaja ~~*Jukka Santala*~~, Suomen ympäristökeskus)
pääjohtaja *Heikki Mannila*, Suomen Akatemia
Head of AI Business Outi Keski-Äijö, Business Finland Oy
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
Ylitarkastaja *Sami Niinimäki*, opetus- ja kulttuuriministeriö

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

Vierailijat

Varatoimitusjohtaja *Tiina Kupila-Rantala*, CSC
Johtaja *Pekka Lehtovuori*, CSC
Senior Applications Specialist *Johanna Törnroos*, CSC

minedu.fi/DL2021

Asialista

1. Kokouksen avaus ja asialistan hyväksyminen

Puheenjohtaja avasi kokouksen CSC:n ja Atosin sopimuksen allekirjoitustilaisuuden jälkeen.

2. Edellisen kokouksen muistio

Kokousmuistio 27.9.2018 pidetystä kokouksesta hyväksyttiin.

3. Ajankohtaiskatsaus DL-osaamisen kehittämiseen ja eri toimijoiden rooleihin

Tiina Kupila-Rantala esitteli ajankohtaiskatsauksen osaamisen kehittämiseen ja eri toimijoiden rooleihin. Hän korosti vahvan dialogin merkitystä. Tutkimusorganisaation omat, paikalliset tukihenkilöt pitää löytyä ennen kuin paikallista tutkimuksen tukea voidaan tarjota CSC:ltä. Oulun yliopiston ja Itä-Suomen yliopiston kampuksista vielä keskustellaan mahdollisina on site -tuen paikkoina.

Puheenjohtaja korosti, että laiteinvestointi ei yksin vie Suomea kärkeen, vaan osaamisen vahvistamiseen pitää panostaa. Käyttäjäorganisaatioilla pitää olla omat resurssit kunnossa hyödyntää DL2021-infrastruktuuria. Jotkut tutkimuslaitokset ovat ostaneet CSC:ltä lisää dedikoitua työaikaa oman laskennan tai datanhallinnan haasteensa ratkaisemiseen.

Keskusteltiin *paikallisesta tutkimuksen IT-tuesta*. ”Datainsinööri” ei ole sinänsä uusi keksintö (vrt. museossa kuraattorit). Bioinformatiikkapuolella tällaisia henkilöitä on ollut jo pidempään. Dataa pitäisi ajatella samalla tavalla kuin kirjastoa (vrt. tiedon löydettävyyys).

Useimmilla yliopistoilla ei ole tällä hetkellä tutkimuksen IT-tuen yksikköä. Jatkossa jokainen tieteenala kuitenkin tarvitsee ”datalaboratorion”. Tutkimusorganisaatioiden on erittäin tärkeää tunnistaa tarve datainsinöörien erityisosaamiselle. Toimenkuvia on erilaisia: riippuu alasta sekä tutkimusryhmien ”kypsydestä”. Tutkimusorganisaatiossa on tärkeää olla tiedossa, mitä palveluja on käytettävissä omassa organisaatiossa, kansallisesti ja kansainvälisesti. Palvelujen tunnettuus on keskeistä. Datainsinöörien palkkaamiseen ei ole tulossa uutta rahoitusta. ***Sihteeristö selvittää mahdollisia kansainvälisiä verrokkeja eri yliopistoista (vrt. esim. yrityksissä chief data officer).***

Todettiin, että CSC on saanut 2 miljoonaa DL2021-kehittämishjelmaan kytkeytyvää rahoitusta, jolla CSC kehittää omaa dataosaamistaan ja auttaa tutkimusorganisaatioita kehittämään omaa osaamistaan.

Todettiin myös, että ***Tieteellisen laskennan yhteistyöfoorumin uudistuva rooli*** on valmistelussa OKM:ssä. Jokaisella korkeakouluilla tulisi olla vähintään yksi henkilö, jonka tehtävänä on tukea tutkimusta IT-asioissa. Data on keskeisessä asemassa.

4. Keskustelu DL2021-kehittämishjelman tavoitteiden toteutumisen ja vaikuttavuuden seurannasta

Johanna Törnroos esitteli CSC:n ehdotuksen DL2021-kehittämishjelman tavoitteiden toteutumisen ja vaikuttavuuden seurantaan. Seurannassa tullaan hyödyntämään myös tutkimustietovarantoa.

minedu.fi/DL2021

Asiakkaiden saaman tutkimusrahoituksen seuranta pystytään jatkossa tekemään systemaattisemmin. Laadullisten mittareiden avulla tarkastellaan, mikä merkitys DL2021-resurssilla on ollut menetelmien kehitykselle, opetukselle ja mahdollisille läpimurroille.

Ohjausryhmä kommentoi, että seurantaan olisi tärkeää lisätä tutkijan näkökulma: kuinka helposti data on saatavilla ja käytettävissä tutkijan näkökulmasta. Pekka Lehtovuori kertoi, että suunnitelmassa on pyritty tekemään sellaisia mittareita, joita voidaan automatisoida. Seuranta pyritään tekemään tutkijoille helpoksi.

Tutkimuslaitosten avaamia aineistoja seurataan (uudelleen käyttö ja ristiinkäyttö). Nyt tutkimuslaitoskäyttäjää on alle 200, onko vuoden päästä suurempi käyttäjämäärä tutkimuslaitoksista. Ohjausryhmä ehdotti, että asiakastyytyvää mitattaisiin tutkimuslaitoksista. Tätä jo tehdään, ja se liitetään osaksi seurantaan. Lisäksi ehdotettiin, että mitattaisiin myös tutkimuslaitosten omia rahapanostuksia laskennan infrastruktuureihin ja palveluihin. Onko tutkimusorganisaatioissa havaittavissa muutoksia?

Ohjausryhmä myös totesi, että **osaamisen kehitystä** voidaan seurata jo nyt, koska se ei ole kiinni infrastruktuurin asennuksista.

Ohjausryhmä toivoi tiivistä raportointia kehittämisohjelman tavoitteiden toteutumisen ja vaikuttavuuden seurannan tuloksista ja toivoi lisäämään suunnitelmaan yhden esittelykerran lisää vuoden 2020 ajalle kyselyjen ja haastattelujen välille. Ohjausryhmä myös kysyi, onko kaikkiin mitattaviin asioihin saatavilla lähtötilanteen tiedot?

CSC:n asiakastyytyvyisyys on ollut perinteisesti korkea. Seuranta on erityisen tärkeää nyt uudessa tilanteessa, kun tulee paljon uusia asiakkaita. Asiakaskunnasta tulee todennäköisesti aiempaa heterogeenisempi, jolloin tilanteen laadullinen seuraaminen on tärkeää. Mittareihin toivottiin erottelua siitä, miten palvelut toimivat toisaalta vanhoille ja toisaalta uusille asiakkaille. Todettiin, että kiinnostavaa olisi koota tietoa myös entisiltä asiakkailta sekä tehdä kysely myös muille kuin CSC:n uusille tai nykyisille asiakkaille. Miten laskennallisen tieteen tutkijat ylipäättään näkevät CSC:n palvelut?

Ohjausryhmä piti myös hyvänä, että katsotaan **yrityskäytön** edistämistä, mm. milloin yritykset tulivat käyttäjiksi. Pekka Lehtovuori vastasi tähän, että CSC ei lähtökohtaisesti hae yritysasiakkaita, vaan yritysten kanssa toimitaan ensisijaisesti tutkimusyhteistyön kautta (CSC–yritys–tutkimusorganisaatio). Kiinnostavaa olisi kuitenkin seurata, paljonko on uusia yritysikäisiä, vaikka yhteistyö toteutuisikin kolmikannan kautta. Tutkijoilta tullaan kysymään, miten esim. tutkimuslaitosten yritys yhteistyö on kehittynyt.

Ohjausryhmässä todettiin, että suuryritykset eivät ole ensisijaisesti kiinnostuneita tällaisesta palvelusta. Osaamisen ja teknologisen tiedon siirtäminen yrityksiin on sen sijaan tärkeää. Tällöin edistetään datan ja laskennan käyttöä suomalaisessa tki-järjestelmässä ja kehitetään osaamista yhdessä. Yrityskäytön avainkohderyhmänä ovat pk-yritykset, joille yhteistyö on erityisen tärkeää hankkeiden tutkimuksellisessa alkuvaiheessa. ***Sihteeristö selvittää, millaista aineistoa Business Finlandin kautta olisi saatavilla, ja tähän asiaan palataan tulevissa kokouksissa.***

minedu.fi/DL2021

Euroopassa tekoälyn kehittämiselle kriittinen kysymys on siinä, miten isoilla yrityksillä olevien arvokkaita aineistojen luovuttaminen tutkimuksen käyttöön tapahtuisi. Voisiko olla tarjolla jokin neutraali paikka, johon anonymisoitu data tuodaan? Tällaisessa ympäristössä tekoälytutkijat voisivat testata uusia algoritmeja. Tämä olisi uniikkia globaalisti. ***Sihteeristö selvittää tähän liittyen ISAACUS-mallia ja mitä siitä voitaisiin oppia.***

Tiina Kupila-Rantala painotti, että CSC:n palvelun erityisyytenä on se, että laskentainfrastruktuurin lisäksi tutkijoille tarjotaan koko palvelupaketti, esimerkiksi ohjelmointi ja parallelisointiosaamista, mikä erottaa CSC:n kaupallisista toimijoista.

5. EuroHPC:n ja EOSC:n valmistelun tilannekatsaus

Erja Heikkinen kertoi **EuroHPC:n** ja EOSC:n valmistelusta. EuroHPC:hen liittyen todettiin, että Suomesta puuttuu yhteiseurooppalainen tutkimusinfrastruktuuri. EuroHPC vahvistaa innovaatiotoimintaympäristöä Suomessa. Ohjausryhmässä myös ehdotettiin, ettei kategorisesti suljettaisi EU:n ulkopuolisia käyttäjiä pois, vaan osallistuminen olisi mahdollista sopimusjärjestelyillä.

EOSC kytkeytyy avoimen tieteen edistämiseen (avoimeen dataan). Todettiin, että Suomi on miettinyt moniin muihin Euroopan maihin verrattuna pitkäjänteisemmin datanhallintaa, kun täällä painotetaan myös tutkimusdatan pitkäaikaissaatavuutta. Suurimmat ongelmat liittyvät kuitenkin pieniin datoihin. Onkin tärkeää korostaa osaamista ja ns. datainsinöörien tarjoamaa tukea. Ohjausryhmä painotti, että EOSC:n valmistelussa ei pidä mennä laitteet edellä, vaan osaaminen ja avoin data edellä.

EOSC-hubissa (yksi EOSCia valmistelevista projekteista) rakennetaan työkaluja EOSC-portaaliin: tuodaan avoimen tieteen palveluja yhteen portaaliin, joka ei ole varsinaisesti mikään pilvi, vaan palveluiden markkinapalkka, josta on helppo lähestyä myös muita palveluja. CSC on mukana mm. EOSC-hubissa ja tekee EOSC-hubin strategiaa.

6. Muut asiat

- Palaute DL2021-vuosiseminaarista 22.11.2018. Seminaarin esitykset ja tallenne: <https://minedu.fi/tapahtumat/2018-11-22/dl2021-kehittamisohjelman-vuosiseminaari>
- Hallitus hyväksyi 5.12. Eettistä tietopolitiikkaa tekoälyn aikakaudella -selonteon lähetettäväksi eduskunnalle. Tiedote ja selonteko (PDF) löytyvät täältä: https://vm.fi/artikkeli/-/asset_publisher/tietopolitiikasta-uusi-politikka-alue-suomeen

7. Seuraavat kokoukset ja alustava listaus käsiteltävistä asioista

- **25.3.2019**
 - o CSC–Valtori–tutkimuslaitokset -yhteistyön kehitystilanne
 - o Tutkimuslaitosten konkreettiset valmiudet DL-infrastruktuurin hyödyntämiseen
 - o Tutkimuksen paikallisen IT-tuen vahvistaminen tutkimusorganisaatioissa
- **15.5.2019**
 - o DL2021-kehittämisohjelmaan linkittyvä lainsäädäntö (Tietosuoja laki, Tiedonhallintalaki, Laki CSC-Tieteen tietotekniikan keskus Oy -nimisestä osakeyhtiöstä)
 - o Tutkimustietovarannon valmistelun eteneminen ja linkittyminen DL2021-kehittämisohjelmaan