



Tutkimuksen digitalisaation politiikkatoimien kokoamisen ja kansallisen datanhallinnan ja laskennan tutkimusinfrastruktuuriympäristön päivittämisen (DL2026) seurantaryhmä

Seurantaryhmän kokous I

Aika: keskiviikko 20.3.2024 klo 10.00–11.30

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

Osallistujat:

johtaja Erja Heikkinen, opetus- ja kulttuuriministeriö, puheenjohtaja
tutkimusjohtaja Annukka Pakarinen, ammattikorkeakoulujen rehtorineuvosto Arene ry (Hämeen ammattikorkeakoulu)
yksikön johtaja Niko-Matti Ronikonmäki, liikenne- ja viestintäministeriö
erityisasiantuntija Tuire Valkonen, liikenne- ja viestintäministeriö
johtava tiedeasiantuntija Anu Nuutinen, Suomen Akatemia
tutkimuspalveluiden johtaja Ella Bingham, Suomen yliopistojen rehtorineuvosto Unifi ry (Aalto-yliopisto)
professori Minna Palmroth, Suomen Tiedeakatemia (Helsingin yliopisto)
johtaja Pekka Lehtovuori, Tieteen tietotekniikan keskus CSC
yksikönjohtaja Saku Anttila, Tutkimuslaitosten yhteenliittymä Tulanet (Suomen ympäristökeskus) (este)
opetusneuvos Petteri Kauppinen, opetus- ja kulttuuriministeriö

Pysyvät asiantuntijat

professori Heikki Mannila, Aalto-yliopisto

CSC:n hallituksen puheenjohtaja Jukka Mönkkönen, Itä-Suomen yliopisto

opetusneuvos Laura Taajamaa, sihteeri

1. Kokouksen avaus ja asialistan hyväksyminen

Puheenjohtaja avasi kokouksen kello 10.02 ja toivotti jäsenet lämpimästi tervetulleeksi ensimmäiseen kokoukseen. Kokouksen asialista hyväksyttiin.

2. Seurantaryhmän järjestäytyminen ja esittelykierros

Seurantaryhmän puheenjohtaja on johtaja Erja Heikkinen, varapuheenjohtaja opetusneuvos Petteri Kauppinen ja sihteeri opetusneuvos Laura Taajamaa opetus- ja kulttuuriministeriön korkeakoulu- ja tiedepolitiikan osastolta, tiedepolitiikan vastuualueelta. Jäsenet esittelivät itsensä ja kertoivat omasta taustastaan aihepiirin tiimoilta.

3. Seurantaryhmän esittely (Liite 1. Seurantaryhmän esittely)

Laura Taajamaa kävi läpi seurantaryhmän käytänteitä. Ryhmän toimikausi on 1.3.2024–31.12.2026. Ryhmä kokoontuu neljä kertaa vuodessa ja lisäksi ensi vuodesta alkaen järjestetään vuosiseminaari. Materiaalit kootaan Tiimeriin, jonne on lähetetty linkki sähköpostitse. Ehdotettiin, että jäsenet vuorollaan alustavat ryhmän keskustelua yhdessä valituista teemoista.

Ryhmän tehtävinä on seurata niin datanhallinnan ja tieteellisen laskennan (ml. tekoälyn) kansallisen tutkimusinfrastruktuurin ja palvelujen päivittämistä (DL2026-hanke) kuin

datanhallinnan ja tieteellisen laskennan eurooppalaista kehitystä. Ryhmä seuraa ja analysoi tutkimuksen digitalisaation vaikutuksia, esittää opetus- ja kulttuuriministeriölle tutkimuksen digitalisaatioon liittyviä toimintamalleja tutkimusedellytysten parantamiseksi sekä ratkaisuja osaajien saatavuuden varmistamiseksi ja kuulee tarvittaessa asiantuntijoita. Käytiin vielä tarkentavaa keskustelua. Ryhmä keskittyy puimaan digitalisaation merkitystä ja vaikutuksia tutkimukselle ja tutkijoille, eri organisaatiolle, hallinnolle ja ohjaukselle. Keskustelussa todettiin, että tieteellisen laskennan yhteistyöfoorumi vastaa kokonaisarkkitehtuuryöstä ja toimii käyttäjien edustajana. Todettiin myös, että Tieteen tietotekniikan keskus CSC:n ohjaukseen on erikseen omat mekanisminsa.

4. Keskustelu tutkimuksen digitalisaatiosta (Liite. HM_ alustus)

Heikki Mannila alusti keskustelua tutkimuksen digitalisaatiosta. Tutkimuksen muutokselle on monia ajureita, tutkimuksen digitalisaation ohella myös esimerkiksi ilmastonmuutos, luontokato, muuttuvat talous ja geopolitiikka vaikuttavat. Muutosta on hyvä hahmottaa myös kokonaisuutena. Tutkimuksen digitalisaatio vaikuttaa kaikilla tieteenaloilla ja koko innovaatiotoiminta-ympäristöön. Tutkimuksen digitalisaatio muuttaa miten, mutta myös mitä tutkitaan ja miten se vaikuttaa yhteiskuntaan. Muutokset tieteen kentällä etenevät usein hitaasti, mutta tekoälyn kehitys viime vuosina on edennyt harppauksin. Muutosta on myös osittain vaikeasti ennustettavissa. Ilmiötä voi lähestyä teknologialähtöisesti, mutta voi pohtia tarvelähtöisesti eli mitä pitäisi saada tehdyksi.

Politiikkatoimia mietittäessä voi ammentaa aiemmista yhteistyössä tehdyistä toimista mm. tutkijakoulutuksen, tutkimusinfrastruktuurien tai avoimen tieteen kehittämisessä. Sisällöllisiä esimerkkejä löytyy esimerkiksi ilmastonmuutostutkimuksen tai bioinformatiikan saralta. Näistä on opittu, että resursseja kannattaa koota yhteen asioiden nopeuttamiseksi. On hyvä keskustella siitä, miten eri toimijoiden tulisi varautua ja erityisesti siitä, mitä tehdään yhdessä, mitä tieteenalakohtaisesti ja mitä tehdään yleisesti. Datan merkityksen kasvu on oleellinen ja näkyy kaikilla aloilla. Datan muodot ja roolit tutkimuksessa vaihtelevat pysyvistä aineistoista (huomioitava pitkäaikaissaavutettavuus) tiettyä hanketta varten mitattuun dataan.

Konkreettisen DL2026-hankinnan lisäksi myös tekoälyn käytön ohjeistuksia tarvitaan ja niitä on tehtykin eri tahoilla. Miten muutoin valmistaudumme ja reagoimme muuttuvaan tilanteeseen? Voidaan seurata rauhassa keskijoukossa tai sitten voidaan tavoitella joillakin alueilla kärkijoukkoa. Tällöin tarvitaan uusia toimintatapoja. Kaikessa ei kuitenkaan voi olla etummaisena: valintoja täytyy tehdä. Loppuun Mannila ehdotti, että teemaa voitaisiin lähestyä esimerkkien kautta, muutosten kautta katsotaan kokonaisuutta ja myös sisältökysymyksiä.

Puheenjohtaja kiitti Mannilaa ajatuksia herättelevästä esityksestä ja avasi keskustelun. Todettiin, että bioinformatiikka on hyvä esimerkki: CSC:llä oli pitkään bioinformatiikkaryhmä, joka järjesti täysiä kursseja tutkijoille, kunnes toiminta siirtyi yliopistoille. Nyt ollaan mukana tukemassa infrastruktuurin käyttöä. Digitalisaation murroksen ja käyttäjäkunnan laventumisen myötä tuki CSC:ltä ei voi olla niin vahvaa. On harkittava, miten varmistetaan osaaminen ja kyvykkyys tukemaan digitalisaatiota.

Esiin nousi kaksi isoa trendiä: HPC:n kyvykkyyden kasvu eksaskaalaan ja tekoälyn esiinmarssi, nähdään aito muutos ja tekoäly siirtymässä tutkimusryhmien käytännön työhön tueksi. Isoilla yhtiöillä on merkittävä rooli ja resursseja tekoälyn kehityksessä. Kun halutaan laventaa HPC-käyttäjäkuntaa, on osaamisen kehittämiseen panostettava. Hyvänä esimerkkinä nostettiin Sveitsi. Ryhmästä nousi toive, että käsiteltäisiin myös datanhallinnan ja laskennan tutkimusinfrastruktuurien käyttäjätilastoja. Keskustelussa todettiin myös, että monialaisuuden huomioiminen on tärkeää. Esitettiin vielä kiitokset esityksestä ja sovittiin, että kalvot voidaan jakaa myös ryhmän hankeikkunan kautta.

5. Datanhallinnan ja tieteellisen laskennan kansallisen ympäristön ja palvelujen päivitys (30 min) (Liite 2. DL2026-tilannekatsaus)

Pekka Lehtovuori esitti datanhallinnan ja laskennan kansallisen ympäristön ja palvelujen päivityksen tilanteen. Kansallinen ympäristö palvelee koko tutkimusyhteisöä, ja LUMI tukee huippututkimusta mm. lippulaivat sekä erityisesti tekoälytutkimusta ja yrityskäyttöä. CSC valmistellut suurteholaskennan tiekartan vuosille 2021–2026. Sen mukaan Puhdin ja Mahdin seuraaja, Roihu, on toiminnassa vuoden 2025 loppupuolella, nykyisen ympäristön elinkaari päättyessä. Uuden rahoitusmallin myötä hoidetaan päivitystä (2023 7,5 milj. euroa; 2024 12 milj. euroa ja 2025 15 milj. euroa). Investoinnit sisältävät FUNET-tutkimusverkon päivitystä miljoonalla eurolla. Käyttäjämäärien kasvu kaksinkertaistunut vuodesta 2020 vuoteen 2023 ja myös käyttötapausten kirjo on laajentunut. Datan ja erityisesti sensitiivisen datan merkitys nousee.

CSC on selvittänyt käyttäjien tarpeita vuoden 2023 aikana DL2021 resurssiallokaatio-pyyntöjen, käyttötilastojen, asiakastytyväisyys-, EasyRetro- sekä verkkokyselyn ja haastattelujen avulla. Näistä huomataan, että laskenta, muisti ja siihen liittyvä tallennustarve, automatisoidut prosessit, arkaluontoisen datan prosessointi, helppokäyttöisyys sekä erilaiset pilvipalvelut nousevat. Nähdään myös kasvavaa tarvetta edustapalveluille ja orkestrointipalveluille. Allas-objektitallennuspalvelua on tarkoitus jatkaa. Tarpeisiin vastataan hankinnan kautta. Lisäksi tavoitteena on saada Roihusta mahdollisimman energiatehokas.

Aikataulusta todettiin, että hankintapyyntö julkaistaneen tällä tai ensi viikolla. Vuoden 2024 aikana uudistetaan pilvipalveluita ja Allasta. Roihu-laitteisto rakennetaan vaiheittain: ensimmäinen vaihe vuoden 2025 loppussa (Puhti poistuu tuotannosta) ja toinen vaihe vuoden 2026 aikana (Mahti poistuu tuotannosta). Todettiin myös, että aikataulu tarkentuu datakeskuksen muutostöiden tarkentuessa. LUMI-konesalia hyödynnetään myös kansallisessa ympäristössä.

Puheenjohtaja kiitti Pekkaa esityksestä, joka selkiytti sopivalla tasolla hankinnan tilannetta. Seurantaryhmän on hyvä kuulla myös jatkossa hankinnan etenemisestä. Lisäksi datakeskuksen kehitys kiinnostavaa.

6. Seuraava kokous

Seurantaryhmän toinen kokous pidetään 27.5.2024 opetus- ja kulttuuriministeriössä etäosallistumismahdollisuudella. Materiaalit jaetaan Tiimerissä noin viikkoa ennen.

7. Kokouksen päättäminen

Puheenjohtaja muistutti vielä Suomen Akatemian Suurteholaskennan kansainvälinen yhteistyö erityisrahoitushausta, tohtoripilotin arviointihankkeen hankinnasta sekä lausuntopyynnöstä liittyen CSC:n infrastruktuurin yhteiskäyttöisyyteen.

Lopuksi puheenjohtaja kiitti osallistumisesta ja päätti kokouksen kello 11.34.