



Datanhallinnan ja laskennan kehittämishjelma

Tutkijoiden ja tutkimusorganisaatioiden työpaja

Aika: Maanantai 13.11.2017 klo 12.30–16.00

Paikka: Paasitorni, Paasivuorenkatu 5 A, Hakaniemi, Helsinki (Siltasaari-sali, 1. krs.)

Työpajan tavoitteena on koota mahdollisimman konkreettinen tilannekuva datanhallinnan ja laskennan tutkimusinfrastruktuuriin, palveluihin ja osaamiseen kohdistuvista tarpeista.

Ohjelma

12.30–12.45 Työpajan avaus: Kehittämishjelman tausta ja tavoitteet

Erja Heikkinen, tiedeasiantuntija, ryhmän päällikkö, opetus- ja kulttuuriministeriö

12.45–13.00 CSC:n kehittyvät palvelut tutkijoille tulevaisuudessa

Tiina Kupila-Rantala, varatoimitusjohtaja, CSC – Tieteen tietotekniikan keskus Oy

13.00–13.05 Pienryhmätyöskentelyn esittely

Anu Nuutinen, johtava tiedeasiantuntija, Suomen Akatemia

13.05–13.10 Siirtyminen pienryhmiin

13.10–14.10 Pienryhmätyöskentely

RYHMÄ A (Petri Myllymäki, pj.): pääsali Siltasaari 1 (1. krs.)

RYHMÄ B (Kai Nordlund, pj.): kokoushuone Karuselli (0 krs.)

RYHMÄ C (Hannele Korhonen, pj.): kokoushuone Johtokunta (4 ½ krs.)

RYHMÄ D (Hannu Salmi, pj.): kokoushuone 501 (5. krs.)

RYHMÄ E (Susanna Pirttikangas, pj.): pääsali Siltasaari 2 (1. krs.)

Jokaisella pienryhmällä on asiantuntijasihteeri, joka vastaa keskustelun raportoinnista.

14.10–14.30 Kahvia tarjolla Paasiravintolassa (2. krs.)

14.30–15.10 Pienryhmätyöskentely jatkuu samoissa kokoustiloissa

15.10–15.15 Siirtyminen pääsaliin

15.15–16.00 Yhteenveto työpajan annista

Pienryhmien puheenjohtajat: Keskeiset nostot jokaisesta ryhmäkeskustelusta (30 min)

Yleisökysymykset (10 min)

Mitä kehittämishjelmassa tapahtuu seuraavaksi? (5 min)

Pienryhmätyöskentelyn ryhmäjako ja kokoustilat

Osallistujat					
Ryhmä	Etunimi	Sukunimi	Tehtävänimike	Organisaatio	Ryhmätyötila
A	Petri	Myllmäki, pj.	Professori	Helsingin yliopisto	Siltasaari 1
A	Merja	Särkioja, rap.	Johtava tiedeasiantuntija	Suomen Akatemia	Siltasaari 1
A	Paavo-Petri	Ahonen	Opetusneuvos	Opetus- ja kulttuuriministeriö	Siltasaari 1
A	Aleksi	Kallio	Kehityspäällikkö	CSC	Siltasaari 1
A	Miki	Kallio	Palvelupäällikkö	Oulun yliopisto	Siltasaari 1
A	Asta	Laiho	Service manager	Bioteknikan keskus, Turun yliopisto	Siltasaari 1
A	Jussi	Paananen	Yliopistotutkija	Itä-Suomen yliopisto	Siltasaari 1
A	Jorma	Palvimo	Professori	UEF	Siltasaari 1
A	Mikko	Sillanpää	Professori	Matemaattisten tieteiden yksikkö, Aalto-yliopisto	Siltasaari 1
B	Kai	Nordlund, pj.	Professori	Helsingin yliopisto	Karuselli
B	Teppo	Häyrynen, rap.	Tiedeasiantuntija	Suomen Akatemia	Karuselli
B	Mikko	Alava	Professori	Aalto-yliopisto	Karuselli
			Manager in high performance computing		
B	Mikko	Hakala		Aalto-yliopisto	Karuselli
B	Karoliina	Honkala	Professori	Jyväskylän yliopisto	Karuselli
B	Petteri	Kauppinen	Opetusneuvos	Opetus- ja kulttuuriministeriö	Karuselli
B	Annakaisa	Korja	Tutkimusjohtaja	Seismologian instituutti, Helsingin yliopisto	Karuselli
B	Kari	Rummukainen	Professori	Helsingin Yliopisto	Karuselli
B	Antti	Snicker	tutkijatohtori	Aalto Yliopisto	Karuselli
B	Tanja	Tarvainen	Associate professor	Itä-Suomen yliopisto	Karuselli
			Department of Geosciences and Geography,		
B	David	Whipp	Associate Professor	Uni. Helsinki	Karuselli
C	Hannele	Korhonen, pj.	Tutkimusprofessori	Ilmatieteen laitos	Johtokunta
C	Sami	Niinimäki, rap.	Ylitarkastaja	Opetus- ja kulttuuriministeriö	Johtokunta
C	Erja	Heikkinen	Tiedeasiantuntija	Opetus- ja kulttuuriministeriö	Johtokunta
C	Seppo	Kaitala	Johtaja tutkija	Suomen ympäristökeskus	Johtokunta
C	Matti	Keränen	Yksikön päällikkö	Ilmatieteen laitos	Johtokunta
C	Turo	Kilpeläinen	rehtori	Kajaanin ammattikorkeakoulu	Johtokunta
C	Eevaliisa	Laine	Erikoistutkija	GTK	Johtokunta
C	Riku	Louhimo	tuotepäällikkö	Työterveyslaitos	Johtokunta
C	Seliina	Päälysaho	Tutkimuspäällikkö	Seinäjoen ammattikorkeakoulu	Johtokunta
C	Jukka	Santala	Johtaja (CIO), Tietokeskus	Suomen ympäristökeskus	Johtokunta
C	Ville	Savolainen	Asiakasratkaisupäällikkö	CSC	Johtokunta
C	Noora	Talsi	Tutkimusjohtaja	Kaakkois-Suomen ammattikorkeakoulu	Johtokunta
D	Hannu	Salmi, pj.	Akatemiaprofessori	Turun yliopisto	Kokoushuone 501
D	Risto	Vilkko, rap.	Ohjelmapäällikkö	Suomen Akatemia	Kokoushuone 501
D	Essi	Eerola	Tutkimusjohtaja	VATT	Kokoushuone 501
D	Anne	Ilvonen	tutkimus- ja kehittämisjohtaja	Karelia-ammattikorkeakoulu	Kokoushuone 501
D	Rami	Kivilahti	Tiimipäällikkö	Lahten ammattikorkeakoulu	Kokoushuone 501
D	Mari	Kleemola	Kehittämispäällikkö	Yhteiskuntatieteellinen tietokirjasto	Kokoushuone 501
D	Tiina	Kupila-Rantala	Varatoimitusjohtaja	CSC - Tieteen tietotekniikan keskus Oy	Kokoushuone 501
D	Krister	Lindén	Tutkimusjohtaja	Helsingin yliopisto	Kokoushuone 501
D	Anu	Nuutinen	Johtava tiedeasiantuntija	Suomen Akatemia	Kokoushuone 501
			Yksikön päällikkö,		
D	Mari	Riipinen	Tutkimusedellytykset	Turun yliopisto	Kokoushuone 501
D	Kirsi	Viskari	Toimialajohtaja	Saimaan ammattikorkeakoulu	Kokoushuone 501
D	Michael	von Boguslawski	Tutkimusasiamies	Arcada	Kokoushuone 501
E	Susanna	Pirttikangas, pj.	yliopistotutkija	Oulun yliopisto	Siltasaari 2
E	Juha	Haataja, rap.	Opetusneuvos	Opetus- ja kulttuuriministeriö	Siltasaari 2
			Head of technology services @		
E	Minna	Harjuniemi	IT center	Helsingin yliopisto	Siltasaari 2
E	Keijo	Heljanko	Associate professor	Aalto-yliopisto	Siltasaari 2
E	Juha	Herrala	Computing Specialist	Tampereen teknillinen yliopisto	Siltasaari 2
			IT-palvelupäällikkö, Tutkimuksen		
E	Meri-Tuulia	Karakainen	IT	Turun yliopisto	Siltasaari 2
E	Aarne	Klemetti	Tutkijaopettaja	Metropolia UAS	Siltasaari 2
E	Heikki	Mannila	Pääjohtaja	Suomen Akatemia	Siltasaari 2
E	Teemu	Mäenpää	Tutkijatohtori	Vaasan yliopisto	Siltasaari 2
E	Pekka	Pöyry	Tieto- ja viestintätekniikan lehtori	TAMK	Siltasaari 2
	Pirkko	Muurinen	Hallinnollinen avustaja	OKM	

Pienryhmissä käsiteltävät teemat

Työskentelyn tavoite: Pienryhmää pyydetään tuottamaan mahdollisimman konkreettinen tilannekuva datanhallinnan ja laskennan tutkimusinfrastruktuuriin, palveluihin ja osaamiseen kohdistuvista tarpeista edustamallaan aloilla / tutkimusorganisaatioissa.

- Palvelukokonaisuuden linjaamiseen ja tarpeiden priorisointiin keskeisesti vaikuttavat asiat
- Jokaista pienryhmää pyydetään käsittelemään kaikkia 5 teemaa, mutta ryhmä voi painottaa niiden sisällä keskeisiä näkökulmia kokoonpanonsa mukaan.

Sessio 1: Datanhallinnan ja laskennan tutkimusinfrastruktuuri, palvelut ja osaaminen

1. Datanhallinnan ja laskennan tutkimusinfrastruktuuri

- Dataintensiivisen laskennan erityisvaatimukset infrastruktuureille
- Datan analysointia, hallintaa, ja hyödyntämistä tukevat alustat
- Big data -ratkaisut eri aloilla
- Laskennalliset mallit ja niihin liittyvät laskenta- ja datan säilytyskapasiteetit
- Suurteholaskennan uudet resurssit, ohjelmistojen ylläpito tulevaisuudessa, datan säilyttämiseen liittyvät ratkaisut
- Millaisia datanhallinnan ja laskennan erityiskysymyksiä on uusilla/kehittyvillä tutkimusaloilla?
- Millaisia datanhallinnan ja laskennan erityiskysymyksiä on erilaisilla tutkimusorganisaatioilla?
- Laskentakapasiteetin jakautuminen eri käyttäjäryhmien kesken, jonotusajat
- Hinnoittelu ja kustannustaso
- Toimintavarmuus (käyttökatkot)
- Datanhallinta tutkimuksen elinkaaren eri vaiheissa, avoin data, tietoturva

2. Palvelut

- Käyttäjätuki: mitä CSC:llä, mitä tutkimusorganisaatioissa?
- CSC:n tarjoamat koulutuspalvelut

3. Osaaminen

- Millaista osaamisen vahvistamista kasvava dataintensiivisyys ja kehittyvät tutkimusmenetelmät edellyttävät alallanne / organisaatiossanne?
- Tieteenalakohtaisten dataosaajien tarve (erityisosaajat)
- Substanssiosaajien osaamisen kehittäminen (esim. tutkijankoulutuksessa)

Sessio 2: Yhteistyö ja priorisoinnin periaatteet

4. Yhteistyö

- Millainen datanhallinnan ja laskennan infrastruktuurikonaisuus palveluineen mahdollistaa korkeatasoisen kansainvälisen yhteistyön?
- Suomen kiintiöt kansainvälisissä datakeskuksissa ja niiden käyttö
- Eri organisaatioiden yhteistyötä tukevat alustat
- Millainen merkitys yritys yhteistyöllä on datanhallinnan ja laskennan näkökulmista tutkimusalallanne / organisaationne edustamilla tutkimusaloilla?

5. Priorisoinnin periaatteet

Mitä näkökulmia ja miten ohjausryhmän tulisi painottaa hankintaa linjatessaan?

- Tutkimuksen monimuotoisuus: Laskenta ---- Suurteholaskenta
- Laskentapalveluiden/ suurteholaskennan edellytykset datanhallinnalle
- Resurssijaon läpinäkyvyys
- Alan tutkimuksen laatu, uudistuminen ja vaikuttavuus
- Kansainvälisen yhteistyön merkitys alalla: kansallinen infrastruktuuri yhteistyön mahdollistajana ja osaajien houkuttelijana
- Laskentainfrastruktuurin käyttäjäkunnan laajuus alalla (nyt/tulevaisuudessa)
- Datanhallintaan ja laskentaan liittyvä osaamistaso alalla
- Muut näkökulmat?

Osallistujien ehdottamia keskusteluaiheita pienryhmiin (edellisen sivun teemoittelua avattu yksityiskohtaisemmin)

Teema	Mille datanhallintaan, laskentaan ja osaamiseen liittyville asioille olisi tärkeä varata keskusteluaikaa tutkijoiden pienryhmissä?
Big data	Big data in human health and disease Bioinformatics
Datanhallinta	Datan "rekisteröinti" tutkimusprosessin aikana Aineistojen jatkohyödyntämisen tehostaminen myös yritysnäkökulmasta, miten tämä pitäisi huomioida? Asiakasnäkökulma monipuolisesti: tutkimushankkeiden elinkaari, nykyiset toimivat konseptit, missä parannettavaa, tulevaisuuden näkymiä Datan hallinta koko datan elinkaaren ajan suunnittelusta uusiokäyttöön ja pitkäaikaissäilytykseen Lainsäädännölliset ja eettiset kysymykset Datan avaaminen ja käyttöoikeudet, rajoitukset niihin Luottamus (dataan ja yleisesti tutkimukseen) Millaisen koko ketjun tulisi olla aineistonhallinnasta pitkäaikaissäilytykseen asti? Miten taataan suurten aineistojen turvallinen jakaminen?
Datatieteet	Avointen datajoukkojen kytkeminen. Lambda-arkkitehtuuri (batch + stream) ja erilaiset data science-sovellukset.
Laskenta	Dataintensiivisen laskennan luomat erityistarpeet käytettävälle infrastruktuurille. Laskennalliset mallit ja niihin liittyvät laskenta ja datan säilytyskapasiteetit. Miten varmistetaan riittävän resurssin saaminen koko (mahdollisesti useitakin vuosia) kestävä projektin ajaksi? Etenkin tilanteissa joissa tarvittavan resurssimäärän arviointi on ennen projektin aloitusta hankalaa. Miten saadaan resurssijakoon läpinäkyvyyttä?
Suurteholaskenta	Suurteholaskennan uudet resurssit sekä ohjelmistojen ylläpito tulevaisuudessa High-performance computing, access and use of massive data storage. Kansallinen GPU-laskentakapasiteetti, yritysten tarjoama ja sen hyödyntäminen. Suuriin data-aineistoihin liittyen datan säilyttäminen ja siihen liittyvät kysymykset Missä mennään kvanttilaskennassa tällä hetkellä? Suurteholaskennan kehittäminen Suomessa (CSC:n superkoneet)
Käyttäjätuki	Tutkimus sekä uusilla, että jo laskentaa vahvasti hyödyntävillä aloilla tarvitsevat käyttäjätukea. Miten tämä tuki toteutetaan eri organisaatioissa ja mikä tässä on CSC:n rooli. Tutkimuksen tukipalveluiden tulevaisuuden osaamistarpeet.
Osaaminen	Datan hyödyntämisen ja jalostamisen osaamisen kehittäminen Miten osaamista olisi tehokkainta jakaa tutkijoille? Tieteenalakohtaisten dataosaajien tarve Tietojohdaminen ja datan merkitys osaamiselle Opiskelijoiden mahdollisuudet
Infrastruktuurien yhteiskäyttö	Tiedon ja datan yhteiskäyttö Yhteistyön kehittäminen infrastruktuurin kehittämisessä

Eri organisaatioiden yhteistyötä tukevat alustat (yhteistyö myös yritysten kanssa).

Miten infrastruktuurit olisivat mahdollisimman laajasti käytettävissä?

**Kansainvälinen
näkökulma**

Datan avaamiseen liittyvät palvelut, mitä tarvitaan kansallisesti ja mitä voidaan koordinoida EU-tasolla

Kilpailukykyinen (laskenta) infrastuktuuri on kilpailuvaltti houkuteltaessa tieteen huippuja ja kilpailtaessa kv-rahoituksesta. Kuinka tätä näkökulmaa tulisi painottaa kohdennettaessa rahoitusta

Kansainväliset e-infrastruktuurit

Suomen kiintiöt kv-keskuksissa ja niiden käyttö. esim. kansainvälisen datakeskuksen ohjelmiston ja aineiston käyttäjä käyttää Suomen kiintiötä kv-laskentakeskuksessa tai CSC-palvelimella.