

Liikenne- ja viestintäministeriö
kirjaamo@lvm.fi
emil.asp@lvm.fi
PL 31
00023 VALTIONEUVOSTO

Viite: LVM138/05/2014

Lausunto kansallisesta big data -strategiasta

Liikenne- ja viestintäministeriö asetti vuoden alussa big data -käyttötyöryhmän. Laajapohjainen työryhmä pohtii big data -kehitystä Suomen näkökulmasta. Työryhmä valmistelee kansallisen big data -strategian. Työ valmistuu 30.6. mennessä.

Strategiassa tunnistetaan keskeisiä sovellusalueita ja yleisiä edellytyksiä big datan hyödyntämiselle. Strategiassa tehdään ehdotuksia toimenpiteistä, joiden avulla tavoitellaan tietoa-aineistojen ja uusien analyysimenetelmien laajempaa hyödyntämistä yhteiskunnan eri sektoreilla.

Suomen Akatemia pitää strategian lähestymistapaa hyvänä ja esitettyjä toimenpiteitä konkreettisina.

Yleiset kommentit

Strategia lähtee liikkeelle big data -käsitteestä. Yleinen *datalähtöisyyden tai datan hyödyntämisen periaate*, eli ajatus siitä että dataa voidaan (ja pitää) hyödyntää laveasti, on kuitenkin keskeisempi kuin datan määrän korostaminen. Big dataan verrattuna pienetkin datajoukot voivat olla erittäin merkittäviä niin julkisen hallinnon kuin yritysten näkökulmasta. Siten datatietoisuuden ja datan hyödyntämisen periaatteen korostaminen on kannatettavaa, ja strategiassa voisi vahvemmin korostaa sitä, että raakadatan määrä ei sinällään kerro tiedon arvosta paljonkaan.

Strategialuonnoksessa tuodaan selkeästi esiin big datan ja avoimen datan yhtymäkohtien merkitys. Tämä ja datan määrän suhteellisen pieni merkitys johtavat siihen, että voitaisiin jopa pohtia tarvetta strategian nimen muuttamiseen. Esimerkiksi ”Datalähtöisyys ja big data – kansallinen strategia” saattaisi olla sovelias; jos halutaan painottaa vielä avointa dataa, niin sekin voitaisiin ottaa huomioon lisänä.

Strategia sopivasti tunnistaa tarvittavia yleisiä edellytyksiä, joista esimerkkinä on koulutuksen, monitieteisen tutkimuksen ja soveltavan osaamisen muodostama, toisiinsa liittyvä kokonaisuus.

Big data on strategialuonnoksessa esitetty pitkälti hyödynnettävyyssnäkökulmasta. Korostumatta jää laajasti tehtävä alan tutkimus, joka on Suomessa erittäin korkeatasoista. Sillä on pitkät perinteet alkaen Teuvo Kohosen itseorganisointuvasta kartasta; suomalaista laadukasta osaamista on myös bioinformatiikan, data miningin (tiedonlouhinnan) ja hahmontunnistuksen alueilla samoin kuin tilastotieteessä ja erilaisissa sovelletun matematiikan osa-alueissa. (Big data on tuore termi: alaan liittyvää tutkimusta ja koulutusta on toki annettu muiden termien alla pitkään.) Tietojenkäsittelytieteellisen tutkimuksen merkitystä voisikin korostaa toimenpiteitä pohdittaessa. Yhteiskuntatieteellisen tutkimuksen rooli on tuotu hyvin esiin.

Kansallista strategiaa pohdittaessa on syytä korostaa koulutuksen merkitystä. Huippuosaamisen lisäksi tulee jatkossa pitää huolta siitä, että kaikessa korkeakoulutasoisessa koulutuksessa saavutetaan tietynlainen *datalukutaito*: kyky arvioida tietyn päätelmän takana olevaa dataa, pohtia suurimpia mahdollisia virhelähteitä jne.

Strategia pyrkii hyvin tunnistamaan ja priorisoimaan Suomen kannalta oleellisia – sekä yleisiä että kaupallisia – sovellusalueita. Merkittävien alueiden valinta on aina vaikeata, ja sen takia onkin tärkeää, että koulutuksessa ja tutkimuksessa luodaan monipuoliset valmiudet, joiden avulla uusiin mahdollisuuksiin pystytään tarttumaan. Siten keskittymistä yksittäisiin teknologioihin tulee välttää ja pyrkiä monipuolisuuteen.

Strategia tuo esiin mahdollisuuden antaa uudennaisia näkökulmia hallinnolle edistää tuottavamman yhteiskunnan ja kestävyysvajeen torjumisen tavoitteita. Myös infrastruktuurit perusedellytyksenä ja niihin liittyvä kansallisen yhteisen suunnittelun ja toimien tarve tuodaan esiin. Nämä ovat molemmat strategian merkittäviä ja tärkeitä näkökulmia.

Esitetyt toimenpide-ehdotukset

Esitetyt toimenpide-ehdotukset ovat konkreettisia ja niitä on tunnistettu monia. Akatemia kiinnittää huomiota siihen, että Tutkimuksen alaotsikossa ei ole esitetty toimenpide-ehdotuksia.

Minkälaisiin big dataa liittyviin toimiin Akatemia on ryhtynyt?

Data-analyysin merkitys modernissa tieteessä on lähes itsestäänselvyys, ja sen takia hyvin monissa Suomen Akatemian rahoittamissa tutkimushankkeissa on merkittävä big data -komponentti.

Suomen Akatemia on tukenut data-analyysiin liittyvää tutkimusta jo pitkään. Alueella on toiminut useita Akatemian huippuyksikköjä pitkälti yli 10 vuotta. Tällä hetkellä Akatemia rahoittaa Laskennallisen päättelyn huippuyksikköä COIN (2012–2017) (maininta s. 14 infolaatikossa). Myös muussa Akatemian rahoittamassa tutkimuksessa tutkitaan big data -kokonaisuuksia: näin on esimerkiksi Syöpägenetiikan tutkimuksen huippuyksikössä. LASTU-tutkimusohjelma on esimerkki tutkimusohjelmasta, jossa data-analyttisillä menetelmillä on selkeä rooli.

Suomen Akatemian luonnontieteiden ja tekniikan tutkimuksen toimikunta on esittänyt vuodelle 2015 ICT2023-ohjelman kolmatta ja neljättä hakua. Kolmannen haun keskeisiksi teemoiksi esitetään suurten tietomassojen ja avoimen tiedon (big data, open data) hallinnan ja analyysin menetelmiä ja sovelluksia.

Suomen Akatemian yhteyteen perustetaan vasta hyväksytyn lakimuutoksen seurauksena strategisen tutkimuksen neuvosto, jonka yhtenä tavoitteena on rahoittaa päätöksentekoa tukevaa tutkimusta (evidence-based policy making).

Pieniä korjauksia: ALGODAN (maininta s. 14 infolaatikossa) oli Suomen Akatemian rahoittama huippuyksikkö vuosina 2008–2013. LASTU (maininta s. 14 infolaatikossa) on Suomen Akatemian rahoittama tutkimusohjelma vuosina 2010–2016.



Minkälainen rooli Akatemialla voisi olla kansallisessa big data -kehittämistyössä?

Suomen Akatemialla on rooli tutkimustoiminnan rahoittajana sekä big datan ja yleisemmin data-analyysin menetelmien tutkimuksessa että eri sovellusalueiden tutkimuksessa.

Suomen Akatemia on halukas osallistumaan kansalliseen big data -kehittämistyöhön siihen liittyvän tieteen ja tutkimuksen asiantuntijana.

Pääjohtaja

Heikki Mannila

Johtava tiedeasiantuntija

Paavo-Petri Ahonen