

Avoim data Suomessa – aktivisteja, EU-komitologiaa ja hallitusvisioita

Me kaikki käytämme avointa dataa, kun tarkistamme lähimmän bussin lähtöajan kännykästämme tai kun suunnistamme navigaattorin avulla määränpäähämme. Avoimen datan varaan asetetaan suuria toiveita eri valtioiden ja koko EU:n kilpailukyvyyn kannalta. Miten tähän on päädytty, vaikka vain toistakymmentä vuotta sitten esimerkiksi Suomen julkishallinnon tietoaineistojen käyttöä ohjattiin lainsäädännön tasolla lähinnä budjettirahoitusta täydentävän asiakasrahoituksen lähteenä?

Avoimen datan historia on Suomessa noin 10 vuoden mittainen. Vuonna 2009 avoimen datan aktiivit järjestivät Suomessa ensimmäisen avoimen datan kilpailun nimellä Apps4Democracy. Kilpailu järjestettiin nimillä Apps4Finland ja Open Finland Challenge vuoteen 2015 saakka. Sen jälkeen avoimen datan tapahtumat ovat jatkuneet mm. lukuisten hackathonien muodossa (Ultrahack, Junction, Hack for Society jne.).

Suomi on datan avoimuudessa kansainvälisesti kärkijoukoissa

Aktiivisten avoimen datan käyttäjien lisäksi on ollut myös aktiivisia avoimen datan tarjoajia. Vuonna 2011 pääkaupunkiseudun kaupungit perustivat Helsinki Region Infoshare (HRI) -palvelun ja alkoivat julkaista julkisia datojaan avoimena datana. Vuonna 2012 perustetun Open Knowledge Finlandin (OKFI) tavoitteena taas on ollut edistää tiedon avaamista, avoimen tiedon hyödyntämistä sekä avoimen yhteiskunnan kehittymistä Suomessa. OKFI ja HRI järjestivätkin yhdessä avoimen datan ensimmäinen vuosikymmen -tapahtuman lokakuussa 2019.

Suomi sijoittuu tänä päivänä hyvin kansainvälisessä vertailussa. Open Knowledge Finland on osa kansainvälistä Open Knowledge International -verkostoa. Se edistää tiedon avaamista ja hyödyntämistä ympäri maailmaa ja ylläpitää esimerkiksi eri maiden julkisen tiedon avoimuutta mittaavaa Global Open Data Index -palvelua. Suomi sijoittuu datan avoimuudessa viidenneksi 94 maan joukossa. Toinen avoimen datan kansainvälistä vertailua tekevä organisaatio, The Open Data Inventory, sijoittaa Suomen kuudenneksi 178 maan joukossa. Suomi on siis tällä hetkellä datan avoimuudessa kansainvälisesti kärkijoukoissa, mutta kehitys tähän pisteeseen ei ole ollut mitenkään itsestään selvää eikä suoraviivaista.

Suomen maksuperustelaki vs. EU:n avoimen datan direktiivi (PSI)

Vuonna 1999 Tilastokeskus avasi kaikille avoimen tilastotietokannan (StatFin) silloisessa vielä varsin uudessa ja nopeasti kehittyvässä internet-verkossa. Tuolloin viraston piti aivan erikseen perustella se, että tietokanta avattiin verkossa veloittamatta kaikille käyttäjille. Miksi? Se johtui maksuperustelasta, jonka mukaan tietokantamuodossa julkaistun datan tuli lain mukaan olla lähtökohtaisesti maksullista. Ilman poikkeusta maksuperustelakiin tietokantaa ei yksinkertaisesti olisi voinut avata ilmaisena kaikille käyttäjille. Senkin jälkeen veloittamattoman tietokannan käyttöä rajoitti pitkään mm. luettelosuoja, jonka perusteella tietojen kaupallista edelleen käyttöä rajoitettiin.

EU:n avoimen datan direktiivi (PSI-direktiivi) vuonna 2003 oli merkittävä säädös, joka vaikutti avoimen datan linjauksiin myös Suomessa. Direktiivissä EU:n komissio määräsi julkisen sektorin informaation uudelleenkäytöstä siten, että kaikki julkisen sektorin tuottama tieto pitäisi lähtökohtaisesti luovuttaa käyttäjille veloituksettomasti vapaasti käytäväksi.

Lainsäädäntö eri EU-maissa julkisen sektorin tuottaman tiedon maksuttomuudesta vaihteli 2000-luvun alussa suuresti. EU:n tilastovirasto Eurostat alkoi soveltaa uutta PSI-direktiiviä käytännössä julkaisemalla kansallisilta tilastoviranomaisilta kokoamansa tilastotiedot veloituksettomina omissa palveluissaan ja tietokannoissaan. Esimerkiksi joitakin Tilastokeskuksen tietoaineistoja ei saanut edelleenkään vapaasti käyttää kaupalliseen käyttöön, mutta samat tiedot sai kuitenkin ladata Eurostatin tietokannoista vapaasti, vieläpä myös kaupalliseen käyttöön. Ja tilanne oli sama monissa muissakin EU-maissa.

Käyttöoikeuksien rajoituksista Creative Commons -lisenssiin

Asiakaspaine tiedon vapaalle jatkoikäilyntämiselle ymmärrettävästi kasvoi vuosien mittaan. Tilastokeskuksen aineistojen käyttöehdoissa sallittiin veloituksettoman tilastotietokannan tietojen vapaa hyödyntäminen myös kaupallisesti vuodesta 2012 lähtien. Tilastokeskus on avannut aiemmin maksullisia tietokokonaisuuksia sekä datan avaamiseen tarkoitettun erillishoidetuksen että oman budjettirahoituksensa turvin (kuten paljon käytetty postinumeroalueittainen avoin aluetietokokonaisuus).

Vuoden 2018 alusta lähtien Tilastokeskuksen avoimen datan käyttöehdot ovat noudattaneet suoraan Creative Commons Nimeä 4.0 -lisenssiä (CC BY 4.0). Vuonna 2014 Creative Commons -lisenssistä tuli Suomen julkishallinnon ja vuonna 2019 myös EU:n standardilisenssi.

INSPIRE direktiivi määräsi kartta-aineistot avoimiksi

Vuonna 2007 hyväksyttiin EU:n INSPIRE-direktiivi, joka määritteli avoimissa kartta-aineistoissa käytettävät standardit sekä määritteli PSI-direktiivin mukaisesti aineiston jakamisen kaikille avoimesti edelleen hyödynnettäviksi. Maanmittauslaitos (MML) avasi vuonna 2012 maastotietoaineistonsa. ”Pari päivää sen jälkeen, kun data oli avattu, oli sitä ladattu enemmän kuin koskaan siihen mennessä historian aikana” kertoo MML:n INSPIRE direktiivin toimeenpanosta vastannut Antti Rainio HRI:n avoimen datan 10-vuotisjuhlatilaisuudessa lokakuussa 2019.



Avoimen tiedon kaupallinen merkitys on valtava, vaikkei sitä välttämättä heti tule ajatelleeksi. Esimerkiksi Yhdysvalloissa on verovaroin tuotettua dataa pidetty varsin pitkään kaupallisestikin vapaasti hyödynnettävänä. Tästä näkökulmasta oli luontevaa, että presidentti Ronald Reagan määräsi satelliittipaikannusdatan avattavaksi jo vuonna 1983.



Kolmenkymmenen vuoden kuluttua noin 3,3 miljoonaa työpaikkaa USA:ssa perustui tämän datan hyödyntämiselle.

EU on määritellyt viideksi tärkeimmäksi avoimen datan osa-alueeksi tilastot, paikka-, liikenne- ja yritystiedot sekä maapalloa koskevan havaintotiedon. EU arvioi PSI-direktiivin suoraksi taloudelliseksi arvoksi 52 miljardia euroa vuonna 2018. Summan arvioidaan kasvavan 194 miljardiin euroon vuonna 2030. Avoin data onkin EU:ssa vahvasti myös osa avointa tiedepolitiikkaa, tasapuolista kilpailupolitiikkaa sekä digitaalista teollisuuspolitiikkaa.

Avoin data tietopolitiikan kantavaksi periaatteeksi

Tiedon avaaminen uudelleenkäyttöön on ollut viime vuosien aikana mukana useissa hallitusohjelmissa, esimerkiksi avoimen tiedon ohjelmassa vuosina 2013-2015. Huhtikuussa 2020 Marinin hallitus asetti Tiedon hyödyntämisen ja avaamisen -hankkeen (2020-2022). Hanke on myös osa avoimen datan PSI-direktiivin tuoreimman, vuonna 2019 tehdyn päivityksen, kansallisesti sitovaa toteutusta.

Hanke edistää hallitusohjelman tavoitetta julkisen tiedon entistä laajempaan ja tehokkaampaan hyödyntämiseen. VM:n asettamassa hankkeessa ovat käytännön tasolla keskeisesti mukana mm. Tilastokeskus, avoindata.fi sekä Digi- ja väestötietovirasto.

Avoimen datan vaikuttavuuden mittaaminen ei ole helppoa

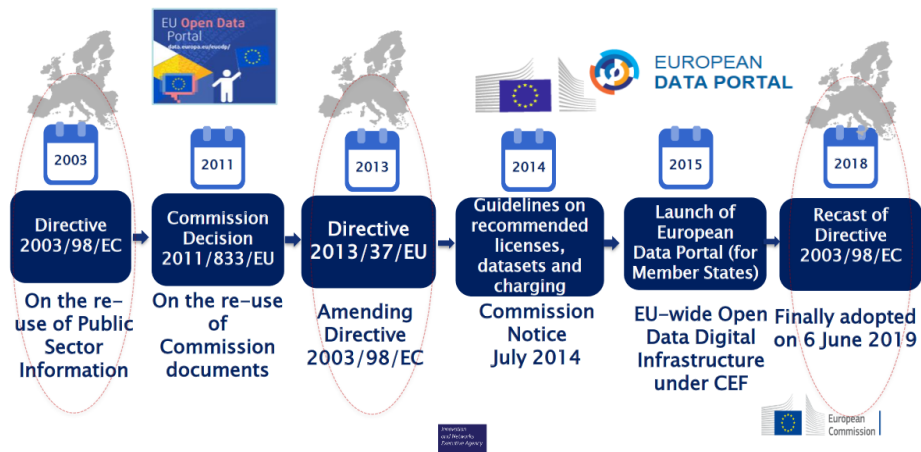
Avoimen datan taloudellinen merkitys on huomattava ja kasvussa. Avoimen datan vaikuttavuudesta on kuitenkin vain vähän tutkittua tietoa mm. siksi, että vaikutusten mittaaminen on varsin hankalaa. Vuonna 2017 Suomessa tehtiin tutkimus *Avoimen datan hyödyntäminen ja vaikuttavuus*.

Tutkimuksessa osoitettiin esimerkiksi, että avointa dataa innovaatiotoiminnassaan hyödyntävät yritykset tekevät uusia tuoteinnovaatioita useammin kuin yritykset, jotka eivät hyödynnä dataa. Joillakin sektoreilla avointa dataa hyödyntävät yritykset olivat kasvaneet muita yrityksiä nopeammin. Myös uusia innovaatioita avointa dataa hyödyntäneet yritykset tekivät enemmän kuin muut, näin erityisesti palveluihin liittyen.



Hyvä esimerkki avoimen datan vaikuttavuudesta on näkövammaisille kehitetty kännykkäsovellus BlindSquare, joka voitti Apps4Finland-kilpailun sovellussarjan vuonna 2012. Palvelu pohjautuu avoimeen GPS-teknologiaan, joukkoistettuun ja avoimeen kartta-aineistoon (OpenStreetMap), sekä joukkoistamalla syntyneeseen avoimeen paikkatietoaineistoon FourSquareen. Ilman näitä avoimia aineistoja tai palveluja ei tätäkään käyttäjilleen kullanarvoista sovellusta olisi olemassa.

Milestones of EU PSI policy



Alkuperäinen EU:n julkisen sektorin informaation uudelleenkäyttöä käsittelevä direktiivi (PSI) vuodelta 2003 on kehittynyt eri vaiheiden kautta vuonna 2019 hyväksytyksi EU:n avoimen datan direktiiviksi. Kaikkien EU-maiden tulee implementoida direktiivi sitovana oman maansa lainsäädännön osaksi viimeistään vuonna 2021.

Termien selitykset:

PSI = Public Sector Information

INSPIRE = Infrastructure for Spatial Information in Europe

Lähteet:

10 vuotta avointa dataa -tapahtuma — Avoindata.fi kiittää osallistuneita!

<https://www.avoindata.fi/fi/artikkeli/10-vuotta-avointa-dataa-tapahtuma-avoindatafi-kiittaa-osallistuneita-0>

Avoin data 10 vuotta: "Datan avoimuus pitää saada lakiin": <https://hri.fi/fi/avoin-data-10-vuotta-datan-avoimuus-pitaa-saada-lakiin/>

Global Open Data Index: <https://index.okfn.org/place/>

Open Data Watch: <https://odin.opendatawatch.com/>

European Commission / Open Data: <https://ec.europa.eu/digital-single-market/en/open-data>

Avoimen datan hyödyntäminen ja vaikuttavuus: <https://tietokayttoon.fi/julkaisu?pubid=18703>

Tiedon hyödyntäminen ja avaaminen: <https://vm.fi/tiedon-hyodyntaminen-ja-avaaminen>

Kirjoittaja:

Markku Huttunen

<https://www.linkedin.com/in/markku-huttunen-7ab61a3/>