



AVOIN JUTTU

Raaka. Data. Auki. Niin **Tim Berners-Lee** kuin **Barack Obamakin** haluavat avata datan. Mistä oikein on kyse?

Mitä tarkoittaa AVOIN DATA

Erityisesti julkishallinnon, mutta myös yritysten ja muiden organisaatioiden tietojärjestelmiin sekä yksityishenkilöille kertynyttä ja lostamatonta informaatiota, johon on avattu maksuton pääsy organisaation ulkopuolisille.

Datan avoimuuden asteeseen vaikuttavat ennen kaikkea tekninen saavutettavuus (eli datan koneluettavuus), maksuttomuus, datan käyttöoikeudet (lisensiointi), sekä datan löydettävyyttä.

Miksi tästä aiheesta piti kirjoittaa juttu?

Internetin kehitys on nyt sekä teknisesti että käytötapojen osalta sellaisessa vaiheessa, että datan avaaminen on kannattavampaa kuin koskaan ennen.

Julkisten organisaatioiden tietokannoissa lojuu valtavat määrät hyödyntämätöntä tietoa. Sitä ei yleensä pääse käyttämään edes naapurivirastosta saati hallinnon ulkopuolelta. Ja jos pääseekin, niin käytöstä pitää maksaa tai se on kohtuuttoman hankalaa.

Julkishallinnon datavarannon hyödyntämistä ohjaava maksuperustelaki lisää byrokratiaa, vähentää datan hyödyntämistä ja synnyttää jopa päällekkäisiä tietovarastoja, kun tehostamispaineet ohjaavat säästämään lyhytnäköisesti. Ehkä suurin syy suljettuihin tietovarastoihin on kuitenkin oman organisaation tulosohejaus ja asenne.

Hallinnon data on nykyaikaisen yhteiskunnan luonnonvara, joka voi paitsi lisätä hallinnon tehokkuutta ja kansalaisten osallistumismahdollisuuksia myös luoda uutta liiketoimintaa.

Avoin data luo hedelmällisen ekosysteemin

Suomen Akatemian rahoittama Somus-tutkimushanke järjesti viime syksynä *Kansalaisosallistujan työkalut* -kilpailun hallinnon avoimen datan ja sovellusten saamiseksi esiin. Kilpailuun tuli 23 hyvää ehdotusta, niin ideoita, pilottoja kuin valmiita sovelluksia.

Kilpailu osoitti koodaajille, että joku on kiinnostunut heidän sovelluksistaan. Hallinnon edustajat taas saivat avaria ideoita datansa hyödyntämiseksi. Ennen kaikkea konseptioijat, koodaajat ja hallinnon edustajat pääsivät käynnistämään keskinäisen keskustelun siitä, miten julkisen datan avoimuutta voitaisiin edistää.

Vastaanvanlaisia mashup-kilpailuja järjestetään ympäri maailmaa. Useimpien tavoitteena on luoda avoimen datan ekosysteemi. Siinä datan jakaja ja hyödyntäjä eivät ole välttämättä vastavuoroisesti yhteistyössä, vaan kokonaisuus hyödyttää kaikkia. Kuka tahansa voi noukkia mitä tahansa ideoita ja aineistoja ja yrittää tehdä paremmin.

Myös netin ripeä kehitys perustuu pitkälti ekosysteemijatteluun. Nykymuotoinen netti ei olisi mahdollinen ilman avoimia, yhdenmukaisia rajapintoja sekä ideoiden ja konseptien lainailua.

Datan lähteillä

Yhdysvaltain liittovaltio avasi toukokuussa 2009 data.gov-palvelun, jonne on koottu tiedot hallinnon tarjoamasta datasta ja sen käyttämisestä. Samanlaisia datakatalogeja on perustettu joka puolella maailmaa. Reilu kuukausi sitten Britanniassa julkistettu

data.gov.uk on amerikkalaisvastinettaan laajempi, monipuolisempi ja paremmin koneluettava linked data -tukensa takia.

Suomen kansallinen datakatalogi on toistaiseksi vain lyhyt linkkilista (suomi.fi/datakatalogi), mutta valtion it-palvelukes-

kuksen ylläpitämä Suomi.fi-portaali kehittää siitä kehittyneempää versiota. Valtion katalogin lisäksi Somus-hanke ja Kirjastot.fi-palvelu kehittävät yhteisöllistä pääasiassa ohjelmoijille ja palveluideoijille suunnattua sivustoa (opengov.fi).

1. Hallinto

Esimerkki: Ympäristöhallinnon Oiva-palvelu tarjoaa tietoa vesivarjoista, eliölajeista ja alueiden käytöstä sekä ympäristöön liittyviä paikkatietoaineistoja.

Hallinnossa meneillään olevat avaamishankkeet tulevat moninkertaistamaan tarjolla olevan datan lähivuosina.

2. Yritykset

Esimerkki: Google tarjoaa muun muassa karttadataa avoimen rajapinnan kautta.

3. Kansalaiset

Esimerkki: OpenStreetMap on vapaaehtoisvoimin ylläpidetty karttapalvelu, jota kuka tahansa voi päivittää esimerkiksi navigaattorilla keräämällään tiedolla.

"Pidän ehkä merkittävimpänä verkon tuomana muutoksena sitä, että tiedon tuottaminen, korjaaminen ja täydentäminen on tullut yksittäisten kansalaisten hyppysiin. Tässä ollaan kuitenkin vasta ihan alkumetreillä." – Jyri Engeström

Miten (hallinnon) data saadaan auki?

1. POLITIIKKA = Luotava yhteinen tahotila tuoda julkiset tietovarannot kaikkien käyttöön, varmistettava päättäjien tuki.

2. JURIDIikka = Sovittava datan käytön pelisäännöt, minkälaisilla ehdoilla dataa voi käyttää ja levittää.

3. TEKNIikka = Julkaistava tiedot datavarannon sisällöstä, muodosta ja rajapinnoista, avattava verkkopalvelu datan jakamiseksi

4. BYROKRATIA = Sovittava toimintamallit ja kehitettävä jakamismyönteistä asennetta.

Muodolla on väliä

Datan hyödynnettävyyteen vaikuttaa suuresti se, minkälaisessa muodossa data on ja minkälaisen rajapintojen kautta pääsee käyttämään. Mitä yleisempiä standardeja

datan muoto ja sen rajapinnat noudattavat, sitä helpompaa datan hyödyntäminen on.

Datan käytettävyyttä voidaan parantaa myös kuvaamalla datan merkityksiä semanttisen webin tekniikoin. Suomessakin rakennetaan kansallisen semanttisen webin infrastruktuuria FinnONTO-hankkeessa.

Datan löytämisen helpottamiseksi tieto siitä pitää julkaista. Datavarantoja esitteleviä koontipalveluita kutsutaan yleensä datakatalogeiksi.

Standardit myöhemmin

Jos data ei ole valmiiksi standardimuotoista, voi sopivien rajapintojen etsiminen, määrittely ja toteutus osoittautua aikaa vieväksi ja kalliiksi.

Kannattaakin pohtia, voisiko datan julkaista sellaisenaan ja kertoa siitä sopiville tahoille. Jos data on jollekin tärkeää, se otetaan käyttöön tavalla tai toisella.

Tarvittavia rajapintoja voi toteuttaa saadun palautteen perusteella. Pohtimalla datan julkaisutapaa yhdessä tarvitsijoiden

kanssa voidaan vähentää turhaa työtä.

Bisnestä datanvälittäjälle?

Rajapintapalveluiden kaupallinen järjestäminen voi myös muodostua suhteettoman kalliiksi.

Verkkorajapintapalveluihin voisikin syntyä uutta liiketoimintaa, jos joku tahoo päättäisi toimia avoimen datan välittäjänä. Se helpottaisi sekä datan julkaisua että käyttöä.

Datan omistajan ei tarvitsisi investoida netin reunalle tulevaan palvelininfraan ja kaistaan sekä niiden varmistamiseen ja turvallisuuteen. Omistaja voisi rekisteröidä oman datalähteensä "datavälityskeskukseen", joka sitten tarjoaisi sen kaikilla mahdollisilla palvelukutsumuodoilla eteenpäin.

Välityskeskukseen voisi joko viedä dataa tai kertoa rajapinnan, jolla keskus saa haettua raakadatan itselleen. Näin datan käyttökin helpottuisi, kun kutsumuodot olisi aina samanlaisia. ■

Miten avoimella datalla voi tehdä rahaa?

Dataa avautuu koko ajan kiihtyvällä tahdilla. Samalla syntyy aivan uudenlaisia bisnesmahdollisuuksia. Jonkun pitää vain keksiä ne.

■ Koosta datalähteistä uudenlaisia palveluita.

■ Käytä dataa tutkimusaineistona tuotekehityksessä

■ Ryhdy avoimen datan välittäjäksi

■ Avaa datasi, niin avoimen datan ekosysteemi rikastuttaa myös sinua. Datan avaaminen ja levittäminen internetissä ei maksa paljoa.

■ Paranna datasi laatua avaamalla käyttäjille mahdollisuus täydentää sitä

Esimerkki: Navigaattorivalmistaja TomTom avasi navigaattorinsa käyttäjille mahdollisuuden ilmoittaa kartoissa olevista virheistä ja tehdä korjauksia. Yhtiö sai nopeasti niin paljon parannusehdotuksia, ettei niiden hankkiminen olisi ollut muilla keinoin taloudellisesti järkevää.

Helsingin seututieto kaikkien käyttöön



Helsingin seudulla ollaan lähdetty edistämään julkisen tiedon avaamista käytännössä. *Helsinki Region Infoshare* -projektin tavoitteena on jakaa Helsingin seutua koskevaa tietoa kaikille kiinnostuneille. Projektia valmistelee Forum Virium Helsinki yhdessä Helsingin kaupungin Tietokeskuksen ja Helsingin seudun kuntien sekä seutuorganisaatioiden kanssa.

Projektissa luodaan verkkopalvelu, eräänlainen seututiedollinen datakatalogi, jonka avulla Helsingin seutua tai sen osia kuvaava tieto on helposti löydettävissä ja käytettävissä. Avattava data on kuntien tuottamaa numeerista ja tilastollista tietoa, joka on aiemmin ollut pääosin vain kuntien työntekijöiden käytössä. Tarkoitus on, että käyttäjä voi vapaasti ja maksutta ladata tietoa omalle koneelleen, käyttää sitä sovelluksissaan tai rakentaa sen pohjalta uusia palveluita.

Avattava tieto liittyy muun muassa asumiseen, koulutukseen, rakentamiseen, työllisyyteen ja yritystoimintaan. Projektin piloteissa selvitetään, miten eri käyttäjäryhmät käyttäisivät tietoa ja testataan tiedonjakomallia sekä rakennettavaa verkkopalvelua.

Esimerkiksi kauppias voisi hyödyntää Helsingin seudun jotakin aluetta koskevaa väestötietoa **suunnitellessaan toimintansa laajentamista tai markkinointinsa kehittämistä**. Opiskelija voisi hyödyntää vaikkapa väestön muutosliikkeitä kuvaavia tietoja tutkielmansa tekemisessä. Seututietoa voidaan myös yhdistellä muihin tietoihin ja vaikkapa esittää ne kartalla.

Viime marraskuussa perustettu Helsinki Region Infoshare on kolmivuotinen projekti, mutta palvelun beta-versio ja pilotihankkeet on tarkoitus käynnistää jo tänä vuonna.

Mallia datakatalogin rakentamiseen on haettu muun muassa Yhdysvalloista ja Britanniasta. Tiedon avaamisessa Suomi ja Helsingin seutu ovat tietenkin oma tapauksensa, jossa pitää huomioida teknisten kysymysten lisäksi myös lainsäädäntö ja sopimusasiat.

Hankkeen tavoitteena on oppia, mitä julkisen tiedon avaaminen tarkoittaa seututasolla käytännössä ja myös jakaa oppeja muiden käyttöön. Voisiko tiedon avaaminen esimerkiksi edistää hallinnon sisäisiä, tiedon kehittämiseen liittyviä projekteja? ■

Pyöräilijä jakaa dataa



Näyttävät yhteisöllisesti kootut tietovarannot kuten Wikipedia ja OpenStreetMap ovat hyvin tunnettuja, mutta jostain syystä usein unohdetaan, että kansalliset ja yritykset voivat tuottaa dataa myös julkishallinnon hyödynnettäväksi.

Yksi esimerkki tällaisesta on kaupungin liikennesuunnittelijoille, rakennusvirastolle ja pyöräilijöille tarkoitettu karttapohjainen yhteistyöformaatti Fillarikanava. Siinä pyöräilijät merkitsevät kartalle arkisia pyöräilyoloihin liittyviä havaintojaan ja ne tallentuvat paikkatiedoksi, joka voidaan huomioida alueita koskevissa suunnittelu- ja parannushankkeissa.

Arkitietoa virkamiehille

Fillarikanava syntyi, kun Helsingin ”pyöräilynedistämisorukoissa” aktiiviset TKK:n tutkijat alkoivat ideoida, miten pyöräilijöiden arkitiedon saisi koostettua virkamiesten työtä helpottavaksi informaatioksi. Tutkijat esittelivät idean Helsingin sähköisten palautekanavien hankepäällikölle, joka innostui pilotoimaan hanketta.

Avoimen datan hyödyntämiselle on Fil-

larikanavassa paljon mahdollisuuksia, sillä esimerkiksi kaupungin paikkatietopalvelun karttasovelluksessa näkyvät katutyöt olisivat pyöräilijöille hyödyllistä tietoa, samoin poliisin rekisterit pyöräilyonnettomuuksista. Google Street View’n integrointi Fillarikanavaan on lähitulevaisuutta, koska sen rajapinta on avoin. Karttapohjaksi on harkittu OpenStreetMapia, joka on Google-karttaa tarkempi. Tiehallinnon Digiroad ja kaupungin omat kartta-aineistot eivät tällä haavaa ole suunnitelmissa niiden hankalan saatavuuden takia.

Fillarikanavassa itsessään on olemassa rajapinnat, joiden kautta pyöräilyhuomioita saadaan ulos, mutta niitä ei ole vielä markkinoitu kovin laajasti. Vahvojen ja hyvin dokumentoitujen rajapintojen tekeminen vaatii kuitenkin panostusta ja Fillarikanavan pienille kehitysresursseille on toistaiseksi riittänyt muitakin kohteita.

Sama koskee Fillarikanavan koodin julkistamista avoimen koodin projektiksi. Avoin data, avoin lähdekoodi ja avoimet formaatit ovat Fillarikanavan toimittajankin omalla prioriteettilistalla melko korkealla, mutta julkishallinnolta niitä tulisi suorastaan vaatia. ■

Toimittajalta Jouni Junkkaala

Aika erikoinen juttu

Teksti: Jouni Junkkaala
>> Kommentit: Santtu Toivonen

Tiedonvälityksen mahdollisuudet ovat nyt paremmat kuin koskaan. Lehtien tekeminen ei silti ole muuttunut juuri mitenkään. Miksei lukijoiden osaamista ja innostusta hyödynnetä, kun siihen keran on niin hienot välineet?

>> *Voisiko lisäksi vielä mainita sen, että ihmiset on nyt somen ansiosta tottuneet osallistumaan? Eli paitsi että välineet on muuttuneet, myös käyttäytymismallit on. Nämä muutokset etenevät rinta rintaan.*

Päätimme kokeilla, voisiko jutun tehdä kokoaan netissä. Kaikki alkoi **Santtu Toivosen** ideasta. Lyhyen aivoriihen jälkeen heitin ajatuksen keskustelupalvelu Qaikuun. Se sai innostuneen vastaanoton, joten aloitimme.

”Jutun tekeminen avoimesti verkossa on hienoa ja hirveää.”

Valitsimme jutulle aiheen, ideoimme jutun sisällön ja rakenteen Qaikuun. Tekstit kirjoitimme EtherPad-yhteiskirjoituspalvelussa. Jutun taittomalli kehittyi Google Docsin esitysohjelmassa. Kaikilla oli mahdollisuus vaikuttaa kaikissa vaiheissa – siis myös muokata lopullista tekstiä.

Lopputulos on näillä sivuilla.

Mitä opittiin?

Jutun tekeminen avoimesti verkossa oli hienoa ja hirveää. Oli mahtavaa päästä tekemään asioita uudella tavalla ja huomata, miten niin moni tuli mukaan. Alussa hirvitti, kun lopputuloksesta ei ollut mitään takeita. Toisen kerran pelotti, kun tekstin tuottaminen ei käynnistynyt odotetusti.

Näillä se tehtiin

Qaiku on suomalaisen Rohean kehittämä keskustelupalvelu, joka voi näyttää ulospäin Twittermäiseltä mikroblogilta, mutta on sitä selvästi parempi keskustelujen synnyttämiseen, seuraamiseen ja arkistointiin.

Qaiku valikoitui jutun työtyöympäristöksi, koska siellä oli valmiiksi aktiivinen ja uusista toimintatavoista innostunut yhteisö.

Googlen nykyään omistama **EtherPad** on erittäin helppokäyttöinen yhteiskirjoituspalvelu, ikään kuin verkkosivulle tallennettu yksinkertainen Word-dokumentti, johon useampi henkilö voi yhtä aikaa kirjoittaa tekstiä ja muokata sitä.

Saimme kuin saimmekin jutun valmiiksi.

Lopputuloksena jää lukijan arvioitavaksi, mutta prosessin osalta pidän hanketta onnistuneena.

Tein monta asiaa väärin.

Olisi pitänyt varata enemmän aikaa, mainostaa asiaa useammalla ihmiselle ja haastaa osallistujia aktiivisemmin. Opimme kuitenkin paljon, ja sehän hankkeen tarkoitus ennen muuta oli.

>> *Tässä voisi ehkä vetää pienen analogian yleensä someen ja osallistamiseen. Sehän kiinnostaa monia yrityksiä. Eli että aikaa ja resursseja pitää varata yhteisön aktivointiin ja hienovaraiseen ”ohjaamiseen”. Pelkkä Facebook-sivun perustaminen ei riitä.*

Avoin ryhmätyöskentely verkossa sopii hyvin ideointiin ja tekstien arviointiin. Tekstin tuottaminen on vaikeampaa, sillä se vaatii motivaatiota ja sitoutumista.

Prosessin vetäjän rooli on ratkaiseva: pitää löytää innostuneet tekijät, varmistaa että keskustelu pysyy käynnissä ja kirjoittaminen etenee.

Jonkun pitää myös tehdä lopulliset välinat ja vastata siitä, että juttuun tulee sopiva määrä merkkejä määräpäivään mennessä. Nyt se tarkoitti sitä, että jouduin itse tuottamaan enemmän tekstiä, kuin olisin halunnut mutta en kuitenkaan tarpeeksi kokeakseni sen omakseeni.

Jutun tuottaminen avoimesti verkossa on vaikeaa ja epävarmaa, mutta siinä on paljon potentiaalia. Laajapohjainen tekijäkaarti ja koko ajan kehittyvät välineet mahdollistavat paljon sellaista, mihin yksittäinen toimittaja ei pysty.

Tämä oli vasta alkua. ■



Seuraavat henkilöt osallistuivat tämän juttukokonaisuuden tekemistä koskevaan keskusteluun. Osa kirjoitti yhden kommentin, jotkut kymmeniä. Tekstiä juttuun tuotti toistakymmentä eri henkilöä.

Tuija Aalto, Kari Ahokas, Christina Forsgård, Kari A. Hintikka, Ferrix Hovi, Jouni K. Juntunen, Petja Jäppinen, Timo Kantola, Helge V. Keitel, Juha Matti Keskinen, Mitro Koivunen, Petri Kola, Suvi Korhonen, Ossi Kuittinen, Jarmo Lahti, Marjatta Laine, Antti Lindström, Karoliina Luoto, Pirjo Näkki, Mace Djala, Tarja Ollas, Ilkka Pirttimaa, Antti Poikola, Harto Pönkä, Samuel Rinnetmäki, Teemu Ropponen, Janne Saarikko, Pekka Sarkola, Tuija Sonkkila, Arto Toivanen, Santtu Toivonen, Kim Viljanen, Kimmo Virtanen

[Lue lisää]
tivi.fi/avoinjuttu