

Hallinto tuottaa jatkuvasti suuren määrän dataa, joka voisi olla tehokkaamminkin käytössä huomioiden tietovarantojemme laadukkuus.

KUVA: ACB / FLICR.com | CREATIVE COMMONS

Johdanto

Julkinen hallinto omistaa ja hallinnoi laajoja, rahallisesti ja yhteiskunnallisesti arvokkaita informaatiovarantoja. Suurin osa näistä raakamuotoisena datana olevista varannoista on tällä hetkellä vain niiden käytössä, joilla on pääsy aineistoihin. Datan jatkojalostukseen päätyy arvioiden mukaan vain pieni osa käytettävistä olevasta. Internetin kehitys niin yhteisöllisiltä kuin teknologisilta ominaisuuksiltaan antaa uusia mahdollisuuksia avoimeen datapolitiikkaan. Kansallisten tietovarantojen maksuttoman käytön salliminen ja sitä tukeva toimintakulttuurin muutos ovat merkittävimpiä asioita, joita suomalaisen tietoyhteiskunnan kehittämiseksi voitaisiin tällä hetkellä toteuttaa. Tämä toisi uusia mahdollisuuksia vastata ajankohtaisiin haasteisiin. Maksuton julkinen data olisi myös suomalaisille instituutioille mahdollisuus uudistua ja oppia verkostoaikakauden yhteistoimintataitoja sekä suunnata suomalaisen yhteiskunnan kehittämistä uudenlaiseen yhdessä tekemisen kulttuuriin.

Viime aikoina ajatukset sallia julkishallinnon datan käyttäjiksi niin yksityiset yritykset, tutkimuslaitokset kuin kaikki muutkin kiinnostuneet tahot on saanut yksimielistä tukea. Keskustelu aihepiiristä on lisääntynyt meillä omien kansallisten tavoitteidemme, mutta myös kansainvälisten esimerkkien innoittamana.

Julkisen datan maksuttoman saatavuuden puolesta ovat viime aikoina esiintyneet monet tahot (ks. LVM 2009, Paukku 2009, Turkki 2009). Myös Arjen tietoyhteiskunnan neuvottelukunta on ottanut kantaa sen puolesta, että julkinen data saatettaisiin avoimesti ja laajasti kansalaisten ja elinkeinoelämän käytettäväksi (Neuvottelukunta 2009). Helmikuussa 2010 liikenne- ja viestintäministeriö ja valtiovarainministeriö asettivat kolme työryhmää selvittämään, miten julkisen informaation saatavuutta parannettaisiin ja kuinka tiedon käyttöä julkisen ja yksityisen sektorin kesken voitaisiin edistää (LVM & VM 2010).

Tätä kirjoitettaessa ajankohtaisin kannanotto avoimen datan puolesta on valtiovarainministeriön työryhmältä, joka on pohtinut tilastotoimen tehostamista ja alueellistamista. Työryhmän esittää useita toimenpiteitä, joiden tavoitteena on helpottaa julkisin varoin luotujen tietovarantojen hyödyntämistä. Parantamisalueina työryhmä nostaa esiin kuntatilastot sekä ympäristö- ja luonnonvaratilastot (VM 2010).

Enää ei nähdäksemme ole tarvetta periaatekeskusteluun avoimesta datasta, vaan erityistä painetta kohdistuu jo tietotaidon lisäämiseen ja käytäntöjen luomiseen orientoivalla opastuksella.

Avoim data monialaisen hyödyn lähteenä

Suomessa on hallinnon avoimeen ja maksuttomaan dataan liittyvässä keskustelussa tähän asti viitattu ensisijaisesti liike-elämän hyötyihin ja kansalliseen kilpailukykyyn. Markkinatalouden kärkimaassa Yhdysvalloissa julkista dataa koskeva lainsäädäntö on jo pitkään perustunut ajatukselle, että verovaroin tuotetun datan tulee olla maksutta kaikkien saatavilla. Tämä lähtökohta on johtanut merkittävään julkisen datan uudelleenkäyttöön ja taloudelliseen toimeliaisuuteen. Suomessa, kuten muuallakin Euroopassa, lähtökohta on ollut varsin erilainen ja suhtautuminen datavarantoihin perinteisesti protektionistista. Edellytykset uudentylaiselle ajattelulle kasvavat kuitenkin jatkuvasti, kun avoimet toimintamallit kehittyvä ainakin osaksi internetin ansiosta.

Viranomaiset keräävät dataa ensisijaisesti omaa käyttöön varten. Arvokkaat uudelleenkäytettävät datavarannot syntyvät muun toiminnan sivutuotteena. Useissa tässä oppaassa esitellyistä esimerkeistä parhaat käyttötavat ovat muotoutuneet tuottajaorganisaation ulkopuolella ennalta arvaamattomissakin käyttöympäristöissä ja toisiin datalähteisiin yhdistettyinä. Monet näistä esimerkeistä ovat yllätyksellisiä niin toimijoiden kuin yhteistyökumppaneidenkin osalta.

Julkisten datavarantojen avaaminen edistää samanaikaisesti ainakin kolmea eri tavoitetta: demokratian ja hallinnon läpinäkyvyyden lisäämistä, uusien markkinoiden ja innovaatioiden syntymistä ja hallinnon sisäisen tehokkuuden lisäämistä. Näitä alueita esittelemme oppan eri luvuissa tarkemmin asiayhteyksiin sitoen. Taloudellisten hyötyjen ohella uudenlainen toimintapolitiikka vahvistaisi osaltaan myös identiteettiämme, ts. antaisi pontta pohdintoihin siitä, miten me tulevaisuudessa elämme nettiakakaudella ja mille asioille kansallisesti tuottamamme informaatio vastaisuudessa rakentuu ja miten hyödynnämme yhdessä tuottamaamme dataa?

Avoim data ja datan koneluettavuus virtaviivaistavat prosesseja, mutta tätä mahdollisuutta ei ole vielä käytetty kovinkaan laajasti. Liikenne- ja viestintäministerin asettama selvitysmies Paul Pauku kritisoi hallintoa kyvyttömyydestä yhteistyöhön liike-elämän kanssa: “Not invented here-syndrooma: valtio rakentaa tyhjistä sitä mitä markkina jo käyttää” (2009). Avoin informaationvälitys ja läpinäkyvyys mahdollistaisivat laajemman näkemyksen, jolla voidaan edistää yhteistyötä. Julkisen datan maksuttomalla jakamisella voidaan välttää myös tahatonta osaoptimointia eli sitä, että virastot säästääkseen omista kustannuksista eivät esimerkiksi osta palveluna väestötietojärjestelmän tarjonnasta, vaan ylläpitävät itse tietokantaa omista asiakkaistaan.

Oppaan sisältö

Toistaiseksi maassamme ei ole ollut avoimen julkisen datan hyödyntämiseen opastusta tai käytäntöjä. Tämä opas pyrkii poistamaan tätä puutetta. Se on käytännönläheinen johdatus ensisijaisesti julkishallinnon työntekijöille, jotka haluavat edistää datan avaamista omassa organisaatiossaan, mutta on käyttökelpoinen myös kaikille muille aihepiiristä kiinnostuneille.

Avoin julkinen data on aiheena ja ilmiönä haastava tarkastelukohde. Siksi siitä on sekä mielenkiintoista että vaikeaa tehdä opastusta. Tässä oppaassa olemme päätyneet seuraavanlaiseen järjestykseen. Alkuun esittelemme aihepiiriä ja siihen liittyviä käsitteitä, toimintamalleja ja ominaisuuksia sekä luomme ideaalitalannekuvaa siitä, mitä avoimen datan hyödyntäminen parhaimmillaan voisi tuottaa maallemme. Seuraavaksi tarkastelu siirtyy enemmän konkreettiselle tasolle ja pohdimme avoimeen datan jakamiseen liittyviä lainsäädännöllisiä, taloudellisia ja teknisiä kysymyksiä. Lopuksi esitämme suosituksia, joita noudattamalla keskustelu ja toiminta avoimen datan alueella etenisi ja tulisi myös käytännöksi.

Yksi oppaan perusteemoista käsittelee julkishallinnon tuottaman datavarannon avaamista maksuttomasti kaikkien käytettäväksi niin, että dataa voidaan jalostaa koneluettavassa muodossa. Opas esittelee työkalujen ja runsaiden tapausesimerkkien kautta mitä kaikkea on hyvä huomioida omia datavaroja avattaessa. Käsittelemme myös hallinnon ja käyttäjäyhteisöjen välistä yhteistyötä, jolla voidaan tukea datan uudelleenkäyttöä ja näin lisätä sen yhteiskunnallista ja kansantaloudellista vaikuttavuutta.

Oppaaseen on myös tiivistetty keskeisimmät kansainväliset opit suomalaisen toimintaympäristöön ja kulttuuriin sovitettuina ajatuksella, että osaa niistä voitaisiin soveltaa heti käytännössä. Yksityiskohtaista opastusta käytännön sovelluksiin on kuitenkin vaikea antaa, sillä hyvät käytännöt ovat monilta osin vasta muotoutumassa. Opas toivottavasti viitoittaa tietä ensimmäisiin kokeiluihin, käytäntöjen kehittämiseen ja toimii keskusteluun virittäjänä.

Oppaan sisällössä on nostettu esille niin ikään keskeisiä kysymyksiä datan teknisestä avaamisesta, käyttöönoton arvioinnista ja yhteisöjen kanssa työskentelemistä. Aihepiirin laajuuden vuoksi käsittelemme yksittäisiä asioita jokseenkin yleisellä tasolla, mutta opas tarjoaa asioista kiinnostuneille runsaasti lähteitä teemojen syventämiseen. Olemme rajanneet käsittelystä monia avoimesta datasta kumpuvia erittäin kiinnostavia teemoja, kuten tietoturvan, hallinnon it-arkkitehtuurin, digitaalisen kuilun, rekisterien käytön, dokumentinhallinnan ja yleisen tekijänoikeus -keskustelun.

Johdanto ja ensimmäinen luku antavat taustaa aiheeseen yleisesti tutustuvalla lukijalle. Johdanto esittelee oppaassa käyttämäämme termistöä. Lisäksi johdannossa esitämme lukijan orientoimiseksi viisi tavoitetta Suomen tahtotilaksi ja hallinnon tietovarantojen avaamiseksi lähitulevaisuudessa. Näitä tavoitteita toivon mukaan kansantajuistavat niissä esiteltyt suomalaiset innovaatiot.

Tavoitteita ei välttämättä tarvitse lukea, mikäli on enemmän kiinnostunut varsinaisesta opastuksesta. Tavoitteet johdattelevat avoimen datan laajempiin viitekehyksiin.

Ensimmäinen luku: *Elävä palvelu* jäsentää organisaation suhtautumista yhteistyöhön, verkostoihin sekä muuhun oman organisaation ulkopuoliseen toimintaan, kun tavoitteena on mahdollistaa julkishallinnon dataan pohjautuvien uusien syntyminen.

Luvut 2, 3, 4 ja 5 avustavat organisaatiota muodostamaan omaa näkökulmaa avoimeen dataan. Lukujen sisältö on koottu tapausesimerkkien sekä haastattelujen kautta. **Luku 2:** *Organisaation näkökulmia avoimuuteen* esittää viitteellisen prosessin organisaation omien tietovarantojen inventointiin ja avaamiseen, sekä työkaluja aineistojen avoimuuden arviointiin eri näkökulmista (käyttöehdot, maksuttomuus, koneluettavuus, löydettävyyys jne.). Teemoja syvennetään juridiikan ja tekniikan osalta seuraavissa luvuissa. **Luku 3:** *Lupa julkaista ja uudelleenkäyttää* esittelee yhteenvetomaisesti joukon keskeisimpiä lakeja sekä asetuksia, joita on hyvä huomioida. **Luku 4:** *Avoimen datan taloudesta* kokoaa viime vuosien keskustelua siitä, miten hallinnon datan tuotantoa ja jakelua rahoitetaan nykyään, miksi siihen on päädytty Suomessa ja miten sitä voitaisiin rahoittaa. **Luku 5:** *Tekninen valmistautuminen* tuo esiin avoimen datan keskeisen ulottuvuuden; sen saattamisen tarjolle automatisoidusti ja ohjelmallisesti käsiteltävässä muodossa.

Luku 6: *Avoimen datan infrastruktuuri* esittelee kansallista avoimen datan infrastruktuuria, kuten datakatalogeja ja poikkihallinnollisia toimijoita, jotka auttavat muita toimijoita datavarantojen avaamisessa ja hyödyntämisessä.

Luku 7: *Suosituksia Suomelle* esittää laajempia toimenpidekokonaisuuksia, joilla voitaisiin tukea kansallista avoimen datan infrastruktuurin luomista ja käytäntöjä. Kehitys lähtee yksilöistä ja organisaatioista, mutta kaikkea yksittäiset toimijat tai organisaatiot eivät voi järkevästi edistää, vaan tarvittaisiin kansallisen tason linjanvetoja. Oppaan loppuun olemme koonneet hyväksihavaittuja liitteitä ja työkaluja.

Data käsitteenä

Suomen kielessä termien data, informaatio (information) ja tieto (knowledge) arkikäytössä ei ole useinkaan selvää eroa. Kotimaisessa puheessa tiedolla viitataan yleensä ja tarkemmin erittelemättä niin dataan, informaatioon kuin tietoonkin. Toisaalta tiedolla on suomenkielessä - muista kielistä poiketen - myös arvollinen merkitys. Esimerkiksi 'tietoyhteiskunta' sisältää enemmän sivistykseen ja kulttuuriin liittyviä latauksia kuin muissa kielissä neutraalimpi 'informaatioyhteiskunta'. Ihminen voi olla tiedoton monessa mielessä. Datan informaation ja tiedon ohella asiaan liittyviä käsitteitä ovat sisältö ja teos, joihin liittyvät muun muassa tekijänoikeudet.

Informaatiotutkimuksen piirissä datalla tarkoitetaan merkeistä ja symboleista koostuvaa potentiaalista informaatiota. Data on raaka-ainetta, josta voidaan jalostaa merkityksellisempää informaatiota, jota puolestaan voidaan välittää edelleen (Vakkari 1999). Keskeisessä roolissa on informaation tulkinta. Informaatio muodostuu tiedoksi kun vastaanottaja tulkitsee saamansa informaation ja antaa sille merkityksen siten, että se muuttaa tulkitsijan käsityksiä jostakin asiasta. Informaatiotutkimuksessa puhutaan yleisesti tiedon arvoketjusta (value chain of information), jossa irrallinen informaatio jalostuu ihmiselle käytännössä hyödylliseksi.

Kaavio. Tiedon arvoketju (Haasio, Savolainen 2004).

DATA -> INFORMAATIO -> TIETO -> TIETÄMYS -> VIISAUS

Tieto- ja informaatiokäsitteet voidaan erottaa toisistaan. Niiniluodon (2000) mukaan informaatio on yläkäsite ja tieto sen alakäsitteenä tarkoittaa tarkempaa perusteltavuuteen ja totuudenmukaisuuteen liittyvää vaatimusta. Karvosen (2000) mukaan arkikielessä tieto ja informaatio ovat melkein sama asia, mutta tieteellisessä käytössä niiden välillä on suuri ero. Informaatiosta voidaan hänen mukaansa erottaa ontologisia tasoja, joista vain ylimmät ovat semanttista informaatiota, tietoa. Data-käsite sen sijaan viittaa raaka-aineeseen, joka voi olla hyvin erimuotoista. Usein data arkikielessä viittaa tietoteknisiin järjestelmiin tallennettuun informaatioon. Informaation ja datan välinen ero on hyvin suhteellinen.

Käytämme oppaassa pääasiassa termiä data (ja siitä muodostettuja yhdysanoja, kuten datalähde, datantuottaja, datakatalogi jne.) tarkoittamaan digitaalisessa muodossa olevaa raaka-ainetta. Käytännössä koko suomalainen alan keskustelu on kuitenkin puhunut tietovarannoista 1990-luvulta alkaen, joten käytämme myös termiä tietovarannot kun viittaamme yleiseen keskusteluun, sekä aiempiin raportteihin ja selvityksiin.

Datalla tarkoitetaan digitaalisesti tallennettua informaatiota, kuten dokumentteja, tietokantoja sopimuksista, transkripteja kuulemisista ja audio- sekä visuaalisia tallenteita tapahtumista. Vaikka ei-elektronisessa muodossa olevat resurssit, kuten fyysiset esineet (esimerkiksi vanhat paperidokumentit), eivät kuulu julkishallinnon avoimen datan piiriin, niin tällaisten resurssien saattaminen mahdollisuuksien mukaan digitaaliseen muotoon on kannustettavaa.

Oppaassa neuvotaan julkisen datan avaamiseen siinä nimenomaisessa mielessä, että data on raaka-ainetta, jota eri toimijat voivat muokata, yhdistellä, suodattaa ja jalostaa. Yhden toimijan tuottama lopputulos voi olla potentiaalisesti raaka-ainetta jollekin muulle toimijalle. Esimerkiksi Finlexin sisältämä ajan tasainen lainsäädäntö ja suomi.fi-portaalin toimitettu sisältö ovat jalostettua informaatiota, mutta koneluettavassa muodossa ne voisivat olla raaka-ainetta sovellettukehittäjille.

Korostamme data-käsitteen avulla myös informaation tarjoamista digitaalisessa tietokoneiden avulla käsiteltävässä muodossa. Esimerkiksi Euroopan

tietoverkkorikollisuutta koskevassa yleissopimuksessa data tarkoittaa tosi-seikkojen, tietojen tai käsitteiden esitystä sellaisessa muodossa, että se soveltuu käytettäväksi tietojärjestelmässä, mukaan lukien ohjelmat, joiden avulla tietokone pystyy suorittamaan jonkin toiminnon (TEM 2009).

Datan ymmärtäminen raaka-aineeksi liittyy läheisesti seuraavaksi käsiteltävään ekosysteemi-käsitteeseen.

Data ekosysteemin osana

Hallinto tuottaa jatkuvasti suuren määrän dataa, joka voisi olla tehokkaammin käytössä huomioiden tietovarantojemme nykyinen laadukkuus. Asiaa lähestytään hahmottelemalla suomalaista avoimen datan ekosysteemiä. Aikaisemmissa keskusteluissa kansalaisia, yksityistä sektoria tai muita organisaatioita ei ole pidetty osapuolena tietovarantojen kehittämisessä, vaan ne on nähty ainoastaan tiedon loppukäyttäjinä. Hallinnon ohella myös kansalaiset, liike-elämä, järjestöt ja tutkimuslaitokset ovat merkittäviä toimijoita avointa dataa käsiteltäessä. Yksityisten toimijoiden tietovarannot eivät ole julkisia, mutta se ei estä näitä halutettassaan tuottamasta avointa dataa. Oppaassa esittämämme toimintatavat, työkalut, periaatteet ja suositukset ovat hyvin sovellettavissa kaikissa niissä organisaatioissa, kuten yrityksissä ja järjestöissä, jotka haluavat lisätä omien datavarantojen avointa käyttöä.

Oppaan yleistavoitteena on osaltaan edistää toimivan avoimen datan ekosysteemin syntymistä Suomeen. Siinä kansallisten tietovarojen hyödyntäminen lisääntyisi merkittävästi avoimuuteen pohjautuvan yhteistyön ja teknologian avulla. Ekosysteemillä tarkoitetaan yleisesti tietyn alueen toiminnallista kokonaisuutta. Voidaan puhua esimerkiksi tietyn järven tai metsän ekosysteemistä, joka koostuu alueen eliöiden ja elottomien ympäristötekijöiden dynaamisesta vuorovaikutuksesta. Avoimen datan yhteydessä ekosysteemi on monitasoinen ja -ulotteinen kokonaisuus, jossa raakadata on jakamisen ja jalostuksen osalta yhteistyön kohde. Raakadatala ei käydä kauppaa, vaan se saatetaan helposti kaikkien toimijoiden saataville.

Datan hankkimisesta vapautuneet resurssit voidaan täysimittaisesti kohdistaa sen hyödyntämiseen. Laajamittainen datan hyödyntäminen tuottaa uusia palveluita, tutkimusta ja tietoa, josta osalla on kaupallista arvoa ja osa edistää demokratiaa, sivistystä ja ihmisten arkea ilman suoria taloudellisia kytköksiä. Lisääntynyt datan hyödyntäminen heijastuu positiivisesti datan tuottamiseen ja lisää jatkuvasti datavarantojen laatua ja käytettävyyttä. Ekosysteemissä hallinnon organisaatiot, kansalaiset ja yritykset ovat sekä datan käyttäjiä että myös sen tuottajia.

Internetin ja tietotyön yleistymisen myötä perinteisten liiketoimintamallien ja rahatalouden ulkopuolisen tuotannon merkitys on kasvanut, joita ilmentävät muiden muassa toimintamallit kuten open source, wikipedia ja sosiaalinen media. Oppaassa ajattelemme datan keräämistä, jalostusta, julkaisua ja

uudelleenkäyttöä kokonaisuutena ja eri tekijöiden välisenä vuorovaikutuksena, eikä vain markkinoina ja vaihtokauppana. Ekosysteemiin liittyy mielikuva kokonaisuuden hyvinvoinnista ja toisaalta omien tarpeiden tyydyttämisestä osittain ekosysteemin rikkauden ja elinvoimaisuuden kautta.

Toinen käyttökelpoinen käsite aihepiirin kokonaisvaltaiseksi ymmärtämiseksi on avoimen datan infrastruktuuri, joka niin ikään käsittää kaikki avoimen datan kanssa toimivat organisaatiot ja järjestelmät, toisin sanoen koko toimintaympäristön. Tämä malli sopii avoimen datan kentän jäsentämiseen esimerkiksi valtiollisella ja kunnallisella tasolla. Esimerkiksi laki paikkatietoinfrastruktuurista (ks. luku 3) kuvaa hyvin ajattelua, jossa ymmärretään tietoinfrastruktuurin perusaineiksena: ”Paikkatietoinfrastruktuurilla tarkoitetaan saataville asetettuja metatietoja, paikkatietoaineistoja ja paikkatietopalveluja, verkkopalveluja ja -teknologioita, tietojen luovuttamista, saatavuutta ja käyttöä koskevia sopimuksia sekä koordinointi- ja seurantamekanismeja.” Hallinnon rooli olisi tuottaa infrastruktuuria kaikkien käytettäväksi ja näin toimia mahdollistajana datan laaja-alaiselle hyödyntämiselle.

Ekosysteemillä korostamme teknisen järjestelmän ja institutionalisoitujen organisaatioiden ohella myös elävää, dynaamisesti muuttuvaa vuorovaikutusverkostoa, jonka osana hallinnon organisaatiot ja yksittäiset ihmisetkin ovat. Infrastruktuuri-ajattelu on oppaassa taustalla ja siihen tartutaan vahvemmin luvussa 6: Avoimen datan infrastruktuuri, jossa käsitellään organisaatorajat ylittäviä avoimen datan hankkeita, kuten kansallinen datakatalogi.

Tahtotila: Julkinen data avoimesti saataville ja hyödynnettäväksi

Datan avaamiseen voi suhtautua hyvin käytännönläheisesti. Oppaan taustaineiston sekä haastattelujen inspiroimina haluamme kuitenkin tuoda esiin laajemman näkökulman datan avaamiseen ja sen mahdollisuuksiin Suomen kannalta. Oheista visiota konkretisoivat oppaan lopuksi kootut toimenpideehdotukset. Vision tarkoituksena on edistää mielikuvaa siitä, mihin kaikkeen avoimen datan ajattelu voisi johtaa Suomessa. Visiomme on tarkoitettu kannustamaan lukijaa avoimen datan edistämiseen yleisellä tasolla, jotta aikanaan voitaisiin myös poliittisten linjausten avulla luoda toimenpideohjelman tässä esitettyjen ja muiden tavoitteiden saavuttamiseksi.

Netinkäyttäjistä harvat miettivät, miten internetin hakukoneiden löytämä informaatio on tuotettu tai miten se on julkaistu verkkoon. Internetin kehitysterendeistä voi kuitenkin lukea, että yhä useammat ihmiset ja organisaatiot ovat alkaneet ajatella, että suuren osan omassa hallussa olevaa ja kertaalleen tuotettua dataa voisi tarjota maksuttomasti muille eikä se olisi keneltäkään pois.

Avoin data on meistä ensisijaisesti lähestymistapa ja asenne samalla tavalla kuin viime vuosina paljon käytetyt Web 2.0 ja sosiaalinen media -termit. Ne eivät ole yksiselitteisiä, kirjasta tai koulutuspäivästä omaksuttavia suoraviivaisia välineitä tai toimintaohjeita. Pikemminkin ne ovat kehityspolkuja, joita

voi ottaa vähitellen käyttöön ja joiden avulla voi kehittää itselleen sopivia toimintarooleja tai soveltaa niitä organisaatioiden toimintaan (esim. Hintikka 2009; Kuntaliitto 2010). Samalla tavalla avoin data on yksi, mutta monia uusia mahdollisuuksia tuova lähestymistapa internet-aikakauteen.

Miten sitten pitää edetä ja toimia, jotta avoin datapolitiikka olisi meillä mahdollista? Olemme luoneet tulevaisuuskuvan tahtotilasta, joka on mahdollinen toteuttaa maassamme jo olevien resurssien, käytäntöjen ja toimintaympäristön ominaisuuksien avulla. Tarkastelemme seuraavaksi tätä tahtotilaa, joka sisältää viisi tavoitetta ja joita esittelemme esimerkkien avulla.

Tavoite 1: Uusia ratkaisuja kehittävät koulutetut suomalaiset

Yhteiskunnan positiivisen muutoksen tekijöitä ovat informoidut yksilöt. He tietävät, mitä yhteiskunnassa tapahtuu ja pystyvät ilmaisemaan itseään ja vaikuttamaan ympäristöönsä sekä yksilöinä että verkostojen ja järjestöjen osana. Julkishallinnon datan avoin jakelu lisää läpinäkyvyyttä ja erilaisten toimijoiden mahdollisuuksia osallistua yhteiskunnan kehitykseen.

Internetin yleistymisen myötä yksittäisten ihmisten mahdollisuudet saada tietoa yhteiskunnallisista ongelmista ja keskustella niiden ratkaisemisesta ovat parantuneet, esimerkiksi bloggaamisen ansiosta. Verkostoituneet yksilöt ovat osoittaneet kykenevänsä laajamittaisiin yhteistyöprojekteihin, joista on merkittävää päivittäistä hyötyä, kuten Wikipedia-tietosanakirjasta. Vähähiilisyttä edistävä Porkkanamafia-kuluttajaliike ja sähkömoottoreita bensa-autoihin asentava Sähköautot - Nyt! ovat onnistuneet tarttumaan ilmastomuutoksen uhkaan tavalla, joka muille toimijoille on ollut vaikeaa. Avoimen datan ympärille saattaa kehittyä samanlaisia ongelmanratkojayhteisöjä, sillä korkean koulutustason ansiosta meiltä löytyy paljon tietotyöhön tottuneita ihmisiä.

Kahden vuoden välein EU ministerit kokoontuvat sopimaan ministeritason sähköisen hallinnon julistuksesta (Ministerial Declaration on e-government). Se on keskeinen Eurooppalainen strateginen asiakirja. Yleensä tätä julistusta seuraa myös teollisuuden julistus. Vuoden 2009 marraskuussa EU ministereiden e-government kokouksessa Malmössä esiteltiin myös kansalaisten avoimesti sosiaalisen median välineillä koostama julistus (Eups20 2010). Tämä julistus syntyi aktiivisten kansalaisten aloitteesta, joukkovoimalla netin yhteisöllisillä ideointi- (UserVoice) ja kirjoitusalueilla (MixedInk). Myös sen suomentaminen tapahtui vapaaehtoisvoimin Suomalaisissa sosiaalisen median yhteisöissä (alustoina Qaiku, Twitter ja Etherpad). Olemme liittäneen julistuksen oppaan liitteeksi (liite 1) havainnollistamaan sitä kuinka aktiivisimmat kansalaiset kaipaavat yhteistyön tiivistämistä hallinnon kanssa.

Case: Asialista.fi

Asialista.fi on valmisteluvaiheessa oleva yhteisöllinen verkkoprojekti. Palvelu kokoaisi kaikki yksittäiseen kunnallishallinnon

päätöksentekoprosessiin liittyvät dokumentit ja hajallaan olevan tiedon yhdeksi kokonaisuudeksi wiki- ja blogi-ohjelmistoilla. Verkosto vapaaehtoisia aktiiveja voisi kirjoittaa yleiskielisiä tiivistelmiä kaupunginhallitusten asialistoista linkittäen niitä muihin verkossa oleviin aineistoihin. Tavoitteena on vähentää kunnallisen päätöksenteon seuraamiseen liittyviä aikakustannuksia (ks luku 4.5) ja vapauttaa kuntalaisten aikaa poliittiseen keskusteluun. Kun yksi ihminen kirjoittaa tiivistelmän yhteiseen palveluun, vapauttaa se aikaa muilta toisiin tehtäviin.

Palvelun kehittäjät ovat käyneet keskustelua kaupunkien tietotekniikkavastaavien kanssa koneluettavan rajapinnan avaamisesta asianhallintajärjestelmiin. Tällöin tietojen kerääminen Asialista.fi-palveluun voitaisiin automatisoida mahdollisimman pitkälle.

Tavoite 2: Talkoot ja yhteistoiminnan perinne

Avoin data auttaa laajentamaan informaatiopohjaisten hyvinvointipalveluiden tuotantoa hallinnon ulkopuolella. 2010-luvun hyvinvointiyhteiskunta on osin aktiivisten ja voimaantuneiden kansalaisten ja yhteisöjen itselleen ja toisilleen tuottama ja julkishallinto toimii mahdollistajana tarjoten maksuttomia dataresursseja.

Suomi on hyvinvointivaltio, jonka perustuslaissa on kirjattu valtiolle velvollisuus taata kansalaisille ihmisoikeuksien ja poliittisten oikeuksien lisäksi myös sosiaaliset oikeudet välttämättömään elintasoon ja huolenpitoon. Hyvinvointi ymmärretään Suomessa yksimielisesti käsittämään sekä henkisen, sosiaalisen, taloudellisen että ympäristön hyvinvoinnin. Ihmisten hyvinvoinnista huolehtiminen on synnyttänyt Suomeen laajan julkisen sektorin, joka suunnittelee ja suurelta osin toteuttaa yhteiskunnan infrastruktuuria ja palveluja. Hallinnon tuottaman datan osalta julkinen sektori on usein ainoa taho, joka jalostaa datasta arkea helpottavia ja hyvinvointia tukevia palveluita.

Suomessa on vahvoja yhteis- ja ilmaistoiminnan perinteitä, kuten talkookulttuuri, jokamiehenoikeus, kirjastolaitos sekä Linux- ja MySQL-ohjelmointihankkeet. Niitä voitaisiin hyödyntää nykyistä laajemmin myös kansalaisten ja julkishallinnon välisessä yhteistyössä ja julkisten palveluiden kehittämisessä (Mokka & Neuvonen 2006). Avoin data on yksi konkreettinen mahdollisuus saattaa yrityksiä, kansalaisia ja järjestöjä yhteen ja osaltaan tuottamaan yhteiskunnan hyvinvointipalveluita yhdessä julkisen sektorin kanssa.

Luvussa 1: Elävä palvelu käsittelemme hallinnon mahdollistajaroolia ja keinoja avoimeen yhteistyöhön

Case: Ymparisto.fi

Suomen luonnon seurantaan ja sen raportointiin käytetään nykyisin viranomaistyötä noin 60 henkilötyövuotta. Vapaaehtoisten raportojien ja harrastajien joukkouttuneeksi (luku 3) työpanokseksi on arvioitu noin 200 henkilötyövuotta.

Luonnosta kiinnostuneiden ihmisten ja hallinnon edustajien vuorovaikutus toimii molempia osapuolia hyödyttäen ja havaintojen varmentaminen huomioiden. Ihmiset tuottavat raakadataa, jonka hallinnon työntekijät tarkistavat ja jakavat arvioinnin ja korjaamisen jälkeen verkossa informaationa, johon mukana olevat vapaaehtoiset voivat reagoida ja kerätä sen pohjalta uutta raaka-ainetta tarjolle. Vastaavaa menetelmää käyttää muun muassa amerikkalainen tartuntatautien ehkäisykeskus Centers for Disease Control and Prevention (CDC). Tässä toimintatavassa kansalaiset lähettävät havaintojaan esimerkiksi sikainfluenssatartunnoista ja CDC Emergency välittää niitä arvioinnin ja tarkistamisen jälkeen muun muassa Twitter-mikrokanavalla.

Tavoite 3: Laadukkaat tietovarantomme kansainvälisen profiloitumisen perustaksi

Suomalaiset tietovarannot ovat niin laadukkaita, että niitä esittelemällä ja hyödyntämällä on mahdollista profiloitua erityysosaajiksi maailmanlaajuisesti. Positiivinen maine antaa suomalaisille huippuosaajille ja organisaatioille mahdollisuuden toimia kansainvälisinä asiantuntijoina ja yhteistyötahoina kansainvälisissä hankkeissa.

Suomeen on aikojen kuluessa syntynyt laadullisesti tarkastellen maailmanluokan tietovarannot ja niiden potentiaalinen voima saadaan tehokäyttöön, kun niitä hyödynnetään mahdollisimman laajasti. Julkisen datan joustavaan tuotantoon, jakeluun, ylläpitoon ja käyttöön tarvitaan kansallisesti tekninen, juridinen ja organisatorinen infrastruktuuri. Tässä kehityksessä Suomi toivottavasti ottaa harppauksia eteenpäin.

Perusedellytys laadukkaiden tietovarantojen ylläpidolle on myös, että dataa keräävillä ja ylläpitävillä viranomaisilla on pysyviä resursseja. Vähintään yhtä tärkeää on, että ihmiset luottavat viranomaisiin ja siihen, että viranomaiset kunnioittavat tietosuojaa ja siihen liittyvää lainsäädäntöä. Luottamus auttaa kansallisten tietovarantojen ylläpidossa, koska ihmiset ja yritykset toimittavat rekisterinpitäjille kysytyt tiedot. Avoimuudesta ja maksuttomuudesta puhuttaessa monet tahot esittävät oikeutetusti huolensa korkean laadun säilymisestä. Resurssointi on taattava siinäkin tapauksessa, että dataa luovutetaan maksuttomasti. Avoin ja laaja jakelu ei saa johtaa datan tuottajan luotettavuuden heikkenemiseen esimerkiksi siten, että yksityisyydensuoja vaarantuu.

Luvussa 6: Avoimen datan infrastruktuuri käsittelemme keskeisiä teknisiä ja organisatorisia muutoksia, joilla Suomen tietoinfrastruktuuria voidaan kehittää.

Case: Väestölaskenta

Väestölaskenta on yksi esimerkki Suomen laadukkaista tietovaranoista. Väestölaskennan tiedot sisältävät asukasluvun lisäksi dataa esimerkiksi koulutuksesta ja työssäkäynnistä. Monissa maissa laskenta toteutetaan YK:n suositusten mukaan kymmenen vuoden välein lomakkeilla. Suomessa väestölaskenta tehdään vuosittain yhdessä yössä nappia painamalla ja ajamalla 37 rekisteriä yhteen. Tämän jälkeen väestölaskennan tiedot ovat hyvin nopeasti käytettävissä kaikille niitä tarvitseville organisaatioille hallinnon sisällä. Samaan pystyvät myös Tanska, Norja ja viimeisimpänä tulokkaana Ruotsi. Myös Hollannilla, Kanadalla, Australialla ja Uudella-Seelannilla on hyvä tietoinfrastruktuuri. Esimerkiksi Yhdysvalloissa, joka on julkishallinnon tietovarantojen avoimen jakamisen esimerkkimaa, rekisterit ja tietovarannot eivät ole läheskään yhtä laadukkaita kuin Suomessa.

Tavoite 4: Dataa myös koneille

Datan avaamisesta käydyissä keskusteluissa koneluettavuuden ja avoimen lisensoinnin merkitystä on usein aliarvioitu. Datan koneluettavuuteen ja avoimuuteen panostamalla saataisiin jo nykyisin julkiset tietovarannot huomattavasti laajemmin ja monipuolisemmin käyttöön. Varsinaisen datan tarjoamisen rinnalla koneluettavuus on keskeinen elementti kokonaisvaltaisten hyötyjen saamiseksi.

Julkista dataa on jo nykyäänkin saatavilla hallinnon organisaatioiden verkkopalveluissa varsin runsaasti. Suuri osa aineistosta on kuitenkin sellaisessa muodossa, etteivät koneet pysty käsittelemään sitä. Datan ohjelmallinen ja automatisoitu hakeminen, käsittely ja yhdistely toisiin aineistoihin on työlästä tai käytännössä mahdotonta.

Esimerkiksi vireillä olevien kaavahankkeiden esittäminen kaupungin kartta-palvelussa noudattaa julkisuuslakia. Useinkaan kaavojen paikkatieto ei ole saatavilla koneluettavassa muodossa, jolloin lähiympäristön kaavahankkeiden automaattinen seuraaminen ei ole mahdollista.

Luvussa: 5 Tekninen valmistautuminen käsittelemme koneluettavuuden toteuttamista

Case: ReittiGPS

Älypuhelimissa toimiva ReittiGPS-sovellus on hyvä esimerkki koneluettavuuden ja automatisoinnin eduista verrattuna toistuvaan aikataulun selaamiseen netissä. Sovellus yhdistää kaksi datalähdettä: gps-anturin kertoman sijainnin ja Reittiopas-palvelun

tietokannan tiedot joukkoliikenteen aikatauluista. Reittiopas-palvelun tietokannoissa on raakadataa julkisten kulkuneuvojen pysäkki-, reitti- ja aikataulutiedoista. Sovellus ottaa gps:sta paikkatiedon, asettaa sen lähtöpisteeksi ja kysyy käyttäjältä määränpäättä. Avoimen rajapinnan kautta ReittiGPS-sovellus hakee datan koneellisesti ja automatisoidusti kännykkään. Käyttäjälle kännykkäsovellus tarjoaa parilla napin painalluksella selkeän opastuksen, mille pysäkillä hänen tulisi kävellä ja mihin kulkuneuvoon nousta.

Tavoite 5: Pienestä liikkeelle

Parhaiten saavutetaan tuloksia, kun kehityshankkeissa etsitään taustaksi tietoa muiden jo toteuttamista samankaltaisista hankkeista, opitaan tehdyistä virheistä ja lähdetään liikkeelle pienin askelin valmiina korjaamaan suuntaa heti kun siihen ilmenee tarvetta. On hyvä muistaa myös, että jonkun on oltava pioneeri ja avattava tietä.

Linux-käyttöjärjestelmän alkutahdit antavat hyvän esimerkin siitä, miten globaalit innovaatiot voivat nykyään käynnistyä. Linus Torvalds viestitti internetissä vuonna 1991, että hän on kehittänyt jotakin vielä keskeneräistä, joka saattaa kiinnostaa muitakin. Hän halusi kuulla muiden kehitysideoita, mutta ei luvannut toteuttaa niitä. Viestin varovaisesta sisällöstä huolimatta siitä muodostui alkusysäys erittäin merkittäväksi kasvaneelle yhteistyöprojektille. Lähes kaksikymmentä vuotta myöhemmin, avoimen lähdekoodin maksuton Linux-käyttöjärjestelmä on kehittynyt kilpailijaksi Windowsin ja Macin OS X:n rinnalle. Maksuttomuudestaan huolimatta Linux on synnyttänyt merkittävää liiketoimintaa, jopa siinä määrin, että esimerkiksi IBM on luopunut oman käyttöjärjestelmänsä kehittämisestä, tarjoaa Linuxia ratkaisunaan ja edistää omaa liiketoimintaansa. Linuxin käyttöjärjestelmän ytimen arvoksi on laskettu miljardi euroa (García 2010).

Toivomme, että hallinnon organisaatiot samalla tavoin kertovat tietovarannoistaan, vaikka ne eivät vielä olisi valmiita avattaviksi ja ryhtyisivät vuoropuheluun niiden käytön lisäämiseksi.

Luvuissa 1: Elävä palvelu ja 2: Organisaation näkökulmia avoimuuteen käsittelemme ketterää toimintamallia datan avaamisen kontekstissa.

Case: Linux

Lainaus Linus Torvaldsin lähettämästä sähköpostista, jossa hän esittelee ajatuksiaan uuden käyttöjärjestelmän ohjelmoinnista. Alkuperäisen viestin teknisiä yksityiskohtia on jätetty tässä pois.

Aug 26 1991, 8:12 am

Hello everybody out there using minix – I’m doing a (free) operating system (just a hobby, won’t be big and professional like gnu) for 386(486) AT clones. This has been brewing since april, and is starting to get ready. I’d like any feedback on things people like/dislike in minix, as my OS resembles it somewhat...

.... Any suggestions are welcome, but I won’t promise I’ll implement them :-)

Linus (torvalds@kruuna.helsinki.fi <mailto:torvalds@kruuna.helsinki.fi>)