



VALTIOVARAINMINISTERIÖ

# Avoimen tiedon vaikuttavuus – esitutkimus

---

Valtiovarainministeriön julkaisuja – 15a/2015

# Kuvailulehti

Julkaisija ja julkaisu-aika	Valtiovarainministeriö, maaliskuu 2015	
Tekijät	Heli Koski, Elinkeinoelämän tutkimuslaitos	
Julkaisun nimi	Avoimen tiedon vaikuttavuus –esitutkimus	
Asiasanat	vaikuttavuus, julkinen hallinto, tietovaranto, tiedonkeruu	
Julkaisusarjan nimi ja numero	Valtiovarainministeriön julkaisuja 15a/2015	
Julkaisun myynti/jakaja	Julkaisu on saatavissa pdf-tiedostona osoitteesta <a href="http://www.vm.fi/julkaisut">www.vm.fi/julkaisut</a> . Samassa osoitteessa on ohjeet julkaisun painetun version tilaamiseen.	
Painopaikka ja -aika	Juvenes Print - Suomen Yliopistopaino Oy, 2015	
ISBN 978-952-251-667-1 (nid.) ISSN 1459-3394 (nid.) ISBN 978-952-251-668-8 (PDF) ISSN 1797-9714 (PDF)	Sivuja 36	Kieli Suomi

## Tiivistelmä

Avoim julkinen data on julkisen organisaation tuottama tai hallinnoima tieto, joka on konekielisessä muodossa ja maksutta kenen tahansa käytettävissä, muokattavissa ja jaettavissa sekä yksityisiin että kaupallisiin tarkoituksiin. Julkisten tietojen avaamisen potentiaaliset hyödyt on ennakkolisissa arvioinneissa arvioitu mittaviksi.

Tietovarantojen hyödyntämisen yhteiskunnallis-taloudellisiin kysymyksiin pureutuva tutkimus on vielä varsin vähäistä, ja datan hyödyntämiseen liittyvistä taloudellisista vaikutusmekanismeista ei ole organisaatiotasolla eikä kansantalouden tasolla juurikaan tutkimustietoa.

Suomi on julkisen tiedon avaamisen kärkimaiden joukossa, ja sillä on mahdollisuus myös nousta edistyksellisimpien maiden joukkoon avoimen datan vaikuttavuuden tutkimuksessa. Julkisen sektorin tietovarantojen avaamisen vaikutukset voidaan jakaa taloudellisiin vaikutuksiin ja muihin yhteiskunnallisiin vaikutuksiin. Julkisten tietovarantojen avaamisen toteutuneita yhteiskunnallis-taloudellisia vaikutuksia ei pystytä kattavasti arvioimaan olemassa olevan tai tällä hetkellä säännöllisesti kerättävän tiedon valossa. Arviointiin tarvitaan seurantamallin ja mittariston kehittämistä ja tämän jälkeen systemaattista tiedonkeruuta vaikutusten jälkikäteiseen mittaamiseen.

Tilastollisen vaikuttavuusmittareita hyödyntävän aineistoanalyysin lisäksi voidaan sekä kansalaisten, yritysten että julkisen sektorin organisaatioiden osalta arvioida tarjolla olevien julkisten tietovarantojen ja erilaisten aineistojen käytettävyyttä ja hyödyllisyyttä käyttäjien omien arvioiden perusteella. Lisäksi olisi syytä miettiä keinoja, miten tuotetaan tietoa avoimen datan hyödyntämiseen liittyvistä parhaista käytännöistä erityisesti julkisen sektorin organisaatioissa ja siihen, millä keinoin nämä parhaat käytännöt saadaan mahdollisimman tehokkaasti leviämään ja hyödynnetyksi eri organisaatioissa.

# Presentationsblad

Utgivare och datum	Finansministeriet, mars 2015	
Författare	Heli Koski, Elinkeinoelämän tutkimuslaitos	
Publikationens titel	Avoimen tiedon vaikuttavuus –esitutkimus	
Publikationsserie och nummer	Finansministeriet publikationer 15a/2015	
Beställningar/distribution	Publikationen finns på finska i PDF-format på <a href="http://www.vm.fi/julkaisut">www.vm.fi/julkaisut</a> . Anvisningar för beställning av en tryckt version finns på samma adress.	
Tryckeri/tryckningsort och -år	Juvenes Print – Finlands Universitetstryckeri Ab, 2015	
ISBN 978-952-251-667-1 (hft.) ISSN 1459-3394 (hft.) ISBN 978-952-251-668-8 (PDF) ISSN 1797-9714 (PDF)	Sidor 36	Språk Finska

## Sammandrag

Med öppen data avses data, som produceras eller administreras av offentliga organisationer, och som finns i maskinläsbar form och avgiftsfritt tillgängliga för vem som helst, för förädling och distribuering i såväl privat som kommersiellt syfte. De potentiella nyttorna med öppnandet av offentliga datalager är betydande.

Forskning som fördjupat sig i de samhällsekonomiska aspekterna med utnyttjandet av datalager har till dags dato varit tämligen begränsat, och det finns knappast någon forskningsinformation om de ekonomiska verkningarna av utnyttjandet av data varken på organisationsnivå eller på samhällsekonomisk nivå.

Finland befinner sig för tillfället i täten bland länderna som öppnat offentliga datalager, och vi har möjlighet att häva oss bland de allra mest avancerade länderna även när det gäller forskningen i genomslagenheten hos öppen data. Verkningarna av den offentliga sektorns datalager kan indelas i ekonomiska verkningar och andra samhällseliga verkningar. De faktiska samhällsekonomiska konsekvenserna av öppnandet av offentliga datalager kan inte utvärderas på ett täckande sätt i ljuset av befintlig eller regelbundet insamlad information. För utvärderingen behövs utveckling av uppföljningsmodeller och mätare, och dessutom systematisk informationsinsamling för efterhandsmätning av effekterna.

Förutom materialanalys som utnyttjar statistiska effektivitetsmätare kan man både för medborgares, företagens och offentliga organisationers del utvärdera tillgängliga datalagrars och olika materials användbarhet och nytthet utifrån användarnas egna bedömningar. Dessutom bör man begrunda medel genom vilka man kunde producera information om bästa praxis för utnyttjandet av offentlig data speciellt inom offentliga organisationer, och om hur dessa bästa praxis kan spridas till och utnyttjas inom olika organisationer på ett så effektivt sätt som möjligt.

## Description page

Publisher and date	maaliskuu 2015; Ministry of Finance, March 2015	
Author(s)	Heli Koski, Elinkeinoelämän tutkimuslaitos	
Title of publication	Avoimen tiedon vaikuttavuus –esitutkimus	
Publication series and number	Ministry of Finance publications 15a /2015	
Distribution and sale	The publication can be accessed in pdf-format in Finnish at <a href="http://www.vm.fi/julkaisut">www.vm.fi/julkaisut</a> . There are also instructions for ordering a printed version of the publication.	
Printed by	Juvenes Print – Finland University Print Ltd, 2015	
ISBN 978-952-251-667-1 (print.) ISSN 1459-3394 (print.) ISBN 978-952-251-668-8 (PDF) ISSN 1797-9714 (PDF)	No. of pages 36	Language Finnish

### Abstract

Various countries have implemented the open (government) data strategy aiming at providing wide access to government data in machine-readable format such that it can be freely used, reused and redistributed by anyone. Reported ex ante evaluations have estimated that the potential benefits of opening up public data resources are substantial.

Very little is known about the underlying economic and organizational mechanisms and implications of open data use at the organizational level or at the level of economy as a whole. To my best knowledge, there is no reported comprehensive country-level ex-post impact assessment of opening up government data.

Currently, Finland is among the leading countries in opening up the government data, and it has also a chance to be among the most advanced ones in the impact assessment of open data. The impacts of opening up government data can be divided to economic impacts and to other social impacts. The prerequisite for this is a careful development of the monitoring and evaluation model for opening up government data as well as a systematic gathering of data for the impact assessment.

Furthermore, the usability and usefulness of different public data resources for consumers, firms and public sector organizations can be assessed via the users' own evaluations. In addition, it is important to assess appropriate means to disseminate and promote efficient utilization of information on the best practices of open data re-use in different organizations.

# Avoimen datan vaikuttavuus – esitutkimus

Heli Koski, Elinkeinoelämän tutkimuslaitos

29.1.2015

## YHTEENVETO

Avoin julkinen data on mikä tahansa julkisen organisaation tuottama tai hallinnoima tieto(varanto), joka on konekielisessä muodossa ja maksutta kenen tahansa käytettävissä, muokattavissa ja jaettavissa sekä yksityisiin että kaupallisiin tarkoituksiin. Julkisten tietojen avaamisen potentiaaliset hyödyt on ennakkolisissa arvioinneissa arvioitu mittaviksi. Tämän takia julkisten tietovarantojen avaaminen on noussut useissa maissa keskeiseksi politiikkateemaksi. Suomen hallitus teki maaliskuussa 2011 periaatepäätöksen siitä, että julkisten tietojen avaaminen laajalti yhteiskunnan eri tahojen uudelleenkäyttöön on yksi hallitusohjelman kärkihankkeista. Toukokuussa 2013 valtiovarainministeriö asetti kaksivuotisen julkisten tietovarantojen avaamiseen mahdollisimman laajasti ja nopeasti tähtäävän avoimen tiedon ohjelman. Tavoitteena on avata kaikki merkittävät julkishallinnon hallinnoimat tietovarannot vuoteen 2020 mennessä niin, että avattavissa olevat tiedot ovat saatavissa ja käytettävissä koneluettavassa muodossa, maksutta ja selkein käyttöehdoin.

Tietovarantojen hyödyntämisen yhteiskunnallis-taloudellisiin kysymyksiin pureutuva tutkimus on ollut vielä varsin vähäistä, ja datan hyödyntämiseen liittyvistä taloudellisista vaikutusmekanismeista ei ole organisaatiotasolla eikä kansantalouden tasolla juurikaan tutkimustietoa. Vaikka useissa raporteissa avoimen datan taloudellisia vaikutuksia on pyritty arvioimaan ennakkolisesti, on jälkikäteinen toteutuneisiin vaikutuksiin perustuva tutkimus vasta aluillaan. Avoimen datan toteutuneista vaikutuksista ja parhaista käytänteistä tiedetään vain yksittäisiltä alueilta ja niistäkin suureksi osaksi case-esimerkkien valossa. Tietävästi missään maassa ei ole vielä arvioitu kattavasti avoimen datan toteutunutta vaikuttavuutta ja hyötyjä.

Suomi on tällä hetkellä julkisen tiedon avaamisen kärkimaiden joukossa, ja sillä on mahdollisuus myös nousta edistyksellisimpien maiden joukkoon avoimen datan vaikuttavuuden tutkimuksessa. Tämä edellyttää kuitenkin huolellista avoimen datan seuranta- ja arviointimallin kehittämistä sekä systemaattista tiedon keräämistä avoimen datan vaikuttavuuden tutkimiseksi ja parhaiden käytäntöjen selvittämiseksi. Tässä raportissa on

esitetty alustavia suuntaviivoja avoimen datan vaikuttavuuden arvioinnin kehittämiseksi erityisesti taloudellisten vaikuttavuusmittarien osalta. Raportissa esitetään katsaus siitä, mitä avoimen datan vaikuttavuudesta tällä hetkellä tiedetään, mitä menetelmiä ja mittareita voidaan tietovarantojen hyödyntämisen yhteiskunnallis-taloudellisen vaikuttavuuden tutkimisessa käyttää, ja minkälaista hyödynnettävissä olevaa aineistoa jo kerätään sekä minkälaista lisäaineistoa tarvitaan. Raportin painopiste on mitattavissa, taloudellisissa vaikutuksissa, mutta myös muita yhteiskunnallisia vaikutuksia käsitellään lyhyesti.

Julkisen sektorin tietovarantojen avaamisen vaikutukset voidaan jakaa taloudellisiin vaikutuksiin ja muihin yhteiskunnallisiin vaikutuksiin. Datan avaamisen taloudellisia vaikutuksia voidaan arvioida yritysten, kansalaisten ja kotitalouksien, julkisen sektorin sekä koko kansantalouden tasolla. Yritysten osalta merkittävimmät taloudelliset vaikutukset voivat materialisoitua uusien tuotteiden ja palveluiden kehittämisen ja toiminnan tehostumisen kautta yritysten ja tuottavuuden kasvuna. Julkisten tietovarantojen avaamisen odotetaan synnyttävän myös uutta yritystoimintaa. Kansalaisille ja kotitalouksille julkisten tietojen avaamisen merkittävimpiä taloudellisia hyötyjä arvioidaan syntyvän aiemmin maksullisen tiedon muuttumisesta maksuttomaksi sekä avoimen datan mahdollistamasta ajankäytön tehostumisesta. Julkisen sektorin organisaatioille tietovarantojen avaaminen tarjoaa mahdollisuuden palveluiden tehostamiseen ja kustannussäästöihin. Mikäli avointa dataa pystytään hyödyntämään riittävän laajalti talouden eri sektoreilla, voivat julkisten tietovarantojen avaamisen vaikutukset näkyä myös kansantalouden tasolla. Muut tietovarantojen avaamisen potentiaaliset yhteiskunnalliset vaikutukset koskevat esimerkiksi hallinnon ja päätöksenteon avoimuutta ja läpinäkyvyyttä, koulutusta, terveyttä, kansalaisten asemaa, vaikutusmahdollisuuksia ja osallistumista, ympäristövaikutuksia, kestäväää kehitystä ja liikennettä.

Julkisten tietovarantojen avaamisen toteutuneita yhteiskunnallis-taloudellisia vaikutuksia ei pystytä kattavasti arvioimaan olemassa olevan tai tällä hetkellä säännöllisesti kerättävän tiedon valossa. Arviointiin tarvitaan seurantamallin ja mittariston kehittämistä ja tämän jälkeen systemaattista tiedonkeruuta vaikutusten jälkikäteiseen mittaamiseen. Pitkäaikainen avoimen datan käytön seuranta ja tilastointi eri sektorien osalta mahdollistaisi avoimen datan taloudellisen vaikuttavuuden kvantitatiivisen analyysin kehittyneitä tilastollisia menetelmiä käyttäen. Mittariston kehittämisen ja aineiston keruun jälkeen voidaan käyttää kustannus-hyötyanalyysia arvioimaan julkisen sektorin tietovarantojen avaamisesta seuraavaa

hyvinvoinnin muutosta. Kustannus-hyötyanalyysi lienee käytännössä toteutettavissa vain keskeisten vaikutusten osalta, joten sen avulla pystytään tekemään arvio avoimen datan vaikuttavuuden ”alarajasta”. Tämän lisäksi voidaan avointen julkisten tietovarantojen hyödyntämisen vaikuttavuutta arvioida laadullisesti. Tältäkin osin keskityttäneen niihin potentiaalisiin vaikutuksiin, joita pidetään yhteiskunnallisesti tärkeinä, mutta joiden euromääräinen mittaaminen on erittäin vaikeaa.

Tilastollisen vaikuttavuusmittareita hyödyntävän aineistoanalyysin lisäksi voidaan sekä kansalaisten, yritysten että julkisen sektorin organisaatioiden osalta arvioida tarjolla olevien julkisten tietovarantojen ja erilaisten aineistojen käytettävyyttä ja hyödyllisyyttä käyttäjien omien arvioiden perusteella. Luonteva paikka käyttäjälähtöisen arvioinnin ja palautteen antamiselle on avoimen datan palvelu, jossa tietovarannot ovat käyttäjille tarjolla. Tämän lisäksi olisi syytä miettiä keinoja, miten tuotetaan tietoa avoimen datan hyödyntämiseen liittyvistä parhaista käytännöistä erityisesti julkisen sektorin organisaatioissa ja siihen, millä keinoin nämä parhaat käytännöt saadaan mahdollisimman tehokkaasti leviämään ja hyödynnetyksi eri organisaatioissa.

## **Sammanfattning**

Med öppen data avses vilka som helst data(lager) som produceras eller administreras av offentliga organisationer, och som finns i maskinläsbar form och avgiftsfritt tillgängliga för vem som helst, för förädling och distribuering i såväl privat som kommersiellt syfte. De potentiella nyttorna med öppnandet av offentliga datalager är betydande. Öppnandet av offentliga datalager har därför blivit ett centralt politiskt tema i flera länder. Den finländska regeringen fattade i mars 2011 ett principbeslut om att öppnandet av offentliga datalager för omfattande återanvändning görs till ett av regeringens prioriterade projekt. Finansministeriet tillsatte i maj 2013 det tvååriga programmet för öppen information, som syftar till att öppna offentliga datalager för allmänt bruk så snabbt och i så stor omfattning som möjligt. Målet är att samtliga betydande datalager som förvaltas av den offentliga förvaltningen enligt tydliga användningsvillkor skulle vara avgiftsfritt tillgängliga och maskinellt läsbara före 2020.

Forskning som fördjupat sig i de samhällsekonomiska aspekterna med utnyttjandet av datalager har till dags dato varit tämligen begränsat, och det finns knappast någon forskningsinformation om de ekonomiska verkningarna av utnyttjandet av data varken på

organisationsnivå eller på samhällsekonomisk nivå. Även om man i flera rapporter försökt göra preliminära uppskattningar av de ekonomiska effekterna med öppen data, håller efterhandsforskningen som baserar sig på faktiska verkningar först på att inledas. Information om faktiska konsekvenser och bästa praxis i fråga om öppen data existerar endast inom några enstaka områden, och även då till stor del som case-exempel. Inget land har veterligen tills vidare gjort några täckande utvärderingar över genomslagningen av och nyttorna med öppen data.

Finland befinner sig för tillfället i täten bland länderna som öppnat offentliga datalager, och vi har möjlighet att häva oss bland de allra mest avancerade länderna även när det gäller forskningen i genomslagningen hos öppen data. Här krävs emellertid utveckling av noggranna uppföljnings- och utvärderingsmodeller samt systematisk informationsinsamling i syfte att understöka verkningarna av öppen data och bästa praxis. I denna rapport presenteras preliminära riktlinjer för utvecklandet av verkningsutvärderingen i fråga om öppen data, speciellt med tanke på ekonomiska parametrar. Rapporten innehåller en överblick om den befintliga informationen om öppen data och dess verkning, om de metoder och mätare som kan utnyttjas vid undersökningen av de samhällsekonomiska verkningarna med utnyttjandet av offentliga datalager, hurdant material som redan insamlas och hurdant tilläggsmaterial som behövs. Rapporten fokuserar på mätbara, ekonomiska konsekvenser, men andra samhällsliga verkningar behandlas också i korthet.

Verkningarna av den offentliga sektorns datalager kan indelas i ekonomiska verkningar och andra samhällsliga verkningar. De ekonomiska verkningarna av öppnandet av data kan utvärderas på företagens, hushållens, den offentliga sektorns samt på hela samhällsekonomins plan. För företagens del kan de viktigaste ekonomiska verkningarna materialiseras i form av företags- och produktivitetstillväxt tack vare utveckling av nya produkter och tjänster och effektivare verksamhet. Öppnandet av offentliga datalager förväntas även generera ny företagsverksamhet. För medborgarnas och företagens del beräknas de viktigaste ekonomiska nyttorna bestå av att information som tidigare varit avgiftsbelagt nu blir avgiftsfritt, samt att tidsanvändningen effektiviseras tack vare öppen data. För organisationer inom den offentliga sektorn erbjuder öppnandet av datalagren möjligheter till effektivisering av service och kostnadsbesparingar. Ifall öppen data kan utnyttjas i tillräcklig omfattning inom olika ekonomiska sektorer kan effekterna av öppnandet av offentliga datalager realiseras även på samhällsekonomisk plan. Övriga potentiella samhällsliga verkningar gäller bl.a. öppenheten



och transparens hos förvaltningen och beslutsfattandet, utbildning, hälsa, medborgarnas ställning, påverkningsmöjligheter och delaktighet, miljökonsekvenser, hållbar utveckling och trafik.

De faktiska samhällsekonomiska konsekvenserna av öppnandet av offentliga datalager kan inte utvärderas på ett täckande sätt i ljuset av befintlig eller regelbundet insamlad information. För utvärderingen behövs utveckling av uppföljningsmodeller och mätare, och dessutom systematisk informationsinsamling för efterhandsmätning av effekterna. Långvarig uppföljning och statistikföring av användningen av öppen data för olika sektors del skulle möjliggöra en kvantitativ analys av genomslagningen hos öppen data med hjälp av utvecklade statistiska metoder. Efter att mätarverket utvecklats och materialet insamlats kan man utnyttja kostnads/nyttoanalyser för utvärderingen av välfärdsförändringen till följd av att den offentliga sektorns datalager öppnats. Kostnads/nyttoanalysen kan i praktiken förverkligas endast för de mest centrala effekternas del, och med dess hjälp kan man fastställa en ”nedre gräns” för öppen datas verkan. Dessutom kan verkningarna hos utnyttjandet av öppna offentliga datalager utvärderas kvalitativt. Även till denna del torde man koncentrera sig på de potentiella verkningar som betraktas som samhälleligt viktiga, men som är mycket svåra att mäta.

Förutom materialanalys som utnyttjar statistiska effektivitetsmätare kan man både för medborgares, företagens och offentliga organisationers del utvärdera tillgängliga datalagrars och olika materials användbarhet och nytthet utifrån användarnas egna bedömningar. Ett naturligt forum för användarvänlig utvärdering och respons utgörs av tjänsten för öppen data, där datalagren görs tillgängliga för användarna. Dessutom bör man begrunda medel genom vilka man kunde producera information om bästa praxis för utnyttjandet av offentlig data speciellt inom offentliga organisationer, och om hur dessa bästa praxis kan spridas till och utnyttjas inom olika organisationer på ett så effektivt sätt som möjligt.

## SUMMARY

Reported ex ante evaluations have estimated that the potential benefits of opening up public data resources are substantial. Consequently, various countries have implemented the open (government) data strategy aiming at providing wide access to government data in machine-readable format such that it can be freely used, reused and redistributed by anyone. In March

2011, the Finnish government made a resolution to open up government data. In May 2013, the Ministry of Finance launched a two-year Open Data Program to accelerate government data openings. In Finland, the target is to open up all major public databases by the year 2020.

Very little is known about the underlying economic and organizational mechanisms and implications of open data use at the organizational level or at the level of economy as a whole. Various reports have evaluated the economic impacts of open data ex ante but the literature lacks ex post analysis focusing on the materialized economic impacts. To my best knowledge, there is no reported comprehensive country-level ex-post impact assessment of opening up government data. The current research-based knowledge concerning the impacts of open data is limited only to narrow areas and largely based on case examples.

Currently, Finland is among the leading countries in opening up the government data, and it has also a chance to be among the most advanced ones in the impact assessment of open data. The prerequisite for this is a careful development of the monitoring and evaluation model for opening up government data as well as a systematic gathering of data for the impact assessment. This report provides guidelines for the ex-post evaluation of the impacts of open data. It summarizes contemporary knowledge concerning the impacts of open data, what methods and measures can be utilized in research assessing the socio-economic impacts of open data, what kind of data are currently available for such research and what kind of data are further needed. This report focuses on measurable, economic impacts but also other social impacts are briefly discussed.

The impacts of opening up government data can be divided to economic impacts and to other social impacts. Economic impacts can be assessed at the level of firms, citizens and households, public sector and the economy as a whole. The most notable potential economic impacts for firms are growth and increased productivity via the efficiency improvements and via the development of new products and services enabled by open data. Opening up government data is also expected to result in the establishment of new companies. For citizens, the most important economic benefits of open data are likely to arise from free access to data instead of using chargeable data resources and from the time savings. For public sector organizations, opening up data resources offers cost savings and an opportunity to improve efficiency of service provision. The economic impacts of opening up government

data may also be observed at the level of economy as a whole if open data can be exploited sufficiently broadly across different sectors of the economy. Other potential social impacts of opening up government data concern, for instance, the transparency of government and decision making, education, health, citizens engagement, environmental impacts, sustainability and transportation.

Contemporarily gathered and available (statistical) data is not sufficient for the comprehensive assessment of the materialized socio-economic impacts of opening up government data. For such assessment, we first need to develop the monitoring and evaluation model and then systematically gather data over the years for the ex post assessment of the impacts of open data. This is a prerequisite for the quantitative impact assessment of open government data. After that, (partial) cost-benefit analysis can be applied for evaluating the ex post changes in the welfare followed by the opening up the government data. Furthermore, the impacts of open data can be qualitatively assessed focusing on those important (expected) social impacts of which measurement in euros is difficult.

Furthermore, the usability and usefulness of different public data resources for citizens, firms and public sector organizations can be assessed via the users' own evaluations. A natural platform for such evaluations and feedback is the open data service providing access to the government data. In addition, it is important to assess appropriate means to disseminate and promote efficient utilization of information on the best practices of open data re-use in different organizations.

## 1. Johdanto

Tietovarantojen eli datan hyödyntämisen merkitys kasvaa jatkuvasti sekä yksityisellä että julkisella sektorilla. Euroopan tasolla komissio pyrkii tukemaan julkisen datan avaamista joulukuussa 2011 julkaisemassaan avoimen datan strategiassa. Sen tavoitteena on edistää tietovarantojen hyödyntämistä jäsenmaissa myös muun muassa tutkimusohjelmien ja infrastruktuurihankkeiden avulla. Suomessa julkisten tietovarantojen avaaminen on nimetty yhdeksi hallituksen kärkihankkeista. Toukokuussa 2013 valtiovarainministeriö asetti kaksivuotisen julkisten tietovarantojen avaamiseen mahdollisimman laajasti ja nopeasti tähtäävän avoimen tiedon ohjelman. Tavoitteena on avata kaikki merkittävät julkishallinnon hallinnoimat tietovarannot vuoteen 2020 mennessä niin, että avattavissa olevat tiedot ovat saatavissa ja käytettävissä koneluettavassa muodossa, maksutta ja selkein käyttöehdoin.

Tietovarantojen hyödyntämisen yhteiskunnallis-taloudellisiin kysymyksiin pureutuva tutkimus on kuitenkin ollut vielä varsin vähäistä, ja datan hyödyntämiseen liittyvistä taloudellisista vaikutusmekanismeista ei ole organisaatiotasolla eikä kansantalouden tasolla juurikaan tutkimustietoa. Useissa raporteissa on avoimen datan taloudellisia vaikutuksia pyritty arvioimaan ennakkollisesti (kts. esim. Vicker, 2011), mutta jälkikäteen toteutuneisiin vaikutuksiin perustuvaa tutkimus on vasta aluillaan. Syynä on osittain tutkimusaiheen haasteellisuus (esim. mittaamisvaikeudet; tietovarantojen käytöstä ei ole systemaattisesti tilastoitua tietoa) ja se, että kysymys on verrattain uudesta ilmiöstä. Erityisesti tutkimustietoa tarvitaan siitä, millaisin edellytyksin ja toimenpitein tietovarantojen avaamisesta saadaan suurimmat yhteiskunnalliset ja taloudelliset hyödyt.

Aihepiirin tutkimiseksi tarvitaan systemaattista ja pitkäjänteistä työtä. Datan avaamisen vaikuttavuuden tutkimisen edellytyksenä on tietovarantojen avaamisen ja avoimen datan vaikutusten arviointimallin kehittäminen. Tätä varten tarvitaan ensinnäkin katsaus siihen, mitä avoimen datan vaikuttavuudesta jo nyt tiedetään, mitä menetelmiä ja mittareita voidaan tietovarantojen hyödyntämisen yhteiskunnallis-taloudellisen vaikuttavuuden tutkimisessa käyttää, ja minkälaista hyödynnettävissä olevaa aineistoa jo kerätään sekä minkälaista lisäaineistoa tarvitaan. Raportoidun esitutkimuksen tavoitteena on tämän tiedon tuottaminen. Raportin painopiste on mitattavissa, taloudellisissa vaikutuksissa, mutta myös muita yhteiskunnallisia vaikutuksia käsitellään lyhyesti.

Raportin luku 2 kuvailee keskeisiä julkisten tietovarantojen avaamiseen tähtääviä toimenpiteitä Suomessa sekä valottaa Suomen sijoittumista kansainvälisessä rankingissa julkisten datojen avaamisen suhteen. Luku 3 esittää, mitä taloudellisia ja muita yhteiskunnallisia vaikutuksia avoimella julkisella tiedolla on arvioitu olevan. Luku 4 luo alustavaa kehikkoa avoimen datan vaikuttavuuden seurantomallin kehittämiseksi. Luku 5 esittää yhteenvedon.

## 2. Tietovarantojen avaaminen – nykytilanne Suomessa ja kansainvälinen vertailu

### Julkinen data avoimeksi vuoteen 2020 mennessä

Avoin julkinen data on mikä tahansa julkisen organisaation tuottama tai hallinnoima tieto(varanto), joka on konekielisessä muodossa ja maksutta kenen tahansa käytettävissä, muokattavissa ja jaettavissa sekä yksityisiin että kaupallisiin tarkoituksiin. Myös yksityiset tahot voivat tarjota avointa dataa, mutta tässä raportissa keskitytään julkisen sektorin tietovarantoihin ja niiden avaamiseen. Tärkeitä julkisen sektorin, myös yrityksissä laajalti hyödynnettyjä tietoja ovat esimerkiksi paikkatiedot, säätiedot, väestötiedot, yritystiedot, liikennetiedot sekä lainsäädännölliset tiedot.

Suomen hallitus teki maaliskuussa 2011 periaatepäätöksen siitä, että julkisten tietojen avaaminen laajalti yhteiskunnan eri tahojen uudelleenkäyttöön on yksi hallitusohjelman kärkihankkeista. Päätöksen jälkeen siirryttiin vuonna 2012 valtionhallinnon tietovarantojen systemaattiseen avaamiseen.<sup>1</sup> Esimerkiksi Maanmittauslaitoksen kaikki digitaaliset maastotietoaineistot avattiin toukokuussa 2012 kansalaisten ja yritysten käyttöön. Vuotta myöhemmin Ilmatieteen laitos avasi säätiedot omassa verkkopalvelussaan. Tavoitteena on avata kaikki merkittävät julkishallinnon keräämät ja hallinnoimat tietovarannot vuoteen 2020 mennessä niin, että tiedot ovat saatavissa ja käytettävissä koneluettavassa muodossa, maksutta ja selkein käyttöehdoin<sup>2</sup>. Kaikkien tietovarantojen avaaminen ei ole mahdollista lainsäädännön puitteissa; tietovarantoja avataan huomioiden lainsäädännölliset rajoitukset

---

<sup>1</sup> Jo ennen tätä periaatepäätöstä tehtiin merkittäviä julkisen datan avauksia. Esimerkiksi Suomen ympäristökeskus, SYKE, tarjosi jo ennen vuotta 2011 rekisteröityneille käyttäjille ympäristö- ja paikkatietoja OIVA –palvelussa (<https://www.wp2.ymparisto.fi/scripts/oiva.asp>).

<sup>2</sup> <http://valtioneuvosto.fi/etusivu/rakenneuudistus395285/tiedostot/ministerioiden-materiaalit-15112013/vm/VM-aineistot-6-15112013.pdf>

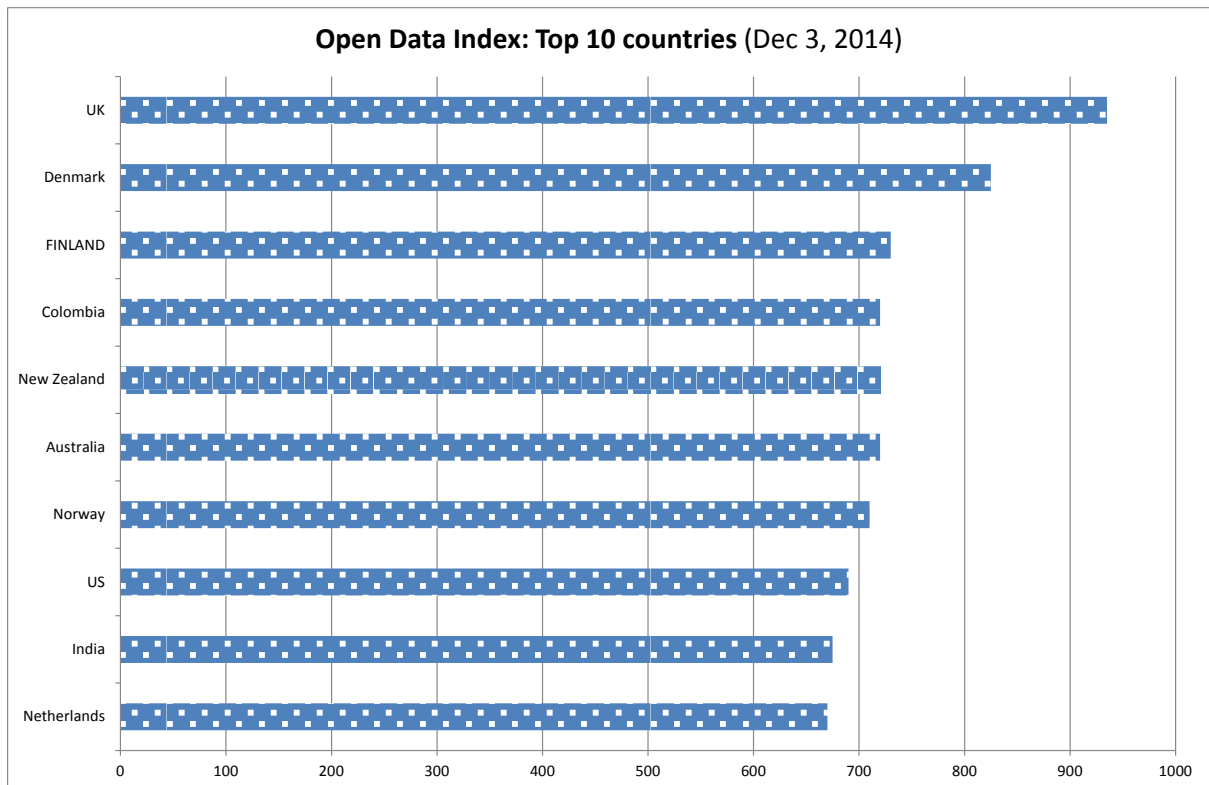
kuten yksityisyydensuoja, liikesalaisuudet, tekijänoikeudet ja kansalliseen turvallisuuteen liittyvät seikat. Tavoitteena on, että tiedot kaikista keskeisistä tietoaaineistoista tulevat saataville syyskuussa 2014 avattuun julkisen hallinnon yhteiseen avoimen datan palveluun [avoindata.fi](http://avoindata.fi):hin.

Julkishallinnon avoimuuden ja läpinäkyvyyden lisäämiseksi tehdään mittavaa kansainvälistä yhteistyötä, jossa tietovarantojen avaamisella on keskeinen osa. Vuonna 2011 perustetun kansainvälisen Avoimen hallinnon kumppanuusohjelman (Open Government Partnership, OGP) tavoitteena on lisätä julkisen hallinnon avoimuutta, vastuullisuutta ja kansalaisten huomioonottamista. Suomi hyväksyttiin jäseneksi OGP -kumppanuusohjelmaan huhtikuussa 2013. Vuoden 2014 marraskuussa OGP:n yhteistyössä oli mukana 65 avoimempaan hallintoon pyrkivää jäsenvaltiota (<http://www.opengovpartnership.org/>). Jäsenmaiden julkaisemissa avoimen hallinnon toimintasuunnitelmissa on julkisten tietojen avaamiseen liittyvillä sitoumuksilla ollut keskeinen rooli (Davies ym., 2013).

## Suomi lähestyy avoimen datan eturintamaa

Julkisen datan avaamisen etenemisestä eri maissa on tehty kansainvälistä vertailua. Open Data Index on Open Knowledge Foundationin käyttämä mittari, jolla kuvataan julkisen sektorin tietojen avaamisen edistymistä ja laajuutta eri maissa. Julkisten tietojen avaamista seurataan kymmenellä osa-alueella: liikenteen aikataulutiedot, valtion budjetti, valtion menot, äänestystulokset, yritysrekisterit, kansallinen karttatietokanta, maan keskeiset taloudelliset ja demografiset tilastotiedot, lainsäädäntö, postinumerot ja ympäristöpäästöt. Open Data Indexin päivittämisen vastuutahot vaihtelevat maittain ja joissakin maissa tietovarantojen avaamisen päivitykset indeksiä varten ovat vapaaehtoisten varassa. Täten Open Data Indexin kansainvälisiä vertailutietoja julkisten tietovarantojen avaamisesta voidaan pitää suuntaa antavina.

Open Data Index -mittaria käyttäen Iso-Britannia, Tanska ja Suomi muodostivat julkisten tietojen avaamisen kärkikolmikon vuoden 2014 joulukuussa. Suomen sijoitus vertailussa oli kolme pykälää parempi kuin edellisenä vuonna, kun taas Yhdysvallat tippui toiselta sijalta kahdeksannelle. Suomessa suurimmat puutteet julkisten tietovarantojen avaamisessa koskivat yritysrekisteritietoja. Myöskään lainsäädäntötiedot eivät olleet täysin avoimesti saatavissa.



Aineistolähde: Open Data Index 2014, Open Knowledge Foundation. <https://index.okfn.org/country/>

Julkisten tietovarantojen avaamisen eturintamassa pitkään olleissa maissa kuten Iso-Britanniassa ja Yhdysvalloissa julkisen tiedon avaamiseen tähtäävät merkittävät politiikkatoimenpiteet aloitettiin jonkin verran aiemmin kuin Suomessa. Iso-Britannian joulukuussa 2009 julkaiseman avointa dataa koskevan strategian ”Putting the Frontline First: Smarter Government”<sup>3</sup> keskeisinä tavoitteina oli edistää taloudellista kasvua, innovaatiotoimintaa sekä vahvistaa kansalaisten ja kansalaisyhteiskunnan roolia tietovarantojen avaamisen ja läpinäkyvyyden lisäämisen kautta. Yhdysvalloissa Barack Obaman vuoden 2009 tammikuussa julkistama muistio tähdensi julkisten tietojen avaamisen merkitystä ja johti avoin hallinto (”Open Government”) –ohjelman julkaisemiseen saman vuoden joulukuussa. Ohjelman yhtenä keskeisenä pyrkimyksenä oli tehdä poliittinen päätöksenteko läpinäkyvämmäksi avaamalla internetiin tietokantoja, joita kansalaiset voivat käyttää hallituksen toimenpiteiden seuraamiseen ja arviointiin. Lainvalvonnan vahvistaminen dataperusteisten sovellusten avulla on ollut yksi tietovarantojen avaamisen peruste esimerkiksi Isossa-Britanniassa ja Yhdysvalloissa. Tarkoituksena on ollut sekä informoida

<sup>3</sup> Chief Secretary to the Treasury (2009). Putting the Fronline First: Smarter Government. London: HMGovernment.

kansalaisia että lisätä kansalaisten osallistumista lainvalvontaan (esimerkiksi rikollisista tai rikoksista tiedottavien sovellusten avulla).

Julkisten tietovarantojen avaaminen on ollut monissa maissa osana avoimempaan ja läpinäkyvämpään hallintoon pyrkivää strategiaa. Datan avaamisen taloudelliset hyödyt kuten uusien tuotteiden, palveluiden ja yritysten syntyminen, kustannussäästöt ja tehokkuuden lisääntyminen sekä taloudellinen kasvu on esitetty tärkeinä julkisten tietojen avaamista puoltavina seikkoina erityisesti monien Euroopan maiden (esim. Tanskan<sup>4</sup>) kansallisissa strategioissa. Julkisen tiedon avaamista perustellaan monissa maissa pyrkimyksillä saavuttaa erilaisia yhteiskunnallis-taloudellisia hyötyjä.

### 3. Avoimen datan vaikuttavuus

#### Datan avaaminen lisää sen käyttöä

Julkisten tietovarantojen avaamisen ja laajemmin tietovarantojen hyödyntämisen toteutuneesta yhteiskunnallis-taloudellisesta vaikuttavuudesta on vielä varsin vähän tutkimukseen perustuvaa tietoa. Avoin data barometri –projektin raporttoiman selvityksen mukaan vuonna 2013 missään tutkimuksen kohteena olleesta 77 maasta ei oltu arvioitu kattavasti datan avaamisen vaikuttavuutta tai hyötyjä.<sup>5</sup> Merkittävä syy tähän on tietovarantojen hyödyntämiseen liittyvän systemaattisen tiedonkeruun ja tilastoinnin puuttuminen. Aineistoanalyysiin perustuvan tutkimustiedon ollessa vähäistä voidaan avoimen datan vaikuttavuutta eri osa-alueilla tässä vaiheessa valottaa pääasiassa potentiaalisten vaikutusten teoreettisen pohdinnan sekä case-esimerkkien avulla.

Lukuisat julkisen datan avaamista koskevat esimerkit viittaavat siihen, että datan avaaminen lisää merkittävästi tietoaineistojen käyttöä. Esimerkiksi Espanjassa perustettiin maaliskuussa 2003 OVC (“Virtual Office of Cadastre”), jonka kautta maantieteellistä tietoa oli laajalti saatavissa ilmaiseksi. Kesäkuussa 2004 datan avaamista jatkettiin ja maantieteellisten tietojen saatavuus parani geoportaali “IDEE”:n avaamisen seurauksena. Tämä geoportaali toi huomattavan määrän maantieteellisiä tietoja ilmaiseksi kenen tahansa käytettäväksi. Tästä

---

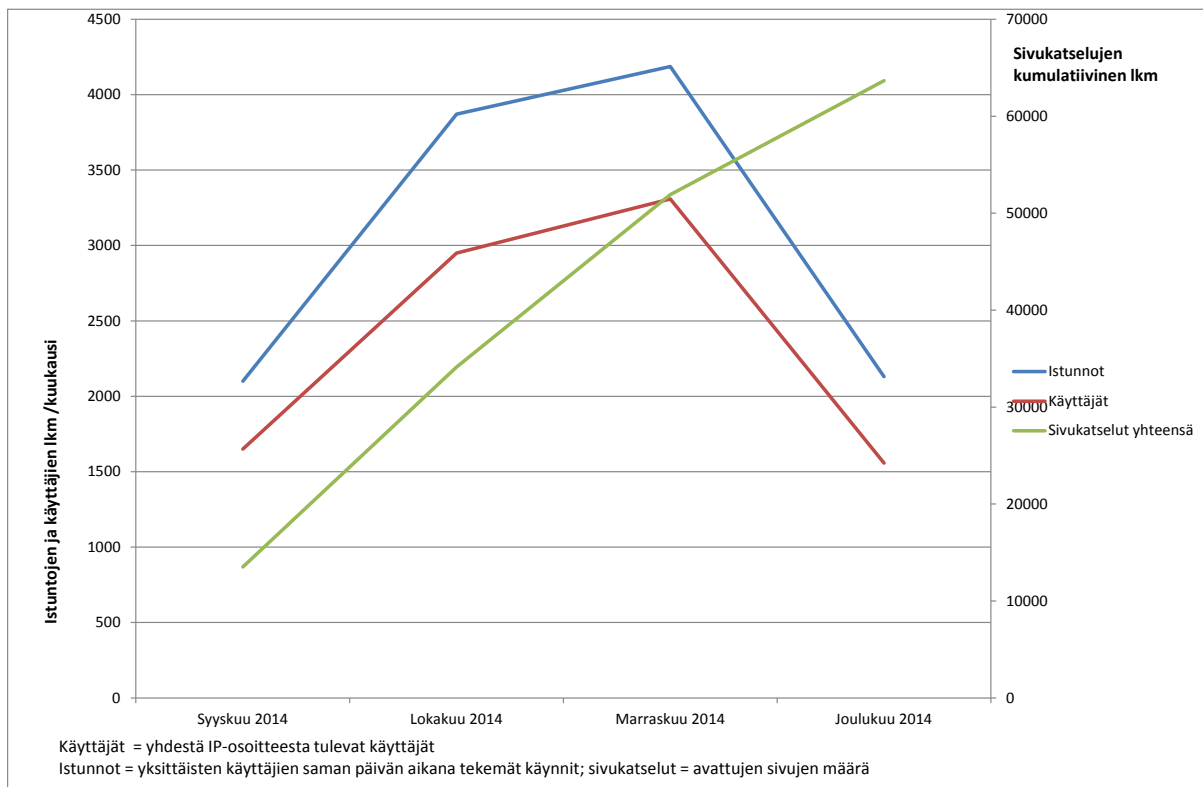
<sup>4</sup> The Danish government/local government Denmark, October 2012: “The eGovernment strategy 2011-2015. Good basic data for everyone – a driver for growth and efficiency.” ([http://uk.fm.dk/publications/2012/good-basic-data-for-everyone/~/\\_media/Publikationer/Imported/2012/Gode%20grunddata%20til%20alle/BasicData\\_UK\\_web\\_2012.10.08.ashx](http://uk.fm.dk/publications/2012/good-basic-data-for-everyone/~/_media/Publikationer/Imported/2012/Gode%20grunddata%20til%20alle/BasicData_UK_web_2012.10.08.ashx).)

<sup>5</sup> Open Data Barometer. 2013 Global Report. Source:<http://www.opendataresearch.org>.



seurasi merkittävä kasvu maantieteellisen datan kysynnässä. Esimerkiksi karttojen käytön määrä kasvoi vuodesta 2004 vuoteen 2005 noin 700 prosenttia (yli miljoonasta käyttökerrasta yli 41 miljoonaan käyttökertaan). Vuonna 2010 vastaava luku oli yli 124 miljoonaa, eli kysyntä oli yli 100 kertaa suurempi vuoteen 2004 verrattuna. Vastaavia esimerkkejä on monia. Suomessa maastotietojen avaaminen toukokuussa 2012 on sekä lisännyt että monipuolistanut maastotietojen käyttöä. Ahonen-Rainion ym. (2014) raportin mukaan uusia käyttäjiä on tullut erityisesti yksityisten henkilöiden ja pk-yritysten joukosta. Julkisen hallinnon yhteisen datapalvelun avoindata.fi:n käyttö on myös kasvanut nopeasti sen aukaisemisen jälkeen. Palvelun aukaisemisen jälkeen sivukatseluja kertyi ensimmäisen kolmen ja puolen kuukauden aikana yhteensä yli 63 000 kappaletta.

#### *Avoindata.fi –palvelun käytön määrä 15.9.2014-31.12.2014*



Kuva: Kuvassa esitetään avoindata.fi-palvelun käyttömäärät joulukuuhun 2014 asti. Vuoden 2015 käyttömäärät ovat jatkuvasti kasvaneet.

Tietovarantojen käytön huomattava lisääntyminen julkisten aineistojen avaamisen seurauksena ei vielä itsessään kerro avoimen datan hyödyntämisen yhteiskunnallistaloudellisesta vaikuttavuudesta. Julkisen sektorin tietovarantojen avaamisen vaikutukset voidaan jakaa taloudellisiin vaikutuksiin ja muihin yhteiskunnallisiin vaikutuksiin. Datan

avaamisen taloudellisia vaikutuksia voidaan arvioida yritysten, kansalaisten (ja kotitalouksien), julkisen sektorin sekä koko kansantalouden tasolla:

- i) *yritys- ja toimialavaikutuksia* arvioitaessa tarkastelukohteena ovat esim. kilpailukyky, innovaatiotoiminta, tuottavuus sekä kilpailu ja markkinadynamiikka (yritysten markkinoille tulo ja sieltä poistuminen, yrittäjyys)
- ii) *vaikutukset kansalaisiin*, kuten vaikutukset kulutukseen, tulonjakoon ja erilaisten väestöryhmien (esim. palkansaajat, erilaisten etuuksien saajat) asemaan.
- iii) *julkistaloudelliset vaikutukset*, kuten vaikutukset valtion ja kuntien talouteen ja viranomaisten resurssien käyttöön, julkisen sektorin työllisyys- ja tuottavuusvaikutukset, vaikutukset julkisiin palveluihin
- iv) *kansantaloudelliset vaikutukset*, kuten vaikutukset tuottavuuteen ja kasvuun kansantalouden tasolla sekä työmarkkinavaikutukset.

Muut tietovarantojen avaamisen potentiaaliset yhteiskunnalliset vaikutukset koskevat esimerkiksi hallinnon ja päätöksenteon avoimuutta ja läpinäkyvyyttä, koulutusta, terveyttä, kansalaisten asemaa, vaikutusmahdollisuuksia ja osallistumista, ympäristövaikutuksia, kestäväää kehitystä ja liikennettä.

### **Avoimesta datasta innovaatioita, kasvua ja tuottavuushyötyjä yrityksille?**

Eri toimialojen yritysten potentiaali hyödyntää avointa dataa liittyy likeisesti yritysten toimintojen tietointensiivisyyteen. Toimialojen tietointensiivisyyttä on mitattu esimerkiksi datan määrällä organisaatiota kohden ja organisaation työllistämällä data-analyytikoilla suhteutettuna työntekijöiden kokonaismäärään. Toimialoista tietointensiivisimpien joukossa organisaatiokohtaisen datan määrällä ovat McKinsey Global Instituten (2011) raportin mukaan rahoituspalvelut, informaatio ja viestintä sekä sähkö-, kaasu- ja lämpöhuolto. OECD:n (2013) raportin mukaan suurin potentiaalinen hyöty tietovarannoista innovaatiotoiminnassa ja tuottavuuden kasvun lähteenä yksityisellä sektorilla on terveydenhuoltosektorilla, mainostoiminnassa, infrastruktuuripalvelujen (esim. kaasu, sähkö, vesi) tuotannossa ja kuljetus- ja logistiikka-alalla.

Tietovarantojen avaamisen vaikutuksissa yritysten toimintaan voidaan täten olettaa olevan suurta toimialoittaista vaihtelua. Tällä hetkellä suhteellisen vähän tietovarantoihin

toimintaansa perustavat yritykset voivat kuitenkin ryhtyä hyödyntämään erilaista dataa laajemmin, kun tietovarantoja avataan ja tietoperusteisten toimintamallien hyödyistä saadaan enemmän tietoa. Yritykset voivat hyödyntää tietovarantoja paitsi uusien tuotteiden ja palveluiden raaka-aineena myös laajemmin päätöksenteon apuna esimerkiksi maatalouden ja rakentamisen tuotannonohjauksessa ja yrityksen toimintojen sijaintia koskevissa päätöksissä. Pitkällä aikavälillä tietovarantojen avaamisesta saatavat hyödyt yrityksissä riippuvat keskeisesti tietoperusteisten toimintojen (esim. tietoperusteisten päätöksentekojärjestelmien hyödyntäminen) käyttöönoton nopeudesta ja laajuudesta eri toimialojen yritysten joukossa. Brynjolfssonin ym. (2011) aineistoanalyysiin perustuvan tutkimuksen mukaan tietoperusteisia päätöksentekomalleja soveltavien yritysten tuottavuus ja kannattavuus oli noin 5-6 prosenttia suurempi kuin niiden kilpailijoiden.

Avoimen datan vaikuttavuudesta yrityksissä on olemassa tutkimustietoa lähinnä maantieteellisten tietojen osalta. Ahonen-Rainio ym. (2014) arvioivat avoimen maastotiedon vaikuttavuutta. Raportissa kartoitetaan maastotiedon käyttötapoja, merkitystä ja hyötyjä käyttäjille, mutta taloudellisia vaikutuksia ei ole pyritty arvioimaan euromääräisesti. Kyselytutkimuksen mukaan yritykset hyötyvät taloudellisesti avoimesta maastotiedosta, koska avoin data mahdollistaa uusien palveluiden ja tuotteiden kehittämisen, avoimen datan käyttö tehostaa toimintaa ja tarjoaa yrityksille kilpailuetua sekä kotimaassa että ulkomailla.

Kosken (2011) empiirisen tutkimuksen mukaan ”tekninen palvelu” -toimialan yritysten joukossa 15 eri maassa paikkatietojen avaamisen ja rajakustannushintoihin siirtyminen oli yhteydessä merkittävästi nopeampaan kasvuun yrityksissä. Paikkatietoja rajakustannushintaan tai ilmaiseksi tarjonneissa maissa paikkatietoja hyödyntävän ”tekninen palvelu” -toimialan yritysten liikevaihto kasvoi vuositasolla keskimäärin 15 prosenttia nopeammin kuin saman toimialan yritysten liikevaihto muissa maissa. Suurten yritysten kasvuun paikkatietojen avaamisella ei kuitenkaan ollut merkittävää vaikutusta.

Avoimen datan hyödyntämisestä yritysten innovaatiotoiminnassa ei myöskään ole paljon tutkimukseen perustuvaa tietoa. Kosken (2012) tutkimuksessa käytettiin 531 suomalaisesta yrityksestä koottua kyselyaineistoa analysoimaan uusien tietoperusteisten tuotteiden ja palveluiden tuottamista suomalaisissa yrityksissä. Tutkimuksen perusteella tietoperusteisia tuote- ja palveluinnovaatioita tuotettiin muita toimialoja todennäköisemmin ammatillista, tieteellistä ja teknistä toimintaa harjoittavissa yrityksissä sekä kiinteistöalalla ja

rakentamistoimintaa harjoittavissa yrityksissä. Kiurun ym. (2012) haastattelututkimuksessa ilmeni, että erityisesti pienet yritykset pystyvät hyödyntämään avattavia julkisia dataja nopeasti uusissa tuotteissaan ja palveluissaan.

### Esimerkkejä avoimen datan hyödyntämisestä yrityksissä



Tutkimukset viittaavat siihen, että julkisten tietojen avaaminen voi toimia innovaatiotoiminnan ja kasvun mahdollistajana erityisesti pk-yrityksissä toimialoilla, joissa data on keskeisessä roolissa yritysten liiketoiminnassa. Kirjallisuudessa on myös viitteitä sitä, että tiettyjen tietoperusteisten markkinoiden koko ja kasvu eri maissa ovat likeisessä yhteydessä kyseisillä toimialoilla hyödynnettävien tietojen avoimuuteen (kts. esim. Pettifer, 2009). Yhdysvalloissa meteorologinen data on ollut - toisin kuin Euroopassa - avointa jo pitkään. Meteorologisten tietotuotteiden ja -palveluiden markkinat ovat kasvaneet Yhdysvalloissa huomattavasti nopeammin kuin Euroopassa: Vuonna 2006 Yhdysvaltojen meteorologisten tuotteiden ja palveluiden markkinoiden koko oli 1.4 USD, kun taas Euroopassa vastaavien markkinoiden koko oli vain 372 miljoonaa USD. Myös markkinoiden kasvuvauhti oli varsin erilainen: meteorologisten markkinoiden keskimääräinen vuositaso kasvuvauhti Yhdysvalloissa ajanjaksolla 2000 - 2006 oli noin 17 prosenttia, kun taas Euroopassa markkinat kasvoivat samaan aikaan vain 1.2 prosentin vuosivauhtia.

Avoimen datan hyödyntäminen liiketoiminnassa tarjoaa mahdollisuuden uusien liiketoimintamallien käyttöönottoon, uusien yritysten perustamiseen ja jopa avoimen julkisen

tiedon hyödyntämiseen perustuvien uusien toimialojen syntyyn. Esimerkiksi Espanjassa vuonna 2012 yli 150 yritystä oli keskittynyt toiminnassaan pelkästään avoimeen dataan liittyvien palveluiden myyntiin perustuvalle ”infomediary” –toimialalle. Tällöin toimiala työllisti noin 4000 henkilöä ja avoimeen dataan liittyvien palveluiden myynnin arvoksi arvioitiin 330-500 miljoona euroa vuosittain.

Espanjan ”infomediary”-toimiala tarjoaa esimerkin julkisen avoimen datan luomista uusista liiketoimintamahdollisuuksista, mutta tietoa yleisemmin siitä, missä määrin julkisen raakadatan uudelleenkäyttö on johtanut uusien yritysten perustamiseen eri toimialoilla ei ole. Kirjallisuudessa on raportoitu pääasiassa yksittäisiä menestystarinoita. Esimerkiksi amerikkalainen televisiokanava The Weather Channel ja GPS-markkinoiden johtava yritys Garmin perustettiin käyttäen julkista raakadataa Garminin markkina-arvo vuonna 2013 oli yli 7 miljardia dollaria (Capgemini Consulting, 2013; The Open Data Economy). Esimerkin julkisen raakadatan hyödyntämisen synnyttämästä yritystoiminnasta tarjoaa myös vuonna 2006 perustettu Climate corp., joka keräsi aluksi säättietoja ja satoennustedataa maanviljelijöille tarjottavia säävakuutuksia varten. Myöhemmin yritys alkoi hyödyntää tietoja laajemmin maatalouksien kylvösuunnittelun avuksi. Vuoden 2013 syksyllä Climate corp. myytiin 930 miljoonan USD hintaan.

### **Avoim data kansalaisten käytössä – säästöjä, parempia päätöksiä ja osallistavaa demokratiaa?**

Julkiset sektorin keräämät ja hallinnoimat tietovarannot ovat verovaroin rahoitettuja julkishyödykkeitä niiltä osin kuin niiden julkisuudelle ei ole lainsäädännöllisiä rajoitteita. Kansalaisilla on täten oikeus avoimeen julkiseen dataan. Yksi avoimen datan merkittävimmistä taloudellisista suorista hyödyistä kansalaisten näkökulmasta liittyy sen mahdollistamaan ajankäytön tehostumiseen. Esimerkiksi reaaliaikaista tietoa käyttäjän arvioiduista matka-ajoista eri liikkumismuotoja ja liikennevälineitä käyttäen tarjoavat helppokäyttöiset sovellukset voivat tuottaa kuluttajille taloudellista hyötyä (esim. ajankäytön säästöt, jotka saadaan ruuhkien välttämisellä).<sup>6</sup> Siltä osin kuin kansalaisille relevanttien

---

<sup>6</sup> Esimerkki tällaisesta sovelluksesta on Helsingin seudun liikenteen tarjoama paikallinen reittiopas: [www.reittiopas.fi](http://www.reittiopas.fi).

julkisten aineistojen avaaminen tarkoittaa tarpeellisen maksuttoman tiedon saataville tuloa aiemmin maksullisen sijaan, kansalaiset hyötyvät aineistoavauksista myös rahamääräisesti.

Julkisen tiedon avaaminen tarjoaa kansalaisille mahdollisuuden parempien henkilökohtaista elämää koskevien päätösten tekoon. Erilaiset paikallista ympäristöä koskevien tietovarantojen avaamiset ja näitä tietoja hyödyntävien sovellusten saataville tulo tarjoavat apuvälineitä päätöksenteon tueksi ja voivat parantaa elämänlaatua. Esimerkiksi tarkat, paikalliset sääennusteet voivat ohjata yksittäisten ihmisten liikkumista koskevia päätöksiä (esim. paikasta toiseen siirtymisen ajankohtaa ja liikennemuodon valintaa) ja parantaa myös liikenneturvallisuutta. Monissa maissa, kuten esimerkiksi Yhdysvalloissa, on kehitetty avoimeen dataan perustuvia kansalaisten turvallisuuden lisäämiseen ja rikosten ehkäisemiseen tähtäviä erilaisten rikosten tapahtumista alueellisesti (esim. kaupungeissa katutason tarkkuudella) kartoitettavia sovelluksia. Näillä voi olla myös asuinpaikan valintaa ohjaava vaikutus.

Poliittisia päätöksiä koskevien tietojen avoimuus tekee päätöksenteosta läpinäkyvämpää, mahdollistaa sen paremman valvonnan ja voi edistää osallistavan demokratian toteutumista. Kansalaisten aktiivinen hallinnon ja päätöksenteon seuranta antaa kansalaisille edellytykset omasta näkökulmastaan parempien päätöksentekijöiden valintaan ja ainakin periaatteessa voi lisätä tätä kautta hyvinvointia. Avoin kysymys on, missä määrin kansalaiset hyödyntävät mahdollisuutta seurata päätöksentekoprosesseja jo avattujen ja tulevaisuudessa avattavien datojen kautta.

### **Datan avaamisella tehokkaampaan julkisten palveluiden tarjontaan?**

Julkinen sektori on yksi talouden tietointensiivisimmistä sektoreista. Se on merkittävä tietovarantojen tuottaja, hallinnoija ja hyödyntäjä. OECD:n (2013) raportissa arvioidaan, että julkinen sektori on yksi talouden sektoreista, joka voi saada merkittävimpiä taloudellisia hyötyjä tietovarantojen käytöstä. Tietovarantojen avaaminen voi hyödyttää julkista sektoria avointa dataa käyttävän yritystoiminnan kasvun kautta syntyvien verotulojen muodossa. Pettifer (2009) arvioi, että meteorologisten aineistojen avoimuus olisi merkinnyt vuositason Euroopassa noin 300 miljoonaa euroa suurempia verotuloja meteorologiselta sektorilta. Tietovarantojen avaaminen julkisen hallinnon sisällä yli maarajojen voi myös lisätä julkisen

sektorin verotuloja, koska verotustietojen kansainvälisellä avoimuudella voidaan vähentää kansainvälistä veronkiertoa.

Julkisen sektorin organisaatioille tietovarantojen avaaminen tarjoaa mahdollisuuden palveluiden tehostamiseen ja kustannussäästöihin. Suljettuja tietovarantoja hallinnoivissa organisaatioissa joudutaan vastaamaan huomattavasti suurempaan määrään asiakkaiden tietovarantoja koskevia kyselyitä kuin tiedon ollessa avointa. Toisin sanoen, datan avaaminen voi pienentää tietovarantoja avanneiden julkisen sektorin organisaatioiden transaktiokustannuksia. Esimerkiksi San Franciscossa julkista liikennettä koskevan datan avaaminen vähensi puhelintiedusteluja noin 22 prosenttia ja tuotti yli miljoonan dollarin vuotuiset säästöt<sup>7</sup>.

Julkisen sektorin organisaatioiden toimintaa koskevan läpinäkyvyyden lisääntyminen voi tarjota keinoja päätöksenteon ja toimintojen tehostamiseen. Avointa dataa voidaan hyödyntää julkishallinnon suunnittelussa ja toiminnan ohjaamisessa entistä kustannustehokkaampaan suuntaan. Sen avulla voidaan analysoida mahdollisia kustannussäästökohteita ja arvioida erilaisista toimenpiteistä tai erilaisiin toimenpiteisiin kannustamisesta saatavien kustannussäästöjen suuruusluokkaa. Esimerkiksi Englannin julkinen terveydenhuoltojärjestelmä NHS (National Health Service) olisi vuonna 2011 säästänyt yli 200 miljoonaa puntaa, jos jokainen lääkäri olisi sydän- ja verisuonisairauksiin liittyvän lääkemääräyksen tehdessään määrännyt patentoidun alkuperäislääkkeen sijasta halvemman, koostumukseltaan samanlaisen geneerisen lääkkeen<sup>8</sup>.

Avoin data voi toimia julkisen sektorin sisäisenä kannustimena julkisten palveluiden tehostamiseksi ja parantamiseksi. Iso-Britanniassa kaikkien sairaaloiden infektioilastojen julkistaminen sekä pahimpien infektiosairaaloiden listaaminen avoimen datan portaalissa (data.gov.uk) kannusti parhaiden käytäntöjen jakamiseen sairaaloiden kesken (Gapgemini Consulting, 2013). Capgeminin selvityksen mukaan infektiopotilaiden vuotuinen määrä väheni noin 5000 potilaasta alle 1200 infektiopotilaaseen. Avoimen datan avulla voidaan tehostaa myös esimerkiksi julkista liikennettä ja parantaa palveluja.<sup>9</sup> Esimerkiksi tietoja

---

<sup>7</sup> Lähde: <http://www.resetsanfrancisco.org/better-government/opacity-and-real-time-information-saves-san-francisco-over-1-million/>.

<sup>8</sup> Lähde: <http://www.economist.com/news/britain/21567980-how-scrutiny-freely-available-data-might-save-nhs-money-beggar-thy-neighbour>.

<sup>9</sup> HSL:n (Helsingin seudun liikenteen) esimerkki osoittaa, että datan avaaminen ja avoimien rajapintojen kehittäminen sovelluskehittäjien käyttöön voi edistää myös joukkoliikenteen palvelutason kustannustehokasta kehittämistä. <http://www.hri.fi/fi/ajankohtaista/datan-avaaminen/hsln-joukkoliikennedata/>

matkustajamääristä eri aikoina/paikoissa voidaan käyttää asiakkaiden tarpeisiin ja kysyntään paremmin vastaavien bussi- ja juna-aikataulujen suunnitteluun.

Julkisten tietovarantojen avaaminen voi edistää myös koulutusta ja tutkimusta. Erilaisten aineistojen saataville tulo parantaa edellytyksiä tutkia erilaisia yhteiskunnallis-taloudellisia ilmiöitä. Kerättyjen aineistojen ja niistä tehtyjen selvitysten tai tutkimusten jakamisella julkisen sektorin organisaatioiden välillä voidaan vähentää päällekkäistä työtä. Samoin koulutuksessa jo kerättyjen aineistojen ja tehtyjen oppimateriaalien avoin saatavuus voi sekä vähentää päällekkäistä työtä että edistää koulutusta.

### **Realisoituvatko avoimen datan hyödyt kansantalouden tasolla?**

Mikäli avointa dataa pystytään hyödyntämään riittävän laajalti talouden eri sektoreilla, voivat julkisten tietovarantojen avaamisen vaikutukset näkyä myös kansantalouden tasolla. Vanhojen yritysten kasvu ja uusien yritysten syntyminen voi näkyä myös koko talouden kasvuna. Toiminnan tehostuminen laajalti julkisen sektorin organisaatioissa ja yksityisen sektorin tuottavuuden kasvu voi puolestaan nostaa talouden aggregaattituottavuutta.

Useissa eri raporteissa on esitetty ennakoivia (ex ante) arvioita avoimen datan taloudellisesta vaikuttavuudesta kansantalouksien ja jopa globaalilla tasolla. Avoimen datan taloudellinen potentiaali on nähty raportoiduissa ennakkolisissa arvioinneissa erittäin suurena. Mittaamisen ja etukäteisarvioinnin vaikeudesta kertoo se, että eri tahojen arviot avoimen datan taloudellisten vaikutusten suuruusluokasta poikkeavat toisistaan huomattavasti. Esimerkiksi Dekkersin ym. (2006) kyselytutkimus julkisen sektorin tietovarantojen ylläpitäjille ja uudelleenkäyttäjille Euroopan eri maissa paljasti vastanneiden henkilöiden näkemyksissä ja arvioissa suuria eroja heidän kotimaansa julkisen tiedon markkinoiden arvosta. Vastausten perusteella Euroopan julkisen tiedon markkinoiden arvon arvioitiin olevan 10 ja 48 miljardin euron välillä. Vickeryn (2011) laskelmien mukaan julkisten tietovarantojen hyödyntämisestä syntyvien markkinoiden arvo EU-alueella oli vuonna 2008 noin 28 miljardia euroa, ja tietovarantojen avaamisesta syntyvien markkinoiden vuotuinen arvo noin 40 miljardia euroa. McKinsey (2014) arvioi, että avoin data (sisältäen julkisen sektorin tiedon lisäksi myös yksityisesti tuotetun koneluetettavan datan) voi tuottaa globaalisti yli kolmen triljoonan dollarin



vuosittaisen lisäarvon seitsemälle keskeiselle avoimesta datasta hyötyvälle talouden sektorille<sup>10</sup>.

Käytettävissä olevien tietojen perusteella ei pystytä arvioimaan siis kovinkaan tarkasti sitä, missä määrin julkisten tietovarantojen avaaminen on vaikuttanut ja tulee vaikuttamaan työllisyyteen, tuottavuuteen ja kasvuun kansantalouden tasolla. Tässä yhteydessä on hyvä muistaa, että vain hyvin harvoilla innovaatioilla (esim. sähkö, höyry, tieto- ja viestintäteknologia) on havaittu olevan merkittäviä taloudellisia vaikutuksia makrotasolla.<sup>11</sup> Tällöinkin vaikutukset ovat tyypillisesti realisoituneet vuosikymmenien viiveellä innovaatioiden ensimmäisestä käyttöönotosta.

### **Avoimen datan strategian riskit ja niiden hallinta**

Keskustelu avoimen datan vaikuttavuudesta on kirjallisuudessa keskittynyt hyötyjen (ja jossain määrin kustannusten) arvioimiseen. Julkisen tiedon avaamisen potentiaalisista ongelmista tai riskeistä keskusteltaessa tyypillisesti esille nousee perustuslailla turvattu kansalaisten oikeus yksityisyyden suojaan ja sen julkisen datan avaamiselle asettamat rajoitukset. Useiden tietojen osalta on varsin selvää (esim. potilastiedot), ettei niitä voida lainsäädännön puitteissa avata ilman asianosaisten lupaa.<sup>12</sup> Esimerkiksi paikkatietojen osalta tilanne on monimutkaisempi. Tietoja voidaan julkaista vain niin, että henkilötasolla tiedot pysyvät anonyymeinä. Se, milloin tietoja voidaan pitää anonyymeina, on kuitenkin ratkaistava tilannekohtaisesti. Tietosuojavaltuutetun toimiston tietosuojaryhmän mukaan ”*Se, voiko tietojen avulla tunnistaa henkilön ja voidaanko tietoja pitää anonyymeinä vai ei, riippuu tietosuojaryhmän mukaan tilanteesta. Kutakin tapausta on tarkasteltava erikseen ottaen huomioon, missä määrin kohtuullisesti toteutettavissa olevia keinoja todennäköisesti käytetään kyseisen henkilön tunnistamiseksi*”.<sup>13</sup>

---

<sup>10</sup> McKinsey (2014) avoimen datan vaikuttavuutta seuraavilla talouden sektoreilla: koulutus, liikenne, kuluttajatuotteet, sähkö, öljy ja kaasu, terveydenhuolto ja rahoitusala.

<sup>11</sup> Tällaisia laajalti yli talouden eri sektoreiden hyödynnettäviä ja talouskasvua makrotasolla edistäviä teknologioita kutsutaan yleiskäyttöisiksi teknologioiksi.

<sup>12</sup> Henkilötiedoista voi kuitenkin tulla avointa dataa asianomaisten luvalla. Esimerkiksi vaikeasta sairaudesta kärsivä voi olla halukas jakamaan itseään koskevia terveystietoja lääketieteellistä tutkimusta edistääkseen (Poikola ym., 2014).

<sup>13</sup> Lähde: <http://www.tietosuoja.fi/fi/index/ratkaisut/matka-aikatiedontuottaminenmatkapuhelinpaikannuksenavulla-paikkatietohenkilötietojavalillinentunnistettavuus.html>

Avoimen julkisen datan käytön potentiaaliset hyödyt on arvioitu monissa tutkimuksissa varsin mittaviksi. Nämä hyödyt eivät kuitenkaan välttämättä realisoidu ainakaan täysimittaisesti. Julkisen sektorin organisaatioiden liian vähäinen kapasiteetti hyödyntää avattavia tietoja on yksi mahdollinen aineistojen avaamisen hyödyntämisen este. Jos avattavia tietoja ei organisoida ja tarjota riittävän helposti löydettävässä muodossa, saattaa avatun tiedon suuri määrä vaikeuttaa tarvittavan tiedon löytämistä. Aineistojen hyödyntäjien näkökulmasta on tärkeää, että avaamisen jälkeen tietoja päivitetään ja pidetään ajantasaisina ja laadultaan korkeatasoisina. Mikäli dataa keräävissä ja hallinnoivissa julkisissa organisaatioissa ei ole riittävää tietotaitoa ja kannustimia ylläpitää ja päivittää avattuja aineistoja<sup>14</sup>, tämä voi johtaa vanhentuneen tiedon käyttöön, epäoptimaalisiin päätöksiin ja vanhentuneen tiedon jatkojalostamiseen ja leviämiseen. Julkishallinnon omien tietojärjestelmien rakentaminen niin, että ne hyödyntävät avoimesti muille tarjottavia rajapintoja ratkaisisi avoimen datan ajantasaisuuteen liittyvän ongelman ja vähentäisi ylläpitokustannuksia. Kuntien valtiolle toimittama data voitaisiin myös helposti saattaa avoimeksi näiden rajapintojen kautta. Laadukas avoin data ei kuitenkaan välttämättä riitä: myös käyttäjien osaamisen puute voi johtaa väärin tulkintoihin ja päätöksiin sekä tietovarantojen alihyödyntämiseen sekä kotitalouksissa että yrityksissä.

Yhtenä avoimen datan riskinä on kirjallisuudessa mainittu myös digitaalisen eliitin muodostuminen ja eriarvoistumisen lisääntyminen. Kaikilla kansalaisilla ei ole samanlaisia valmiuksia ja kykyjä hyödyntää avattuja tietovarantoja. Siten avoin julkinen data ei välttämättä automaattisesti, ilman muita toimenpiteitä johda osallistavan demokratian lisääntymiseen.

Julkkisten tietovarantojen avaamisen jälkeisten jatkotoimenpiteiden huolellisella suunnittelulla (esim. datojen päivitysten suhteen) julkisen datan avaamiseen ja käyttöön liittyvät riskit voidaan minimoida. Sekä tiedon tarjoajien että kansalaisten koulutuksella (esim. tiedonlukutaidon ja lähdekriittisyyden opettamisella) on merkittävä rooli. Huijboom ja van der Broek (2011) listaavat keinoja, joilla avoimen datan strategiaa on toteutettu. Käytetyt keinot - joita voidaan monelta osin hyödyntää myös julkisen datan avaamiseen liittyvien riskien hallintaan - voidaan jakaa seuraavaan neljään pääluokkaan:

---

<sup>14</sup> Peled (2011) raportoi, että toukokuussa 2011 Yhdysvalloissa avoimen datan portaali data.gov:in tietojen tarjoamiseen osallistuneista 169 organisaatiosta liki 30 prosenttia ei ollut tehnyt portaaliin datapäivityksiä vähintään vuoteen.

- 1) koulutus (esim. erilaiset alustat, joilla käyttäjät voivat jakaa kokemuksiaan ja ideoitaan avoimeen dataan perustuvista projekteista, ohjeet ja konferenssit, workshopit ym.),
- 2) oma-aloitteinen toiminta kansallisella tasolla julkisen datan avaamisen edistämiseksi
- 3) taloudelliset keinot (esim. kilpailut, sovellustenkehityskisat, avointen dataportaalien rahoitus) ja
- 4) lainsäädäntö ja valvonta (esim. julkisen sektorin tietovarantoja koskevat lainsäädännölliset toimenpiteet, teknologisten yhteentoimivuutta lisäävien standardien kehittäminen, avoimen datan projektien seuranta/valvonta).

#### **4. Avoimen datan vaikuttavuuden seuranta- ja arviointimallin kehittäminen**

Julkisten tietovarantojen avaamisen toteutuneita yhteiskunnallis-taloudellisia vaikutuksia ei pystytä kattavasti arvioimaan olemassa olevan tai tällä hetkellä säännöllisesti kerättävän tiedon valossa. Arviointiin tarvitaan seurantamallin ja mittariston kehittämistä ja tämän jälkeen systemaattista tiedonkeruuta vaikutusten jälkikäteiseen mittaamiseen. Systemaattista tiedonkeruuta varten tässä kartoitetaan alustavasti sitä, mitä mittareita voitaisiin käyttää, minkälaista hyödynnettävissä olevaa aineistoa jo kerätään ja minkälaista lisäaineistoa tarvitaan. Tämä tarjoaa pohjan seurantamallin ja mittariston yksityiskohtaiselle kehitystyölle. Se kattaa pääasiassa mitattavat, taloudelliset vaikutukset, mutta sitä voidaan täydentää laadullisella arvioinnilla.

Mittariston kehittämisen ja aineiston keruun jälkeen voidaan käyttää (osittaista) kustannus-hyötyanalyysia arvioimaan julkisen sektorin tietovarantojen avaamisesta seuraavaa hyvinvoinnin muutosta. Kustannus-hyötyanalyysissa pyritään määrittämään kaikille avoimen datan vaikutuksille rahamääräinen arvo, ja laskemaan tämän perusteella vaikutusten kokonaisarvo ja kannattavuus yhteiskunnan näkökulmasta. Kustannus-hyötyanalyysi lienee käytännössä toteutettavissa vain keskeisten vaikutusten osalta, joten sen avulla pystytään tekemään arvio avoimen datan vaikuttavuuden ”alarajasta”. Lisäksi voidaan avointen julkisten tietovarantojen hyödyntämisen vaikuttavuutta arvioida laadullisesti<sup>15</sup>. Tältäkin osin keskityttäneen niihin potentiaalsiin vaikutuksiin, joita pidetään yhteiskunnallisesti tärkeinä,

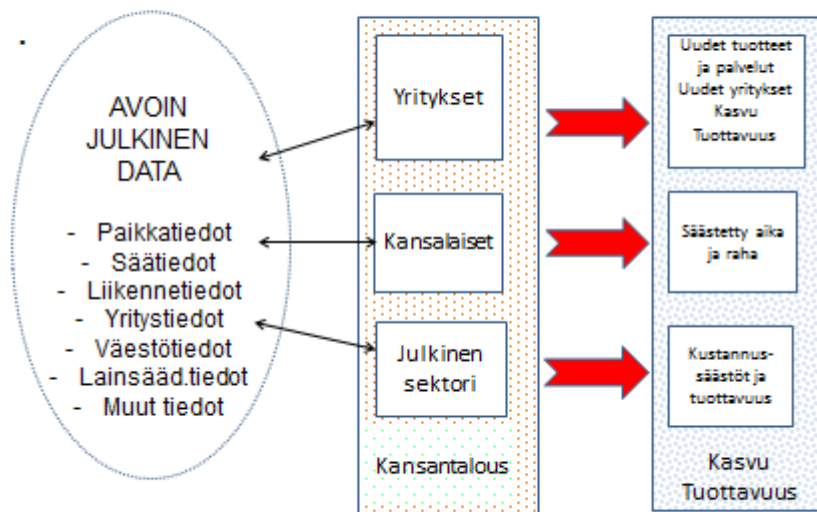
---

<sup>15</sup> Laadullista arviota datan avaamisen vaikutuksista on tehty esimerkiksi Liikenne- ja viestintäministeriön (2013) raportissa.

mutta joiden euromääräinen mittaaminen on erittäin vaikeaa (esim. kansalaisaktiivisuuden lisääntyminen, demokratian toteutuminen).

Yksi keskeisistä avoimen datan arvioituista vaikutuksista yrityksissä liittyy avoimen datan käyttöön uusissa tuotteissa ja palveluissa. Onnistuneet innovaatiot ovat sekä tuottavuuden että taloudellisen kasvun keskeinen lähde. Avoimen datan vaikutukset yritysten innovaatioimintaan ja laajemmin yritysten kasvuun ja tuottavuuteen ovat arviointimallin ja seurannan keskiössä yritysvaikutusten osalta. Liikevaihdon ja työntekijöiden määrän kasvu ovat usein käytettyjä yrityksen kasvua kuvastavia mittareita. Kokonaistuottavuus mittaa aikaan saatua tuotosta suhteessa käytettyihin panoksiin eli yrityksissä sitä kuinka paljon tuotteita ja palveluita saadaan tuotettua käytetyllä työllä ja pääomalla. Tyypillisesti yrityksen tuottavuuden mittarina käytetään työn tuottavuutta, joka lasketaan jakamalla arvonlisä työntekijöiden lukumäärällä tai tehtyjen työtuntien lukumäärällä. Yritysten toiminnasta kerätään vuosittain taloudellisia tietoja, joita voidaan käyttää yritysten kasvun ja tuottavuuden mittaamisessa.

*Avoim data: merkittävimmät odotetut taloudelliset vaikutukset*



Kansalaisten ja kotitalouksien osalta julkisen datan käytön euromääräisten hyötyjen arvioiminen on haasteellista. Periaatteessa niiltä osin kuin aiemmin maksullista tietoa on

tarjolla ilmaiseksi, voitaisiin erilaisten tietojen käyttömäärien perusteella tehdä karkea arvio kansalaisten tiedon hankkimisesta saamasta rahallisesta säästöstä. Käytännössä tieto olisi saatavilla vain kyselytutkimuksen avulla, koska Internetin avoimen datan palveluiden käyttötilastot eivät erittele kansalaisia muista käyttäjistä. Avoimen datan käytön avulla säästetyn ajan ja rahan määrän arviointi lienee turhan vaikea kysymys vastattavaksi. Kyselytutkimuksen avulla voitaisiin kuitenkin arvioida karkeammalla tasolla avoimen datan käyttöä ja sitä, minkälaisia julkisia tietoja kansalaiset käyttävät sekä julkisten tietovarantojen hyödyntämisen vaikuttavuutta kotitalouksissa.

Tärkeän avoimen datan vaikuttavuuden arvioinnin osa-alueen muodostavat julkisen sektorin organisaatiot. Julkisten tietovarantojen avaamisella on vaikutuksia sekä dataa avanneisiin organisaatioihin että avattuja tietoja hyödyntäviin organisaatioihin. Avoimesta datasta saatavat odotetut kustannussäästöt ja tuottavuusvaikutukset ovat erityisen kiinnostuksen kohteena julkisella sektorilla. Julkisen hallinnon tuottavuutta mitataan tuotettujen palvelujen määrän ja palvelujen tuottamiseen käytettyjen panosten määrämuutosten perusteella. Vaikutusten kartoittamiseksi tulisi pohtia voitaisiinko kehittää menetelmä tai tekninen ratkaisu, jonka avulla pystytään mahdollisimman pienin kustannuksin mittaamaan ja seuraamaan (ainakin keskeisten) julkisen hallinnon organisaatioiden avoimen datan sekä muiden palveluiden tarjontaa ja käyttöä sekä palveluiden tarjontaan käytettäviä resursseja.

Vuosittaisin toteutettavalla ”Tietoja valtion tietohallinnosta” –kyselyllä kerättävien tietojen hyödynnettävyyttä avoimen datan vaikuttavuuden mittaamisessa on syytä arvioida. Lisäksi on syytä pohtia tarvetta lisätä kyselyyn kysymyksiä avoimen datan vaikuttavuuden mittaamiseksi. Erityisen tärkeää olisi saada arvio datan avaamisen aiheuttamasta lisätyöstä tai säästetystä ajasta julkisen sektorin organisaatioissa. Tilastoitujen ammattiryhmien työajan keskihintojen avulla voidaan tällöin laskea hinta datan avaamisesta aiheutuneelle lisätyölle tai sen avulla saaduille aikasäästöille. Tilastokeskus kerää tietoja ”Valtion tuottavuustilastoon” vuosittaisella kyselyllä valtion tilivirastojen ja laitosten tuotoksesta, työ- ja kokonaispanosten määrämuutoksista sekä näistä lasketuista työn ja kokonaistuottavuuden muutoksista. Myös tuottavuustilastojen käytettävyyttä vaikuttavuusarvioinnissa on tarpeen pohtia.

Alla esitetty taulukko tiivistää avoimen datan odotetut keskeiset vaikutukset ja esittää mittareita vaikuttavuuden systemaattiseen seurantaan ja arviointiin. Tämän lisäksi tarvitaan mittareita avoimen datan hyödyntämisestä ja käytöstä talouden eri sektoreilla. Avoimen datan

käytöstä ja sen roolista yritysten liiketoiminnassa, kotitalouksissa ja julkisilla sektorilla ei kerätä tällä hetkellä systemaattisesti tietoja. Pääasiallinen keino saada tietoa avoimen datan käytöstä ovat kyselytutkimukset. Kansalaisten avoimen tiedon hyödyntämistä voitaisiin kartoittaa esimerkiksi Tilastokeskuksen vuosittain toteuttaman kotitalouksille suunnatun ”Tieto- ja viestintätekniiikan käyttötutkimus” –kyselyn puitteissa.

*Avoimen datan taloudellisten vaikutusten arviointi: odotettujen vaikutusten mittarit*

	Odotettu vaikutus	Mittarit	Potentiaalisia tietolähteitä
Yritykset	Uudet tuotteet ja palvelut Uudet yritykset	Avoimeen dataan perustuvien uusien tuotteiden ja palveluiden määrä Avoimen datan palveluiden/tuotteiden tuottamien tulojen osuus liikevaihdosta	Kyselytutkimukset; kysymysten lisääminen esim. ”Tietotekniikan käyttö yrityksissä” –kyselyyn ja CIS –kyselyyn).
	Kasvu	Liikevaihdon muutos, työntekijöiden määrän muutos	Tilastokeskuksen yritysrekisteri Asiakastiedon tilinpäätöstietokanta
	Tuottavuus	Arvonlisä per työntekijä tai tehdyt työtunnit	Asiakastiedon tilinpäätöstietokanta
Kansalaiset	Säästetty aika ja raha	Käytetyn ilmaisen tiedon määrä x aiempi hinta Arvio aikasäästöistä	Kyselytutkimukset; kysymysten lisääminen ”Tieto- ja viestintätekniiikan käyttötutkimus” –kyselyyn
Julkinen sektori	Kustannussäästöt Tuottavuus	Käytetyn työajan muutos x työajan hinta Tuotetut palvelut suhteessa käytettyihin resursseihin	Automatisoitu tietojen ja palvelujen tarjonnan seuranta Kyselytutkimukset; Tietoja valtion tietohallinnosta –kysely ja siihen lisättävät kysymykset Valtion tuottavuustilasto
Kansantalous	Kasvu Tuottavuus	Arvonlisän eli BKT:n muutos BKT asukasta kohden	Tilastokeskuksen tilastotietokannat

Tilastokeskuksella on lakisäätöisiä säännöllisesti yrityksille tehtäviä kyselyjä, joihin voitaisiin lisätä kysymyksiä avoimen datan käytön vaikuttavuuden systemaattisen seurannan tukemiseksi. Tällaisia kyselyjä ovat muun muassa vuosittain toteutettava ”Tietotekniikan käyttö yrityksissä” –kysely ja joka toinen vuosi toteutettava CIS (Community Innovation Survey) –kysely. Keskeisiä kartoitettavia kysymyksiä ovat muun muassa missä määrin ja

mihin tarkoituksiin (esim. innovaatiotoiminta, logistinen suunnittelu) avointa dataa yrityksissä käytetään sekä minkä tyyppistä dataa (esim. paikkatietoja, yritystieto) yritykset toiminnassaan hyödyntävät. Kyselytutkimuksen avulla voitaisiin kartoittaa myös avoimen datan roolia uusien yritysten perustamisessa, niiden toiminnassa ja työpaikkojen luomisessa.

Usein sama julkisen sektorin organisaatio sekä tarjoaa että käyttää avointa dataa. Eräs mittari tietovarantojen käytölle julkisella sektorilla on julkishallinnon organisaatioiden keskenään jakaman tiedon määrä. Mahdollisimman pitkälle automatisoitu tietojen käyttömäärien ja julkisen sektorin organisaatioiden välillä liikkuvien tietojen seurannan käyttöä avoimen datan vaikuttavuuden tutkimuksessa julkisen sektorin organisaatioiden osalta olisi eräs pohdittava vaihtoehto. Koko kansantalouden tasolla avoimen datan tarjontaa ja käyttöä voidaan mitata esimerkiksi avatun tiedon määrän ja sen käytön määrän sekä käyttäjien määrän perusteella. Näiden aggregaattimittareiden avulla voidaan seurata tietovarantojen avaamista ja saada tietoa siitä, mitkä aineistot ovat käyttäjille tärkeitä. Myös avointa dataa hyödyntävien uusien sovellusten ja palveluiden määrä heijastaa avoimen datan hyötyjä. Systemaattisten tietojen kerääminen tältä osin on kuitenkin haasteellista.

Kyselytutkimuksilla avoimen datan käytöstä kerättäviä tietoja voidaan hyödyntää avoimen datan vaikuttavuuden tutkimuksessa yhdistämällä ne erilaisiin tilastoaineistoihin kuten Tilastokeskuksen yritysrekisteriaineistoon ja Asiakastiedon tilinpäätöstietokantaan. Pitkäaikainen avoimen datan käytön seuranta ja tilastointi eri sektorien osalta mahdollistaa avoimen datan taloudellisen vaikuttavuuden kvantitatiivisen analyysin kehittyneitä tilastollisia menetelmiä käyttäen. Muiden avoimen datan yhteiskunnallisten vaikutusten tutkimuksessa voidaan myös hyödyntää kyselytutkimuksia. Muiden vaikutusten analysointiin sopivat pääasiassa laadulliset tutkimusmenetelmät.

## 5. Yhteenveto

Julkisten tietojen avaamisen potentiaaliset hyödyt on ennakkolisissa arvioinneissa arvioitu mittaviksi, ja siksi julkisten tietovarantojen avaaminen on noussut useissa maissa keskeiseksi politiikkateemaksi. Suomessa on tavoitteena avata kaikki merkittävät julkishallinnon keräämät ja hallinnoimat tietovarannot vuoteen 2020 mennessä. Julkisen datan avaamisen yhteiskunnallis-taloudellisen vaikuttavuuden tutkimus on kuitenkin vasta aluillaan. Tällä

hetkellä avoimen datan vaikutuksista ja parhaista käytänteistä tiedetään vain yksittäisiltä alueilta ja niiltäkin suureksi osaksi case-esimerkkien valossa. Tietävästi missään maassa ei ole vielä arvioitu kattavasti avoimen datan *toteutunutta* vaikuttavuutta ja hyötyjä. Suurimpia syitä tähän ovat mittaamisvaikeudet ja se, ettei julkisten tietojen käytöstä ole olemassa tilastoitua tai systemaattisesti kerättyä tietoa.

Suomi on tällä hetkellä julkisen tiedon avaamisen kärkimaiden joukossa, ja sillä on mahdollisuus myös nousta edistyneisimpien maiden joukkoon avoimen datan vaikuttavuuden tutkimuksessa. Tämä edellyttää kuitenkin huolellista avoimen datan seuranta- ja arviointimallin kehittämistä sekä systemaattista tiedon keräämistä avoimen datan vaikuttavuuden tutkimiseksi ja parhaiden käytäntöjen selvittämiseksi. Tässä raportissa on esitetty alustavia suuntaviivoja avoimen datan vaikuttavuuden arvioinnin kehittämiseksi erityisesti taloudellisten vaikuttavuusmittarien osalta.

Tilastollisen vaikuttavuusmittareita hyödyntävän aineistoanalyysin lisäksi voidaan sekä kansalaisten, yritysten että julkisen sektorin organisaatioiden osalta arvioida tarjolla olevien julkisten tietovarantojen ja erilaisten aineistojen käytettävyyttä ja hyödyllisyyttä käyttäjien omien arvioiden perusteella. Luonteva paikka käyttäjälähtöisen arvioinnin ja palautteen antamiselle on avoimen datan palvelu, jossa tietovarannot ovat käyttäjille tarjolla. Lisäksi olisi syytä miettiä keinoja, miten tuotetaan tietoa avoimen datan hyödyntämiseen liittyvistä parhaista käytännöistä erityisesti julkisen sektorin organisaatioissa. Tärkeä kysymys on myös, millä keinoin nämä parhaat käytännöt saadaan mahdollisimman tehokkaasti leviämään ja hyödynnetyksi eri organisaatioissa.



## Lähteet

Ahonen-Rainio, P., Mäkelä, J. ja Virrantaus, K. (2014) Menetelmä avoimen maastotiedon vaikuttavuuden arvioimiseksi. Aalto-yliopisto, Maankäyttötieteiden laitos.

Bargiotti, L., De Keyzer, M., Goedertier, S. ja Loutas, N. (2014). Value based prioritisation of open government data investments. European Public Sector Information Platform, Topic Report No. 2014/08.

Brynjolfsson, E., Hitt, L. M. ja Kim, H.H. (2011). Strength in Numbers: How does data-driven decision-making affect firm performance? Working Paper Available at SSRN: <http://ssrn.com/abstract=1819486>.

Capgemini Consulting, 2013; The Open Data Economy. Unlocking Economic Value by Unlocking Government and Public Data.

Davies, T. Perini, F. ja Alonso; J. M. (2013). researching the emerging impacts of open data. ODDC conceptual framework. World Wide Web Foundation.

Dekkers, M., Polman, F., te Velde, R., ja de Vries, M. (2006). MEPSIR. Measuring European Public Sector Information Resources. June 2006.

Fornefeld, M. Boele-Keimer, G.Recher, S. and Fanning, M. (2009). Assessment of the re-use of public sector information (PSI) in the geographical information, meteorological information and legal information services. Micus Management Consulting.

Huijboom, N. ja Van der Broek, T. (2011). Open data: an international comparison of strategies. *European Journal of ePractice*, March/April 2011, 1-13.

Kiuru, P., Mäkelä, J. ja Huvio, P. (2012). Avoimen jlkisen tiedon hyödyntämisen potentiaalista suomalaisissa yrityksissä. Aalto-yliopiston raportti.

Koski, H. (2011). Does marginal cost pricing of public sector information spur firm growth? *Etla Discussion Paper* No. 1260.

Koski, H. (2012). The role of data and knowledge in firms' service and product innovation. *Etlä Discussion Paper* No. 1272.

Liikenne- ja viestintäministeriö (2013). Liikenteen ja viestinnän avoin tieto. Työryhmän raportti. *Liikenne- ja viestintäministeriön julkaisuja 10/2013*.

[http://www.lvm.fi/c/document\\_library/get\\_file?folderId=2497123&name=DLFE-19419.pdf&title=Julkaisuja%2010-2013](http://www.lvm.fi/c/document_library/get_file?folderId=2497123&name=DLFE-19419.pdf&title=Julkaisuja%2010-2013)

McKinsey Global Institute (2011). Big data. The next frontier for innovation, competition and productivity.

[http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation).

OECD (2013). *Supporting investment in knowledge capital, growth and innovation*. OECD Publishing. doi: [10.1787/9789264193307-en](https://doi.org/10.1787/9789264193307-en)

Peled, A. (2011). When transparency and collaboration collide: The USA open data program. *Journal of the American Society for Information Science and Technology* 62, 2085-2094.

Pettifer, R.E.W. (2009). PSI in European meteorology – An unfilled potential. Mimeo. 25.4.2009.

Poikola, A, Kuikkaniemi, K. ja Kuittinen, O. (2014). My data – johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen. Liikenne- ja viestintäministeriö.

<http://www.lvm.fi/julkaisu/4420389/my-data-johdatus-ihmiskeskeiseen-henkilotiedon-hyodyntamiseen>.

Pollock, R. (2008). The economics of public sector information. University of Cambridge. Mimeo. December 2, 2008.

Salokannel, M. (2012). Opening-up public data – a spur for democracy and information economy. *IPRinfo* 2/2012.

Ubaldi, B. (2013) Open government data: Towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance*, No. 22, OECD Publishing.

Verhulst, S., Noveck, B.S, Caplan, R., Brown, K ja Paz, C. (2014). The open data era in health and social care. The governance lab, NHS, England.

Vickery, G. (2011). Review of recent studies on PSI re-use and related market developments. Information Economics, Paris. <https://ec.europa.eu/digital-agenda/en/news/review-recent-studies-psi-reuse-and-related-market-developments>