

## PERSONAL VERSION

This is a so-called personal version (author's manuscript as accepted for publishing after the review process but prior to final layout and copyediting) of the article: Lindman, J, Kinnari, T & Rossi, M 2014, 'Industrial open data: Case studies of early open data entrepreneurs' in *Proceedings of the 47th Hawaii International Conference on System Sciences*.  
[http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=6758695](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6758695)

This version is stored in the Institutional Repository of the Hanken School of Economics, DHANKEN. Readers are asked to use the official publication in references.

# Industrial open data: Case studies of early open data entrepreneurs

Juho Lindman  
Hanken School of  
Economics  
[juho.lindman@hanken.fi](mailto:juho.lindman@hanken.fi)

Tomi Kinnari  
Aalto University School of  
Business  
[tomi.kinnari@aalto.fi](mailto:tomi.kinnari@aalto.fi)

Matti Rossi  
Aalto University School of  
Business  
[matti.rossi@aalto.fi](mailto:matti.rossi@aalto.fi)

## Abstract

*Open data entrepreneurship is required to create novel services and sustainable value networks based on government released datasets. However, the business is still in its infancy. This paper investigates the emerging open data value network structure based on empirical findings from 14 Finnish organizations. The data was collected through interviews of early adopter open data entrepreneurs during the spring of 2012. We identified business models such as: saving costs with co-creation, creating new user interfaces by combining data from several sources, and analyzing and visualizing data. Understanding the business models, as well as the emerging OD value network, will help companies to better reap the benefits of open data, while contributing to academic discussion on how to establish an open data service ecosystem.*

## 1. Introduction

Many have touted open data as the oil of 21<sup>st</sup> century (e.g. HICSS-45 keynote by Mr. Vivek Kundra), but viable business models of open data have not yet materialized. It can be argued that open data is in a stage, where commercial successes and widely adopted services based on it are needed to demonstrate its potential. To emerge, open data needs examples of long-term successes that impact business or quality of life. The datasets released thus far include for example scheduling data for public transport, weather data, epidemic data and population statistics. The enrichment of this data can take place through linking it to available sensor data, or the creation of location based tags or services that use contextual data on handheld mobile devices. An example of these is the Finnish Meteorological Institute that releases weather radar data as open data.

As the amount of technically and legally open data grows, so does the need to understand the dynamics and business cases of services made possible by the release of previously closed data sources. As noted

earlier [1, 2] there is a research gap concerning the viability of the service businesses based on open data. Most research to date has been dealing with the initial release of the datasets and related technical issues [3].

This empirical paper reports the experiences of pioneer entrepreneurs of open data, who build their company offering on released open datasets. Our research questions are:

- 1) *what are the actors and their positions in the emerging value network and*
- 2) *which business and revenue models are utilized by the early open data entrepreneurs?*

We limit our empirical investigation to Finnish entrepreneurs, but speculate that similar approaches are likely also in other national contexts.

The paper is structured as follows: First the paper defines open data and then proceeds with a review of relevant literature on open data, business models and value networks and their interplay. The empirical part that follows draws on the experiences accumulated over a two-year research project. The paper reports the results of our exploratory case study of 14 companies focusing on analyses of the industry landscape and the business models of the companies. The findings include a description of the emerging value network and the different profiles of the actors and their positions from the perspective of the industry data stream.

## 2. Open data business

Service businesses can use different business models [4-6] to capture value. Work on open data business builds on earlier work on legal, technical and commercial requirements of data openness [2].

The terms of the released governmental or private sector datasets allow commercial organizations to build novel services on the data. However, service design is crucial for the users to find and continue to use the services (ibid.).

## 2.1. Open Data definition

The definition of Open Data in the scope of this article is: *Data, which is legally accessible through the Internet in a machine-readable format.* It does not have to be completely free of charge or restrictive licenses, but experimenting with the data, and running a small-scale-business should be legal. This exploratory usage of the data may also help the data owner to see the potential of data re-usage, influence a change in position on data re-usage permission, allowing the developers to access the data more easily. This developer “activism” is very common in the open data community (Kuk & Davies, 2011).

Technically open data needs to be in a linked format or in another format that is easily readable by a computer (for example comma-separated values (.csv), Excel spread sheet (.xls), or even PC-axis (.px) formats). In addition, all websites and text documents are included. However, scanned paper documents (.pdf) or other image files are not considered machine-readable in this paper.

This definition complements the open data definition of OKF (Open Knowledge Foundation): *accessible as a whole, free-of-charge or at most with a reasonable reproduction costs, redistributable, reusable, in a data format which does not cause technological obstacles, and without discrimination against persons or groups nor against any particular fields of endeavor* [7]. We extend this definition to include situations when companies scrape data from a website without an explicit legal permission or mash up commercial and open data together from several sources to create a better user experience.

## 2.2. Open Data services value network

Services consist of a service act, a type of customer relationship, the customizability of the service, the nature of demand, a delivery mechanism, and the attributes of the service product [8]. Service-dominant logic assumes that exchanges are fundamentally based on a service as an application of competencies to benefit another [6].

In service computing services are considered acts performed by one entity for a different entity, the entities including human actors and organizations, as well as computerized services systems. Services can be described through their intrinsic (the service itself), and extrinsic (the networks needed for service development, provisioning and consumption) attributes [9, 10].

A linked data value chain has four entities: the raw data provider, the linked data provider, the linked data

application provider and the end-user [11]. The raw data provider publishes raw data, the linked data producer utilizes the raw data to produce linked data, and finally the application provider utilizes the linked data to produce a human readable output for human end-users.

These roles are closely connected with three types of data artifacts: raw data, linked data, and human-readable data [11]. In addition, data service provide consultancy services as a revenue source [12] and thus the main revenue sources of open data are consulting, transforming raw data into linked open data, and developing applications on top of the data.

Poikola, Kola & Hintikka in their 2010 (Finnish) book [13] list 10 roles in the open data value chain. Seven of these roles are considered from the data publishing perspective: Data recorder, data refiner, data aggregator, data harmonizer, data updater, data publisher, and registry maintainer.

In addition, they see three end-users for the data: an application developer utilizing the data as part of their service; a data interpreter utilizing data in their research, commercial, or democratic activities; a human, a company, or an organization as an end-user utilizing these applications or interpretations; a data updater, a registry maintainer, a data aggregator, a data harmonizer, and a data interpreter as an end-user.

Kuk and Davies [1] have studied the role of agency and artifacts in assembling open data complementarities: the resulting artifacts create a recursively independent artifact stack. The phases in this artifact stack are the cleaning of data, making the data linkable, writing software to analyze or visualize the data, sharing the source code of the software in a revision control system such as github, and finally allowing other developers innovate new services on top of the source code. Kuk and Davies [1] emphasize the uncoordinated co-operation between the agents involved in the process. This co-operation is organized around the intermediary artifacts (cleaned data, linkable data, source code, shared source code, and service technologies), each of which reinforces the value of the previous artifact [1].

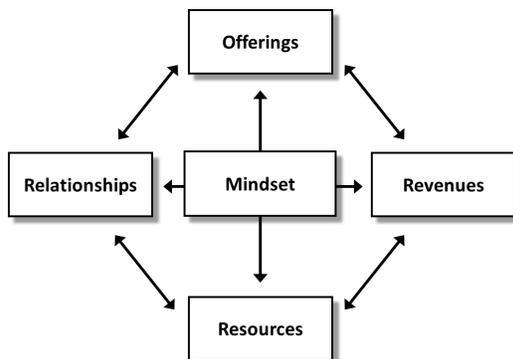
**Table 1: Summary of open data business models from earlier literature**

Topic	Roles in value chain
Linked data value chain Latif et al. [11]	Raw data provider Linked data provider Linked data application provider End user

<b>Roles in data value chain</b> <b>Poikola et al. [13]</b> <b>(Translated from Finnish)</b>	Data recorder Data refiner Data aggregator Data harmonizer Data updater Data publisher Registry maintainer Application developer Data interpreter End user
<b>Tammisto and Lindman [12]</b>	Raw data provider Open Data consultant Linked Data developer Applications developer End users
<b>Phases of artefact stack</b> <b>Kuk and Davies [1]</b>	Cleaned data (producer) Linked data (producer) Software source code (developer) Software source code (sharer) Service technologies (innovator)

### 2.3. Business model framework

A business model is an analytical tool used to understand and analyze companies better, but a discussion on the scope and exact definition of a business model is ongoing. For example Chesbrough and Rosenbloom [14] define a business model as something that “*provides a coherent framework that takes technological characteristics and potentials as inputs, and converts them through customers and markets into economic outputs*”. Research on business models often focuses on the elements they include [4]. We follow this convention and consider business models for software business to consist of offering, revenue model, resources, relationships, and mind-set [5].



**Figure 1: Business model elements as defined by Rajala [5]**

To ground our analysis, we adapt Rajala’s [5] definition of business models as “*concise representation of how an interrelated set of elements – the offering, relationships, resources, revenue model and management mind-set – are addressed to create and capture value in defined markets*”. Thus we can use this framework to analyze the five different elements separately (see figure 1 above). These separate value-creating elements are (not in any particular order):

1) An offering is a value proposition that a software firm offers its customers and other stakeholders, and with which it positions itself in the market.

2) Resources are the assets and capabilities that are needed to develop and implement a given business model. They can be tangible (personnel, equipment, etc.) or intangible (brand name, relationships, etc.). In essence, they are the internal source of advantage, or the core competency of a company.

3) Relationships are the means to access external resources and capabilities.

4) A revenue model includes the revenue sources, pricing policy, cost structure, and revenue velocity. It is the firm’s means to capture value.

5) A management mind-set distinguishes a business model as something that stems from the values, emotions, and attitudes of management instead of cognitive, rational thinking and planning.

### 3. Method

Value is created through the assemblage of open data complementarities and therefore the value network in which the companies operate is essential (see, e.g. [1]). To be able to formulate a comprehensive view, our research seeks to not only understand the business benefits of the companies, but also the wider value network which is needed for them to survive. To reach our aim we opted for an interpretative and exploratory multiple-case design that included in-depth data collection and analysis [15]. An interpretative approach helped us to take into account the interplay of the organizations and their environment [16]. The descriptions were derived from key informant interviews. In terms of systematic data collection, a series of formal face-to-face semi-structured interviews was conducted. Since the aim was to lay emphasis on the depth, nuance, complexity, and comprehensiveness of the data, interviewing was considered to be the most appropriate method for data collection.

The case companies were selected among Apps 4 Finland competition submissions. Apps 4 Finland (A4F) is an annual application innovation contest run by Forum Virium, encouraging developers to create

new applications using open data. The research examined A4F submissions from years 2009, 2010 and 2011, and during this time period there were altogether 193 submissions posted to the competition. Out of these 193 submissions, 17 had business activity during the time of research. One company developed six of these 17 works and two were developed by another one, thus resulting in 11 separate companies that were contacted for an interview. Of the contacted companies, three did not answer to our inquiry and one refused, resulting in seven interviewed companies.

These seven interviews were the starting point for the research. We extended the list of respondents as new interesting names and companies came up in the interviews. This snowball sampling technique provided seven additional respondents, thus increasing the total amount of respondents to 14. All respondents were from different organizations. Table 2 below lists the interviewed organizations. These interviews do not allow us to draw statistical generalizations, but instead serve as a comprehensive subset of pioneering companies to discuss the roles in the open data ecosystem of Finland.

**Table 2 List of interviewed organizations**

#	Organization	Description	Title	Type of dataset(s) used in business
1	Louhos	Community driven collection of open source code and tools to analyse Finnish open data	2 co-founders	Examples: voting data from parliamentary plenaries, Helsinki Region Environmental Services Authority, Statistics Finland, National Institute for Health and Welfare, and many others
2	Hahmota Oy	Financial data visualizations	CEO	Case dependent data directly from the client. Demo visualizations made with data from Statistics Finland and CIA World Fact book
3	Fresh Bits	Mobile apps, including journey planner	Software developer	Commuter traffic APIs, Helsinki Service Map
4	Gemilo Oy	Ways to improve internal communication and work practises in companies and public administration	Social network developer	HILMA electronic notification system for notices of public procurement contracts, OpenStreetMap
5	Skyhood Oy	Service platform connecting jobseekers and employers	CEO	Screen scraped vacancies from Employment and Economic Development Offices, Jobseek service Aarresaari, Statistics Finland, and some private data sources
6	Forum Innovations Oy	Discussion forum, bringing citizens together to discuss	Co-founder	Screen scraped data from the parliament, user generated data
7	Suomen Turvaprojektit Oy	Security, safety training and consultancy	CEO	Lists and regulations from IATA DGR, ICAO TI and EU on prohibited hold luggage and cabin luggage articles on flights
8	Cloud'N'Sci Ltd	Algorithm-as-a-service platform that connects third-party algorithm developers and business	CEO	Case dependent, but e.g. Statistics Finland
9	Essentia Solutions Oy	Mobile apps, including journey planner	CEO	Commuter traffic APIs, Helsinki Service Map

10	KAMU Ry	Service that allows comparison of campaign statements with the actual parliament voting records.	Chairman of the association	Voting data from parliamentary plenaries, parliament members' answers in voting aid applications
11	Logica	Multinational IT and management consultancy	Development Director	Case dependent
12	Flo Apps Ltd	Data visualization and website creation	CEO	Case dependent, but e.g. HelMet Library database, The National Biography of Finland, Commuter traffic API, OpenStreetMap
13	Suomen Asiakastieto Oy	Information services for corporate management, risk management, sales and marketing	Director	Main data sources: Finnish Trade Register, Statistics Finland, Joint business information system of the National Board of Patents and Registration and the Tax Administration, Legal Register Centre, Tax Administration, District Courts, Finnish Communications Regulatory Authority, Enforcement units. In addition, company collected data directly from companies and partners.
14	Helsingin Seudun Liikenne	Commuter traffic service in the greater Helsinki region	Project manager	N/A (Data publisher of commuter traffic API)

To answer the first RQ: "*what are the actors and their positions in the emerging value network*" the companies were grouped and placed under different value network profiles based on their offering as described by the respondents.

Thus, the companies categorized under one value network profile share similar offerings, but might differ in other aspects of their business models. At times the grouping was not straightforward as some organizations had versatile business practices occupying several positions in the value network. These borderline cases were grouped based on their primary value adding functionality.

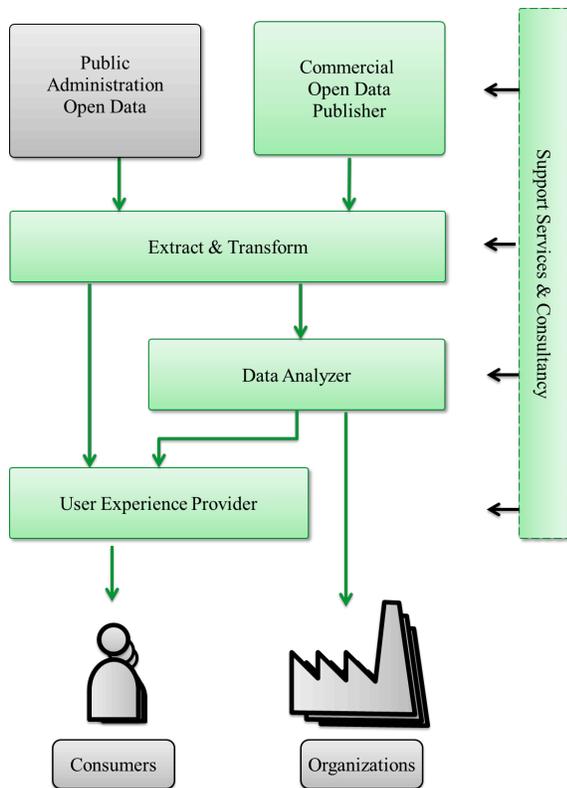
To answer the second research question, "*which business and revenue models are utilized by the early open data entrepreneurs*" the questions posed to the respondents were structured around the five business model elements [5]: offering, resources, relationships, revenues and management mind-set. The questions were adapted to each particular case company because of differences in for example the size of the company. The analysis was carried out in order to report a description of these elements in the profiles that follow.

#### 4. Findings

We identified five value network participant profiles. These are data analyzer, extract & transform,

user experience provider, open data publisher, and support service and consultation. The five profiles are highlighted in Figure 2. Arrows between the profiles represent a flow of open data artifacts. An open data artifact is refined at each step as it proceeds in the value network. During the process raw data is finally converted into a valuable user experience or analysis.

These five value network participant profiles will be described next, in more detail. Furthermore, a summary of the offerings, revenue models, resources and relationships of the five participant profiles is depicted in Table 3 later.



**Figure 2: Open data value network**

#### 4.1. Extract and transform

The Extract and transform profile searches open data from different sources and transforms data into a more readily re-usable format. A proper transformation process is required, if the succeeding analysis combines several data sources and thus demands standardized data. Typically data requires administration to ensure integrity. This administration removes double records, missing information or otherwise incorrect information. Finally the data is stored into a data format suitable for the subsequent analysis.

Most of the examined companies performed their data extraction and transformation themselves before analyzing the data further. Respondents agreed that it takes more time to find, extract, and transform a desired piece of data than it takes to perform the actual analysis. One case organisation, Louhos, focused on making extract & transform faster and simpler for other companies.

Louhos is a community driven collection of open source code and tools that help to process Finnish open data. The Louhos community releases free source code and instructions on how to fetch data from different sources, how to plot it on map and how to cross

analyze it with other data sources. In addition to this extract and transform activity, they also perform and release analyses and visualizations of data.

Currently Louhos does not have a revenue model. The organisation's resources consist of voluntary work from the co-founders and the surrounding open data community. The founders responded that they were hoping for more programmers to get involved in the project so they could start adding new features and data sources. The founders are very active in the Finnish open data scene, and they use these relationships to increase the awareness of the toolkit and to invite new active developers to join the community.

#### 4.2. Data analyzer

Data analyzers create visualizations or algorithm based analyses of data and usually cross-analyze data from multiple sources to gain more powerful results. Part of the data might be from open data sources and part from internal or commercial sources. Before the data can be analyzed, it needs to be extracted from its original source and transformed to an analyzable format. Thus data analyzers have a dual role in the value network. Our data analyzers varied quite a lot in size; the smallest were start-up companies while the largest generated 40-million euro annual revenue.

Customers paying for data analyses were commercial companies, funding agencies or public authorities. Utilized revenue models were thus business-to-business centric. We could not find services targeted to paying consumers using this profile. Therefore, there is no arrow between data analyzers and consumers in Figure 2.

As an example company from this value profile, Asiakastiето offers detailed information on Finnish companies, citizens and properties on an individual level. The respondent said that 95 % of their revenue is based on individual level information. They have developed powerful algorithms to estimate, e.g., credit ratings based on data from several public data sources, such as the Trade Register, the National Board of Patents and Registration of Finland and the Finnish Business Information System. In addition to these public data sources, they also use internally collected data, such as balance sheets and customer data acquired directly from companies. Asiakastiето also exchanges data with foreign credit rating organizations. Altogether, Asiakastiето uses tens of different data sources.

Their revenue is based on selling credit ratings or other information products to banks, retailers and other businesses. Products have been priced based on the monetary value of the risk taken away from customers.

### **4.3. User experience provider**

User experience providers utilize open data sources to create interactive services for end-users. The service can be accessed through a web-browser or with a native mobile application. In both cases, open data has a central role in the offering. Often these open data sources were amended with commercial sources in order to create more a more appealing user experience. The user experience provider was the most popular part of the value network. We found seven companies fitting within this profile. Like data analyzers, user experience providers usually needed to extract and transform the data by themselves. In addition, in some cases they also processed and analyzed the data, thus performing three activities within the value network.

The services were targeted for consumers, and therefore also the earning logic used typical business-to-consumer models. As listed in Table 3, these include advertisements, subscriptions and one-time fee.

An example company is Duunitori.fi, a service platform connecting jobseekers and employers. Their data comes from both public sources, such as the Employment and Economic Development Office, which is amended with several private data sources and social media feeds. The resulting data mash-up is visualized on a map to increase usability and clarity.

When asked about their revenue model and resources, the respondent claimed that developing services on top of open data is relatively cheap, which leads to a situation where end-users are unwilling to pay for them. Therefore, the revenue needs to come from other sources, such as advertising or by capitalizing on user data collected from jobseekers. At the time of interview, the company was still searching for the best way to monetize the service.

### **4.4. Commercial open data publisher**

Commercial open data publishers release private data resources for the free use of the public. Commercial data openers are especially interesting because they bring a new horizon in to the open data value network - instead of utilizing existing open data they increase the open data stack by releasing new data. Revenue model for releasing private data might include cost savings through crowd sourcing. Increasing transparency and facilitating new service innovation also offer benefits.

The Helsinki Region Transport Authority (HRT) is a municipal corporation responsible for planning and procuring public transportation in Greater Helsinki. In 2001 they released a web-based Reittiopas journey planner service to aid finding the fastest transportation between two points. In 2009 HRT opened an API for

developers so they can create applications using the transportation-planning engine. In May 2012, over 650 developers had registered to get access to the API and over 30 third-party applications using the API were listed on HRT webpage. These applications include a native mobile client for all major mobile platforms.

The respondent from HRT admits that developing and updating similar offering with in-house resources would likely have been an impossible task for them. As a result, the partly taxpayer funded municipal corporation has decreased costs.

The resources required for keeping the service running are the maintenance of the API and creating new API versions to keep up with developer requirements. Relationships with the developer community are also very important in order to keep the community active. HRT has organized an application contest to increase the awareness of the API and encourage new developers to create new applications with it.

Another well-known example of commercial open data publisher is the several private data openings of Helsingin Sanomat, the largest subscription newspaper in Finland. However, Helsingin Sanomat was not included in the original list of respondents because they forbid commercial re-use of the data and therefore do not comply with open data guidelines set by Open Definition.

Regardless of the utilized Creative Commons non-commercial license, Helsingin Sanomat data has received a lot of attention from the open data community. For example, in April, 2011 Helsingin Sanomat released data from its web-based service helping voters pick a favorable candidate in the parliamentary elections. The data consisted of answers given by the candidates for the service. Just a week after the data was released, Helsingin Sanomat had already received 15 applications and visualizations using the data. Because of the popularity of the data release, Helsingin Sanomat has continued to release candidate answers in a raw data format in later elections.

### **4.5. Support services and consultation**

Support services and consultation companies assist other value network profiles in their business. Thus, they are not directly involved in the production and refinement of open data artifacts. Instead, they assist other companies that operate in some of the other value network participant profiles. This distinction is highlighted with arrows and dotted lined box in Figure 3.

Support services and consultation operates with typical business-to-business revenue models, such as

project work or service-based pricing. They use a combination of in-house resources and partners to solve the customer problems efficiently.

Flo Apps offers tailored visualizations and web-services for its customers. Parts of these were created using open data, but other data sources were also used. The CEO of Flo Apps told that approximately 30 % of their income comes from open data related projects. The figure is relatively low considering that Flo Apps is one of the first companies in Finland to commercially exploit the possibilities of open data. Flo Apps has actively participated in Apps 4 Finland

competition with six submissions over three years. The respondent said that success in these application contests has helped them position Flo Apps as one of the most known open data consultancy and subcontracting partners in Finland.

In Table 3 below we have highlighted the different business models we identified. In what follows, we analyze this table, draw conclusions and discuss avenues for further research.

**Table 3 Summary of the business models of the five value network profiles**

Value network profile (number of companies)	Offering	Revenue model	Resources	Relationships
Extract & transform (1)	Find and convert raw open data into a format allowing further analysis and processing.	No revenue model, the organization operates pro-bono	Open data and the volunteer developers working on the toolkit	open source community, open data community, and public administration and private data publishers
Data analyzer (3)	Create visualizations or algorithm-based analysis to generate new knowledge from the data.	Project work, product based transaction pricing, and modular ecosystem	open data, extracted & transformed data, and private or commercial data	Open data community, public administration, and global information providers
User experience provider (7)	Create interactive user interfaces with help of open data sources.	Advertisements, one-time fee, subscription, donations, licensing and freemium	Open data, extracted & transformed data, scraped data, private or commercial data	Open data community and subcontractors
Commercial open data publisher (1)	Publish data to be freely used by the community	Cost saving through crowd-sourcing of new user interfaces	Maintenance and updating of the API towards the published data	Open data developer community to increase awareness and usage of the data
Support service and consultation (2)	Supports other companies in the open data value network.	Project work and service-based pricing	In-house programming, consulting and subcontracting	Open data community and small open data companies as innovation partners for large technology companies

## 5. Discussion and conclusions

To answer the first research question about actors in the open data value network we identified five value network participant profiles and described their business models. A summary of the value network profiles along with the number of companies within each profile and a list of revenue models within each profile was shown above in Table 3.

From the table we can see that the companies were distributed very unevenly among the five value network profiles. This finding and discussion below answer the second question about what business

models are employed by the early open data entrepreneurs. Only one entity is focusing solely on extraction and transformation, whereas half of the studied companies were categorized under user experience provision and three under data analysis. This seems intuitive as these roles have a smaller entry barrier, but also because they have more versatile and direct revenue model possibilities.

### 5.1. Impacts for practitioners

For businesses there are three main opportunities. The first is to utilize raw open data as a source

material for providing new services. The second is to utilize community processed open data as a source in new services. The third is to actively publish open data and use crowds to analyze and rake the data.

The first opportunity is to use an open data source as raw data in new services or analyses. Adopting raw data requires that the extraction and transformation of the raw data be done in-house. This was quite popular among the data analysers, partly because their data needs were so specific that no readily chewed data was available, and partly because so little external extract and transform activity has been taking place. This is rather labor-intensive and prone to discontinuities for example if the data source format changes.

The second opportunity is to use third-party created analysis as a source in new services or analyses. These third parties can be hobbyist analysers in the open data community, or commercial companies focusing on a specific part in the chain of open data artifact assemblage. Third-party analysis reduces the risk and effort needed in to conduct the extraction, transformation and data analysis in-house. At the same token it is restricted to the set of analyses and processed data available on the “data market”. In addition, the right data needs to be found.

The third opportunity is to publish private data in order to create cost savings by analyzing, visualizing or creating user interfaces with the help of the open data community. Aitamurto and Lewis have studied open APIs from four big news organizations in their [17] article, and found that open APIs accelerate R&D processes and generate new means of commercializing content, especially in niche segments otherwise difficult to serve. Publishing private data to the public pursues exactly the same goals. However, caution and discretion should be exercised when opening up private corporate data. Exposing critical data assets, which constitute the company’s core competence, might jeopardize the entire business. Therefore, not all private data should be opened, but in the right cases opening up will bring considerable business benefits.

It can be speculated that for the open data industry to emerge there needs to be more players occupying the roles of “extract and transform” and “open data publisher” because these are needed by the user experience providers to create new services. These players did not have clear revenue models and this will make other entrants wary of entering the field. It can be speculated that when the field matures, these are the key back end service providers. It seems clear that at least in the start these back end providers need some, perhaps public, funding to ensure the availability of the data.

## 5.2. Limitations

There are obvious limitations for this kind of descriptive research, as we have been studying early entrants whose business models are not fully fleshed out yet. At the same time we believe that it is important to understand how these early examples work to help the next wave of hopefuls enter the field.

We speculate that similar approaches are likely in other national contexts, even though the legal constraints might vary across different countries – as well as access to financing. The Finnish government has good quality datasets and has been moderately fast in releasing them. Thus service companies are in moderately good position also globally. Most of the services so far have targeted local users – a constraint mainly posed by the content of the datasets.

## 5.3. Avenues for future research

This study calls for more research on the business impacts of open data. In the future it is interesting to see how the industry forms and which business and revenue models turn out to be sustainable. The legal and technical challenges need to be overcome, but so do the ones related to developing innovative services on opened datasets. An international comparison of services might be one way to increase the reach of services or alternatively copy parts of the concepts in different countries and using their local datasets. Open data research could also benefit from building more links related research areas, such as big data and analytics as well as to research on open source software.

New research questions that are derived from our findings include:

- What kind of open data services are able to attract user-base and become popular?
- What kinds of datasets are able to attract developer interest?
- What are the different legal restrictions posed in different countries for opening data or service design and how to overcome these issues?
- How to mitigate conflicts in the value chain, especially between the data publisher and service provider?
- How to evaluate the risks involved with data opening?
- How to find ways to address the privacy concerns of individuals whose data is collected and opened?

## 6. References

- [1] G. Kuk and T. Davies, "The Roles of Agency and Artifacts in Assembling Open Data Complementarities," presented at the 32nd International Conference on Information Systems, Shanghai, China, 2011.
- [2] J. Lindman, M. Rossi, and V. K. Tuunainen, "Open Data Services: Research Agenda," in *46th Hawaii International Conference on System Sciences (HICSS)*, 2013, pp. 1239-1246.
- [3] T. Berners-Lee, "Design issues: Linked data," ed, 2006.
- [4] A. Osterwalder, "The business model ontology: A proposition in a design science approach," *Institut d'Informatique et Organisation. Lausanne, Switzerland, University of Lausanne, Ecole des Hautes Etudes Commerciales HEC*, vol. 173, 2004.
- [5] R. Rajala, "Determinants of business model performance in software firms," Helsinki School of Economics, 2009.
- [6] S. L. Vargo, R. F. Lusch, and M. A. Akaka, "Advancing Service Science with Service-Dominant Logic," in *Handbook of service science*, ed: Springer, 2010, pp. 133-156.
- [7] (2013). *Open Definition*. Available: <http://opendefinition.org/okd/>
- [8] W. E. Warren, C. Abercrombie, and R. L. Berl, "Adoption of a service innovation: a case study with managerial implications," *Journal of Services Marketing*, vol. 3, pp. 21-33, 1989.
- [9] L. B. Methlie and P. E. Pedersen, "Business model choices for value creation of mobile services," *info*, vol. 9, pp. 70-85, 2007.
- [10] S. Alter, "Viewing systems as services: a fresh approach in the IS field," *Communications of the Association for Information Systems*, vol. 26, 2010.
- [11] A. Latif, A. U. Saeed, P. Hoefler, A. Stocker, and C. Wagner, "The linked data value chain: A lightweight model for business engineers," *Proceedings of I-Semantics*, pp. 568-577, 2009.
- [12] Y. Tammisto and J. Lindman, "Open Data Business Models," in *The 34th Information Systems Seminar in Scandinavia, Turku, Finland*, 2011.
- [13] A. Poikola, P. Kola, and K. A. Hintikka, "Julkinen data," 2010.
- [14] H. Chesbrough and R. S. Rosenbloom, "The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies," *Industrial and corporate change*, vol. 11, pp. 529-555, 2002.
- [15] R. K. Yin, *Case study research: Design and methods* vol. 5: SAGE Publications, Incorporated, 2008.
- [16] G. Walsham, *Interpreting information systems in organizations*: John Wiley & Sons, Inc., 1993.
- [17] T. Aitamurto and S. C. Lewis, "Open APIs and News Organizations: A Study of Open Innovation in Online Journalism," in *International Symposium on Online Journalism, Austin, TX, April*, 2011.