D3.3.2.3 Detailed sample profiles of Finnish open data companies

Editor:	Tomi Kinnari
Author(s):	Tomi Kinnari, Juho Lindman, Matti Rossi
Confidentiality:	Public
Date and status:	31.11.2012 - Version 1.0

Participants	Name	Organisation
	Tomi Kinnari	Aalto School of Economics
	Matti Rossi	Aalto School of Economics
	Juho Lindman	Hanken School of Economics

Johdon yhteenveto

Julkaistun avoimen datan määrän jatkuvasti kasvaessa siirrymme liikkeenjohdollisista kokeiluista kestävään arvontuotantoon. Median toimialalla tämä tarkoittaa uusien toimijoiden ja liiketoimintamallien läpimurtoa. Nämä uudet mallit sisältävät esimerkiksi perinteisen mediamateriaalin lisäksi elävää dataa, kehittyneitä visualisaatioita tai paikkatietoon perustuvia palveluita.

Tämä tutkimus jatkaa aikaisempaa analyysiamme avoimen datan kehityksestä Suomessa. Keskiössä on kysymys siitä miten uudet toimijat tuottavat arvoa: mitkä ovat ansaintalogiikat, suhteet toiseen yrityksiin ja tarjoomat? Siinä missä aikaisemmin luokittelimme liiketoimintamalleja, analysoimme nyt yritysten tunnistettuja toimintaprofiileja.

Tässä raportissa kuvaamme mediayhtiöiden arvoverkon ja tunnistamme kolme mahdollisuutta uuteen liiketoimintaan ekosysteemissä. Nämä ovat: 1) raakadatan käyttö datajournalismissa, 2) kolmannen osapuolen analyysien käyttö artikkelimateriaalina, 3) oman datan julkaisu tehokkuuden, kustannussäästöjen ja läpinäkyvyyden parantamiseksi.

Raportti laajentaa Next Median dokumentissa *New Value Network Profiles for Open Data* aloitettua työtämme sekä Tomi Kinnarin samasta aiheesta tehtyä pro gradu-tutkielmaa. Materiaalina analyysissa ovat keväällä 2012 kerätyt haastattelut ja muu materiaali 15 vastaajayrityksessä. Raportti sisältää kaikkien yritysten yksityiskohtaiset profiilit.

Executive Summary

Open data continues to open novel areas for new applications and service development. Now, as many governmental datasets have become available, we are moving from exploration phase towards business value. This development in the media sector introduces new actors and business models, for example, in the forms such as live data streams, advanced visualizations and context dependent service.

Based on continued analyses of the Finnish software landscape, we analyze how these new actors are able to produce value: what are their sources of revenue, relationships and offerings. We previously identified the business models of the players, and we now research the detailed profile descriptions.

In this report, we place the media companies in the value network and derive three avenues of opportunity within business data ecosystem. These are to (1) utilize raw data as a source in data journalism, (2) utilize third party analyses as a source for article ideas and new content, and to (3) publish media's own data as open data in order to gain cost savings and transparency.

This report extends the analysis conducted in the Next Media document (2012) *New Value Network Profiles for Open Data* and Tomi Kinnari's Master's Thesis written about the same topic. The research reports analyses utilizing the research corpus collected during spring 2012. Cases are based on fieldwork in 15 companies. They describe the detailed profiles of all the respondent organizations.

Table of Contents

Table of Tables

Table of Figures

List of Acronyms and Abbreviations

A4F	Apps 4 Finland (pronounced "apps for Finland")
API	Application programming interface
HTTP	Hypertext Transfer Protocol
M2M	Machine-to-machine
RDF	Resource Description Framework
URI	Uniform Resource Identifier

1 Introduction

1.1 Purpose of this study

This paper continues the work published in Tekes Next Media report *New Value Network Profiles for Open Data* by Kinnari, Lindman and Rossi (2012). In the report, Kinnari et al. studied 15 Finnish open data companies and identified their business models and the value network in which they operate.

This document analyses the case companies in more detail and presents descriptions of all the cases.

Comparison of the two documents:

New Value Network Profiles for Open Data:

- Open data business models in research

- Interview process description

- Preliminary findings from the interviews, including 8 business models

Detailed sample profiles (this document);

- Re-cap of the research process of the previous document

- More detailed descriptions of the sample profiles used to make the analysis

- More detailed analysis of the open data value network

- Conclusions for the media companies

1.2 Open data

The general definition of open data is data which is accessible as a whole, free-of-charge or at most with a reasonable reproduction costs, redistributable, reusable, in a data format which doesn't cause technological obstacles, and without discrimination against persons or groups nor against any particular fields of endeavour (Open Definition).

However, many of the interviewed companies operated with commercial or partly commercial data, sometimes even without an explicit legal permission of the data owner. Sometimes the usage of the data had ultimately led to a situation where the data owner changed its position on the data re-usage, and was more willing to open its data sets. In order to include these cases, Kinnari et al. (2012) applied a broader definition for open data:

Data, which is accessible through Internet in a machine readable format. It doesn't necessary have to be completely free of charge or free or licenses, but it should allow experimenting with the data, and even running a small-scale-business without restrictions. Technically the data can be in a linked- or in any other machine readable format.

1.3 Business model elements

This report utilizes a business model definition proposed by Rajala (2009) to analyze the open data companies' operations. Rajala (ibid.) defines business model as a:

"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".

These five interlinked elements are defined by Rajala (ibid.) as:

- **Offering** is a value proposition that a software firm offers its customers and other stakeholders, and with which it positions itself in the market.

- **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.

- **Relationships** are the means to access external resources and capabilities.

- **Revenue model** includes the revenue sources, pricing policy, cost structure, and revenue velocity. It is the firm's means to capture value out of its offerings.

- **Management mind-set** distinguishes business model as something that stems from the values, emotions, and attitudes of management; instead of cognitive, rational thinking and planning.

2 Value network profiles

2.1 Prior research

The research done by Kinnari et al. (2012) was based on 15 case companies and 14 interviews. The case companies were selected from Apps 4 Finland contest submissions, focusing on companies that had continued the development and had business activity. Altogether 11 companies were found, out of which seven responded to an interview request. (Kinnari & al., 2012)

In addition to these seven companies, a snowball sampling technique was employed to make sure the research is interviewing the right people. Furthermore, Ville Meloni, an open data expert and one of the organizers of the Apss 4 Finland contest from ForumVirium, was interviewed and asked for guidance in selecting the right interview subjects. These two additional sources revealed eight more case companies, resulting altogether 15 case companies. (Kinnari & al., 2012)

Most of the case companies were examined with face-to-face open-ended interviews, but some were interviewed over Skype or e-mail. One case company (Helsingin Sanomat) was examined only with secondary evidence found from internet. (Kinnari & al., 2012)

New Value Network Profiles for Open Data report by (ibid.) grouped similar case companies together based on their offering, thus creating value network profiles of open data. Therefore, for example, all the seven user experience providers were grouped under one profile, even though they might have different revenue models. Although offering was decided to be the pivotal factor for categorization of the companies, it was not straightforward as many companies had so versatile business practises that they occupied several positions. These borderline cases were decided based on their primary value adding functionality. Extract & transform profile ended up having only one company representing it, but all the other profiles have several ones. (Kinnari & al., 2012)

The research by Kinnari et al. (2012) found eight business models within the open data sphere:

1. Crowd-sourced client development

2. Create valuable user experience

3. Create visualizations

4. Algorithm-based analysing

5. Crowd-sourced analysing

6. Extract, transform, and sell

7. Consultation and project work

8. Machine-to-machine communication

These eight models were presented in four distinct value network profiles (see the list and *Figure 1* below):

1. Extract and transform

2. Data analysers

3. User experience providers

4. Commercial data openers
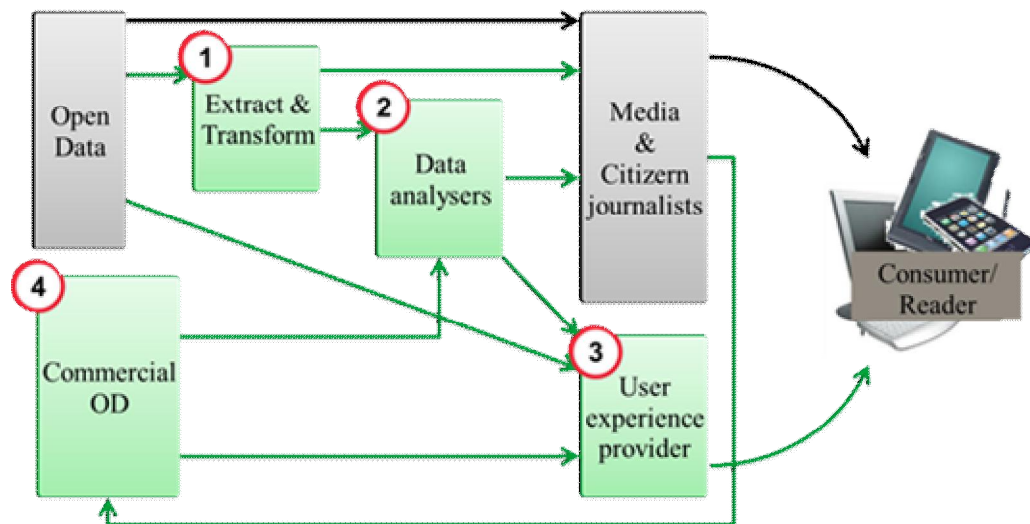
Figure 1 Open data value network presented in New Value Network Profiles for Open Data (Kinnari et al., 2012)

2.2 The value network utilized in this report

In a Master's Thesis *Open data business models – Finnish case study* (Kinnari, 2012), the open data value network was arranged slightly differently, adding a fifth network element to the framework. This "support services and consultancy" profile assists other companies in processing and refining the data, but does not directly locate in the data processing path. Support services and consultation will be analysed further in *Chapter 7* of this document. The value network, as understood in this document, is presented in *Figure 2* below.
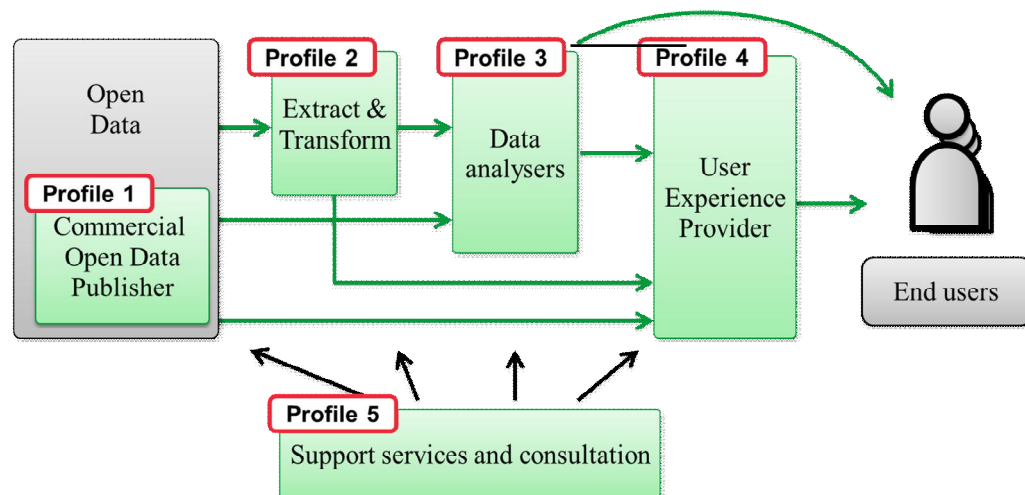
Figure 2 Open data value network utilized in this document

This value network includes five case profiles. Companies within each profile have similar offering, but differ in other aspects of business model. For example,

two companies might offer a mobile client (user experience), but have different revenue models. These five value network roles, and the case companies from which they are derived from, are described in the following five Chapters. Each chapter includes an overall description of the value network profile, and detailed case company descriptions and analysis.

3 Data analysers

Data analysis is an obvious part of open data value network. The interviews revealed multiple types of data analysers. Some were analysing the data to create new visualisations, others were cross-analysing different data sources with advanced algorithms in order to provide valuable knowledge. Some analysers did their job only to serve the common good, while others had strong business model. *Figure 3* below summarizes this profile; it stands between the open data providers and end users. The case companies operating within this profile are: Hahmota Oy, Asiakastieto, and Cloud'n'Sci.
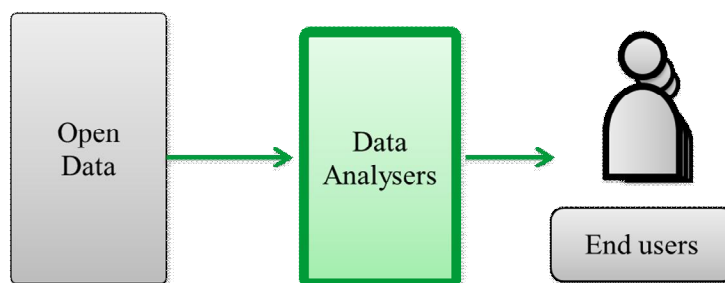
Figure 3 Data analysers

3.1 Data visualizers

Visualization is a powerful way to communicate key points of the data for general crowd. Likewise, for a typical end-user the raw data is basically worthless without an appealing interpretation of it. Therefore, it comes as no surprise that visualizations are very popular within the open data community; there are lots of different visualizations all over the web done by hobbyist visualizers.

3.1.1 Case Hahmota Oy Tax-tree

3.1.1.1 Description

Offering of Hahmota Oy is a visualization of financial data in a tree-like shape. The CEO of Hahmota Oy explains that their visualization principally offers a new metaphor for the basis of conversation. The tree metaphor has allowed their clients to invent a new terminology in their discussion; they talk from leaves, branches, roots etc. In a way, the visualization is like Google Maps for financial data. Their customers could be private companies, governments, governmental bureaus and public utilities alike. The CEO envisions that if they could create a

new visualization method to be used with financial reports, it would open a completely new market for the product. (CEO, Hahmota Oy).

Resources utilized are the financial data to be analysed, a proprietary analysing engine generating the visualization, and personal work to pick the best figures to be analysed. The CEO states that typically their clients offer the necessary data to be analysed, thus Hahmota Oy doesn't need to extract or transform any data. However, sometimes they have made example visualizations for prospect customers as a starting point of the sales process, and this has required some data extraction as well.

Relationships are formed with various companies who wish to use the Finance Tree visualisation engine in conjunction with their own product or service, for example a finance analytic company. The respondent from Hahmota ponders also, that an international partner would increase their visibility.

Revenue of Hahmota Oy comes from one-time project fees charged directly from their customers. Pricing of the project is composed of two parts: Consultation hours and creation of the actual visualization. Their CEO states that at the moment 70 % of their work time goes to consultation regarding how to present the data and 30 % to the final visualization, but in the future the consultation time should become smaller as Hahmota gains a larger portfolio of previous works which help new customers to decide how to present their own figures.

Management mind-set was entrepreneurial; the newly found company was searching for growth and looking forward to make real business with their concept.

3.1.1.2   Analysis: Create visualizations and monetize by selling project work

Hahmota collects data from a client, and uses its unique visualization engine to create a tree-like visualization out of the data. The data doesn't have to be open, but Hahmota has become famous for the concept of visualizing open data from municipal authorities named tax-tree. The tax-tree visualization, however, was based on imaginary municipal financial data, and worked as an advertisement for the company. The actual revenue comes from tailored projects for other companies and government bureaus. Their project-based business model requires lengthy sales work, private contracts, committing to timetables etc. Thus, Hahmota Oy business model resembles software subcontractor's business model, only with a special asset of visualizing vast amounts of data.

3.2   Algorithm based analysis

Another way to analyse data is to utilize advanced algorithms and scrutinize the raw data on numerical level. Our study found two cases companies operating on algorithm level; Asiakastieto and Cloud'n'Sci. Both had strong business intentions, but very different business models. Whereas Asiakastieto leaned on transaction based pricing of information products, Cloud'n'Sci offered a platform where third-party algorithm providers could connect to business customers. Yet,

in both cases their raison d`être was to utilize algorithms to refine raw data into something which has value to their customers.

3.2.1 Case Asiakastieto

3.2.1.1 Description

Offering of Asiakastieto is information from Finnish companies, private citizens, and properties. The respondent from Asiakastieto states that they collect precise data on individual level, with the accuracy of personal identity number, business ID or real estate number. By cross-analysing this data with advanced algorithms, Asiakastieto can give a risk rating for each individual and company in Finland. The respondent said that 95 % of their business is based on this individual level knowledge. Since this knowledge is used in important credit decisions, there is no room for mistakes in the data.

Resources of the company are raw data sources, and both employees and algorithms which make the data analysis. Asiakastieto has been extracting data from various public data sources for almost 100 years. Nowadays, the extraction is usually done over digital interfaces with automatized processes, but some data sources still require manually scanning paper documents into a digital format. Typical data sources include, but are not limited to, Finnish Business Information System (YTJ), Trade Register (Kaupparekisteri), National Board of Patents and Registration of Finland (PRH), and Statistics Finland (Tilastokeskus). In addition to public data, they also collect unique data directly from companies with questionnaires and financial statements. Thus, not all of their data sources are open, as open data should be accessible over the internet, but majority of their data sources are publicly available nevertheless.

Relationships include partnerships with EU-wide and global information providers for information exchange, partnerships with Finnish companies for balance sheet exchange, and collaboration with Finnish authorities in various work groups relating to legislative preparation.

Revenue comes from selling information products to customers with transaction-based pricing. The company generates 40 million euros annual revenue from 25000 customers. The majority of the revenue comes from small amount of big customers, but the Finnish entrepreneurial scene as whole is still well represented as a customer base.

Asiakastieto's products are priced in relation to the benefit that the client can achieve with the information. For example, the efficacy of a credit risk analysis can be tested with historical data, and thus the potential benefit can be proven for the customer.

3.2.1.2   Analysis: Algorithm-based analysing

Asiakastieto is perhaps the leading player to monetize public data in Finland. The company is founded in 1906, and today generates 40 M€annual revenue, mostly by utilizing publicly available data sources. Currently they employ 150 employees, and during the almost 100 years of operation, they have earned reputation of a trustworthy information provider. The respondent from Asiakastieto said that if a credit is applied in Finland, it is likely that at some point the credit request goes through Asiakastieto's information systems.

Asiakastieto analyses the data with mathematical algorithms, similarly as Hahmota Oy, but doesn't necessary provide eye-catching visualizations or user interfaces. Instead, they combine several data sources and refine the data in order to give new knowledge and new valuable insight for its customers. Thus, they provide analysing on algorithm-level. Of course they also provide an easy to use user interface to let their customers access these analyses, but although important, the user interface is not the key part of their offering. The user interface has changed over time, and embodied the leading technologies of each era. Initially postal service and telephone was used as an interface, then text based terminal connections, and nowadays internet and information system integration are utilized (Parpola & Kiljala, 2005).

The biggest difference, however, is Asiakastieto's pricing model, which in essence is product-based transaction pricing. Each bit of information they have in their databases is productized and charged at fixed fee based on how many times it has been requested by customers. Thus, the entire business model is based on "create once, sell many times"-type of information product. This pricing model has proven to be very efficient and profitable for Asiakastieto.

3.2.2   Case Cloud'N'Sci

3.2.2.1   Description

Offering of Cloud'N'Sci Ltd is an algorithm-as-a-service platform where third-party algorithm developers and business world problems are connected. That is, a third-party developer with ingenious algorithm can sign in to the platform and offer his solution to the market. These solutions are then packaged and sold for businesses with various algorithm needs. For the businesses Cloud'n'Sci offers a selection of risk-free algorithm solutions whose worth can be calculated before the investment. This differs remarkably from traditional ground-up algorithm development, which utility is usually unknown before the algorithm is ready.

Resources of the service are the actual platform on top of which the algorithm modules can run, the third-party algorithm developers creating new modules, sources of data, and an algorithm architect who has a responsibility from the entire service towards the business customer. The algorithm architect knows what modules are on offer in the platform and takes responsibility that the system delivers promised results for the business customer. The data source can be any

public or private, as long as there is a module that extracts the data from its source in to the platform.

Relationships are used to market and increase knowledge of the platform. The respondent from Cloud'n'Sci says that they have considered collaboration with Helsinki Region Infoshare-project, because it would result in obvious synergies. However, their focus is first to prove the service concept, and then continue finding new partners.

Revenue comes from a revenue share model between Cloud'n'Sci, algorithm developers, data sources, user interface providers, and the algorithm architect, who is responsible of the whole value chain. The actual split is decided per algorithm solution basis, and it varies depending on the importance of the different players in the solution. Some algorithms might be so central to the solution, that their share of the profit will be proportionally larger.

Management mind-set of the company is very business oriented. The CEO says that the service is born global; there is no reason to limit the service only to Finland. In addition, the fixed costs are kept minimal by making the service as self-served as possible. Thus adding a new algorithm module, creating a new algorithm solution by combining the available modules and data sources, and splitting the revenue of a certain service can all be made by third parties, without interference from the maintenance.

3.2.2.2 Analysis: Algorithm-based analysing

Cloud'n'Sci differs quite a lot from Asiakastieto in their business model. Whereas Asiakastieto sells its own information products with the transaction based pricing, Cloud'n'Sci is providing a marketplace where third-party algorithm providers can sell their services onwards. Cloud'n'Sci has prudently productized their platform, including the revenue share models, but their company is still too young to make a judgement about their business model.

From the open data perspective, the Cloud'n'Sci platform is agnostic to the type of data the algorithms are calculating; it could be open or private. The CEO states that for them open data is just one data source among others, and that if an open data source proves to be vital they are willing to compensate for data provider. In fact, the CEO saw the whole freeness of data as an issue, because an ecosystem where one can freely reap the benefits of the data, which someone else has published, doesn't necessary encourage to publish more data.

4   Extract and transform

In order for the raw data to be analysed, it must be available in a format allowing further processing and handling. Thus, in order to analyse any data, the data needs to be extracted from its original source and transformed in to a meaningful format. This activity is what the "extract and transform" entity does. To be clear, no analysis of the data is done at this stage, only extraction and transformation, respectively.

Open data is typically published in a number of different forms, ranging from Excel files to proprietary formats, which are not necessary compatible with each other. For example, governmental bureaus often publish their data in various formats, and make no effort to standardize it between the bureaus. In addition, a data publisher might alter the data structure over time, thus making inconsistency with the previously released historical data. Also, the data sources have no guarantee on how long the historical data will be available in the first place, thus storing it in a third-party database would solve many problems.
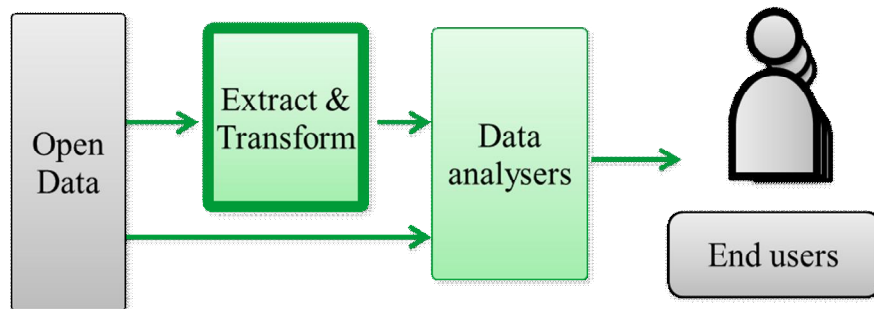
Figure 4 Extract & Transform

The transformation process is even more important if the analysis includes several data sources; the data must be arrayed the same format and in the same scale. This is very cumbersome as there are no guarantees in which scale the data publishers have decided to release their data, and typically, some conversions are necessary.

Usually the data must be also administered in order to ensure its integrity. The data might include double records, missing information or otherwise incorrect information, which needs to be corrected. Part of this work can be done with clever algorithms, but often it requires hours of labour.

This research found extract and transform being part of every open data companies' data processing practises, but it also found one entity specializing in extract and transform activities.

4.1   Extract and transform as integrated part of data analysing

The boundary between data analysers and data extractors and transformers is not always clear. In practise, most of the data analysers, which were introduced in *Chapter 3*, operate also as data extractors and transformers. This is simply because the raw data is rarely available in a meaningful format, and the analysers need to convert the data by themselves. Analysers often use external providers, such as Sorvi, as one data source but the data available through these providers is limited. For example, the respondent from Hahmota Oy said that sometimes the customers bring proprietary data that needs to be analysed. Since this is private data, it needs to be transformed by the analyser himself.

Asiakastieto's business model could be also described as extract, transform and analyse, as extraction and transformation are a big part of their overall process. The respondent from Asiakastieto said they still have to digitalize paper

documents in order to gather enough data to perform analysis. In addition, they also store historical data thus making time-series analysis possible.

4.2 Extract & transform as separate business

Since any data analysis or visualization requires data extraction, it is obvious that there is a need for a separate player as well. Sorvi toolkit by Louhos is probably the best-known example of data extraction and transformation in Finland. However, their company case below reveals that, at the moment, they don't have business intentions. In addition, they have no plans to store historical data.

This leaves open the question whether there would be room for a commercial operator within this value network profile. On the one hand, many interviewees emphasized that raw data as such has no value before it is made valuable by novel analysis and visualization. Based on this, it would be unlikely that someone would be willing to pay for extracted and transformed raw data. On the other hand, several analysers and user experience providers expressed the concern that considerable amount of their time is consumed in data transformation and administration tasks. Most of the public data sources are of surprisingly poor quality, and until the time-consuming transformation and administration process, they are principally useless. Therefore, it would be easy to imagine that refined and trustworthy raw data service would create value for the analysers.

With commercial resources, the service could be improved further. For example, in addition to extraction and transformation, this they could also store the data for further reference. Over time, the stored data would make a time-series analysis possible. In addition, there is no guarantee how long the original data publisher keeps the records available, but on a third-party database the information would be accessible from the entire time span, even if the data publisher would remove some of the older data. Storing the data can be useful also in cases where the data publisher changes its data structure frequently; the interface to the third-party data storage can be kept unaltered.

4.2.1 Case Louhos

4.2.1.1 Description

Offering of Louhos is a comprehensive software library for R-language, named soRvi-toolkit, assisting analysts to extract and analyse open data from several sources. The toolkit offers automatic data fetching-routines supporting several open data sources ranging from municipalities to World Bank and from Finnish postal numbers to OpenStreetMaps (Louhos website, n.d.). In addition to data fetching, the toolkit offers analysis routines to process the data onwards in R. For example, plotting county-level information on top of a Finnish map is made very simple with the toolkit. However, the toolkit doesn't have any central storage for the data; it is a script which extracts the data from the original source each time the toolkit is run.

In addition to the toolkit, Louhos also creates plenty of analysis and visualizations and publishes those in their Louhos-blog along with the example R-source code to replicate the visualization.

Resources required to build and maintain the service are light, principally just the programming knowledge of the founders. The source code has been released in GitHub, and thus anyone can continue the development of the toolkit (soRvi GitHub, n.d.). The founders said they are hoping that other programmers would get interested in the project and start adding new features, datasets, and countries in to the toolkit.

The founders are very active in the open data scene in Finland. These **relations** are used to increase knowledge of the toolkit and to invite other active developers into the community. They are co-operating, among others, with Helsinki Region Infoshare, HS Open, Apps 4 Finland, Open Knowledge Foundation, and Kansan Muisti.

Revenue at the time of interview was zero, and the project ran on the founders will to advance open data. Since there was no revenue, the resources have been kept minimal as well, and the founders are hoping the open source community would help with the development. Although they didn't generate revenue at the time of interview, the founders had ideas about future income from consulting and other supporting tasks.

Since there was no revenue generation, the **mind-set** of the founders is pro-bono. However, they pondered that in the future a combination of commercial and volunteer activity might be the most feasible path onwards.

4.2.1.2   Analysis: Extract and transform

Louhos saves a lot of other companies' time, because everyone doesn't have to go through the laborious work of figuring out the source data formats and parsing out the relevant data into tables in order to execute the actual data analysis. With help of Louhos, the analysers can focus on real analysis, instead of spending time in mundane data extraction and transformation tasks. In addition, since integrating a new data source has been made effortless, the analysers are more likely to employ multiple data sources resulting in a more extensive combinatory analysis. Practically all the interviewed companies expressed a need to extract and transform the data before processing it onwards. Many of the companies did this work in-house before the actual analysis or visualization, but it seems evident that there is a need for a separate entity extracting and transforming as well.

Considering how much added value extract & transform delivers, there is surprisingly little commercial activity in this field. Louhos and their soRvi-toolkit is perhaps the only corresponding entity in Finland. However, at the moment they do not have commercial activity.

Louhos is also the only interviewed entity publishing their source code freely in an Internet repository. This activity is described by Kuk and Davies (2011) as an important part of the overall open data complementarities assemblage, and

according to them, enables further innovation. The respondents from Louhos said they are hoping for more open source developers to join the project, in order to keep adding new data sources nationally and globally.

5   User experience provider

User experience providers are the only entity directly in contact with consumer end-users. The core idea is to utilize open data sources to create a valuable application for the end-users. The interaction can be done either through a mobile or through a web user interface. This research found four revenue models within this value network profile: advertisements, licensing, direct product sales, and donations. As with data analysers, sometimes the user experience providers also need to extract the data from its original source. In addition, they often process and analyse the data as well, so in some cases they perform three types of activities in the value network.

User experience providers are by far the most popular part of the value network. This study found seven companies operating in this role: Duunitori, Reitit, ReittiGPS, Hilmappi, Kansanmuisti, Pikkuparlamentti, and Mitamukaanlennolle.fi. This might be because of diverse revenue possibilities, the offering is easier to conceive and assess because the creator is also a consumer himself, and also the entry barrier is lower because company relationships are not required to sell anything. In addition, the application markets provided by all the major mobile phone operating systems make the sales process easier.

The value network including the user experience provider as well as the two previous profiles is sketched in *Figure 5* below. Most of the user experience providers utilized the raw data directly from its source, thus performing extract & transform and analysis tasks as well. However, some companies, such as Duunitori.fi and KAMU utilized also data generated by other analysers or extractors & transformers.
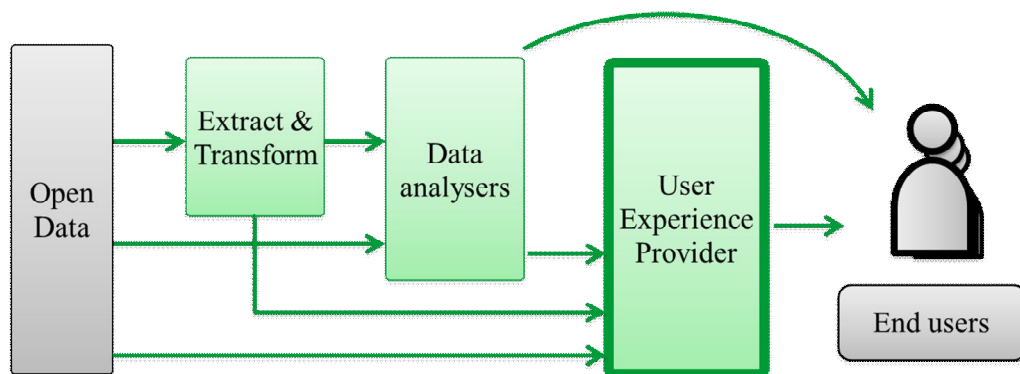

Figure 5 User experience provider

5.1 Monetize with advertising (two-sided markets)

Advertising model is a form of indirect content sales, where the advertisers pay for the content in exchange to have their ads shown to the audience (Afuah, 2004, p. 69). In general, this activity is referred as two-sided markets, where two separable but coupled products are offered for two different markets.

The idea of two-sided market is to under-price one component in order to implement price discrimination in markets with positive network externalities (Parker and Van Alstyne, 2005). In other words, on the right-hand side a company can give away products with prices at or below zero in order to increase the customer base, and then charge the left-hand side market for the expenses. The reasoning is that if the markets are coupled, the network externalities on the right-hand market affect also the left-hand market's demand curve. Therefore, by inducing high demand with under-priced products on right-hand market, the demand curve of the coupled left-hand market moves outward. Outward moving demand curve grows the revenue by making it possible to raise the price of a product while increasing the amount of sold products at the same time. (Parker and Van Alstyne, 2005).

5.1.1 Case Duunitori.fi

5.1.1.1 Description

Duunitori.fi, by Skyhood Oy, scrapes job openings from the Finnish Government's Employment and Economic Development Office (mol.fi) website and plots them on a map. These visualized job openings are then enriched with data from several resources, including Tilastokeskus, Reittiopas, Yritystele, Great Place to Work, Facebook, etc. The result is an interesting mash-up of data fetched from different sources, **offering** jobseekers a hub to find all the relevant information from the employer with the job opening.

Resources to run and maintain the service are kept minimal; the service runs on virtual servers at external partners' facilities. According to the CEO, most work went in to development and building stage of the service, which took tens of person-months altogether.

Relationships are built with labour unions and municipalities, and Duunitori.fi has some pilot cases on the webpages of these partners. The respondent from the company stated that they are better partner for these organizations, because Duunitori.fi has all the job offerings also from the blue-collar segment.

When asked about the **revenue model**, the CEO remarks that since open data is free, and developing services on top of it is reasonably cheap, it leads to situation that users are not ready to pay for them. How these services could then generate revenue, ponders the CEO of Skyhood. He answers to his own question that it could be by selling the users similarly as Facebook, which exploits its user data as raw material and sells it onwards for advertisers. In a way, Duunitori.fi has a

similar subtext; it is bringing job creators and job seekers together. Even if the end users would be reluctant to pay, the businesses are willing to open up their wallet because they want to find new employees. (CEO, Skyhood)

At the moment, the revenue model of Duunitori.fi is based on advertisements, the more visitors the page attracts the more advertisement revenues are possible. In addition, the company is doing custom advertisement campaigns with key partners by offering them increased visibility.

5.1.1.2 Analysis: Create valuable user experience and monetize with advertising (two-sided markets)

The core idea behind the business of Duunitori.fi is to combine different sources of data, both public and commercial, in order to create an eye-catching user experience where the raw data has been enriched and made valuable for the end-user. The revenue model is based on attracting as many users as possible, and then selling the user masses for advertisers and businesses who are hiring employees. This two-sided market business model is typical for internet portals and media companies (Hagiu and Wright, 2011; Rochet and Tirole, 2003, p. 992).

5.1.2 Case Pikkuparlamentti.fi

5.1.2.1 Case description

Pikkuparlamentti.fi **offers** an objective and independent web page, which brings citizens together to discuss about topics of their interest. According to the respondent from Forum Innovations Oy, their website offers more quality, independency, and objectivity than other discussion forums. The idea is that when people search information relating to a certain topic they probably have also some insight from the topic, which they can share with others through the forum.

Resource of the site, in addition to the discussion forum itself, is data gathered from Parliament of Finland. At the time of interview, the data was gathered manually by posting a link to a particular decision proposal and letting users to comment and discuss it. The site has been created by the founders themselves, and thus no subcontracting has been used. (Founder, Forum Innovations)

Revenue model of the site was at the time of interview still at a start-up phase, but they had plans on advertisement-based revenues with banners. They had also plans on selling software for municipalities and government bureaus to help them clarify their decision making process. For example, they were participating in a tender from Ministry of Justice, Finland for an online debating module.

Management mind-set was entrepreneurial oriented, as the respondent stated that they had intentions to build a start-up company already at the idea generation phase. Third place in Apps 4 Finland 2010 competition gave the encouragement and starting capital to find the private limited company.

5.1.2.2 Analysis: Create valuable user experience and monetize with advertising

Pikkuparlamentti.fi has similar business model as Duunitori.fi; combine open data with other data sources in order to create a valuable user experience and monetize it with advertisements. However, the representative from Forum Innovations said that at the moment of interview their user base was not large enough to attract advertisers, but yet he saw advertising as one of the monetizing options. In addition, they have also plans on licensing and selling their platform for other companies, similarly as with mitamukaanlennolle.fi (see the next Chapter) does.

5.2 Monetize with licensing

Licensing is a way to monetize patented or copyrighted inventions. Rather than productizing the invention by themselves, a licenser may decide to license other firms to use the knowledge. The licenser can charge a fixed up-front fee, a per-unit fee for every unit that a licensee sells, or a combination of these (Afuah, 2004).

5.2.1 Case Mitamukaanlennolle.fi

5.2.1.1 Description

Mitamukaanlennolle.fi **offers** a web service where airline passengers can check what items are allowed to take on to the plane either in carry-on luggage or in cargo hold luggage. The passengers are more satisfied when they know beforehand whether certain items, such as medicine, are allowed on board.

The main **resource** of the service is a database consisting of 1600 items and their security information. The database is a combination of open and closed data, which has been gathered from International Air Transport Association Dangerous Goods Regulations (IATA DGR) manuals, International Civil Aviation Organization (ICAO) data, European Union regulations, and their own information based on several years of security training of airport officials. (CEO, Suomen Turvaprojektit Oy)

Revenue model of the mitamukaanlennolle.fi service consists of advertisements and licensing. In addition to advertisements, other airports, European Union, and IATA have been interested in the service, and Suomen Turvaprojektit has already licensed the service to Norwegian and German airports. They are expecting the licensing revenue to increase when they expand to even more countries in the future. (CEO, Suomen Turvaprojektit Oy)

In addition, the CEO reveals that there are several researches, which show that if the security check at the airport goes smoothly the passenger is more likely to spend money in the tax-free shops before departure. Certainly, if the luggage is packed correctly the security check will be easier and less stressful.

In addition, by informing the passengers on how to pack, mitamukaanlennolle.fi is creating a lot of value for the airport security as well. Less security personnel will be needed, because the amount of unnecessary luggage openings and item confiscation will be reduced. In fact, the CEO has calculated that in Helsinki-Vantaa airport alone, there are 9000 cigarette lighters confiscated each month from the cargo luggage. Since it takes between 5 and 10 minutes to open a luggage, it will save between 9000 and 18000 hours of work from the security officials annually.

Relationships are used very effectively. The company has outsourced the advertisement selling to a partner, leaving them more time to focus on the core service. In addition, the service itself was developed by a subcontractor, and it is maintained by another subcontractor. Efficient outsourcing has let Suomen Turvaprojektit Oy to focus on updating the item database.

5.2.1.2 Analysis: Create valuable user experience and monetize with a combination of advertising and licensing

The business model of mitämukaanlennolle.fi is very similar as with Duunitori.fi. Suomen Turvaprojektit Oy is essentially combining raw data from several sources and creating a valuable user experience on top of it. The relationships are used very efficiently, and all non-core elements of the service have been outsourced to external partners. At the moment, the revenue is gathered from advertisements and also licensing. Suomen Turvaprojektit Oy licenses the software and the database for foreign partners, such as airports, and in the future, the licensing model might bring significant income for the company.

5.3 Monetize with direct product sales

Direct product sales is the simplest way of earning revenue. It includes production model, subscription model, and fee-for-service model. Afuah (2004, pp. 68-69) defines the production model as the "automaker model", where each car is paid by the customer directly. In subscription model, the customer pays flat fee for the right to use the product for a period of time. In fee-for-service model, the customer pays only for the service that they use.

5.3.1 Case ReittiGPS and Reitit

5.3.1.1 Description

Offering of both services is a better user experience of Reittiopas service for mobile terminals. Their applications utilize GPS information from the handset to determine current location of end-user, and deliver fast route check-up in to a desired destination. Both of the services were developed for Apple iPhone, but Reitit had in addition an iPad version.

Resources of the services are mainly the public transportation APIs offered by Helsinki and Tampere transportation authorities. In addition, both have integrated Helsinki service map in order to provide a directory of services and points of interest within the city. The client doesn't require a back-end server, because it connects directly to the public transportation APIs, and thus it is easy to maintain.

Relationships are limited to HSL and Apple's App Store, as both perform all activities in-house. Since the offerings of the two companies are very similar, the competition has forced the companies in to a price war. Both respondents said that price is an important decisive factor when a customer is selecting an application from the App Store.

Revenue model for both companies is simply a production model, i.e., charging the end-user for buying the application from Apple's App Store. Initially, however, neither of the founders had plans on making money with the clients, just to create a better Reittiopas service. However, despite the initial expectations of the founders, the applications have turned out to be quite popular in App Store. Apple doesn't give out sales figures of the applications, but representatives from both companies said that they have at least momentarily reached top-10 in Finland. In spite of the success, both developers stated that the revenue is not enough to quit their day jobs.

Both companies offer also a free version of the product with some features removed and with an additional banner for advertisement. The companies have used the free version's banner space only to advertise their premium version, but neither ruled out the possibility that at some point of time in the future the banner space might be sold to outside advertisers as well thus generating additional revenue source.

5.3.1.2 Analysis: Create valuable user experience and monetize with production model

Why exactly are the third-party developers building clients on top of the HSL Reittiopas API? The representatives from both ReittiGPS and Reitit (formerly known as Reitit for iPhone) responded that the initial reason to start developing was their own frustration in the usability of Reittiopas website with mobile phone's browser.

Both companies utilize open data provided public transportation authorities to create a valuable user experience for mobile phones. Whereas the previous case companies have monetized their service through advertisements or licensing, ReittiGPS and Reitit create revenue with sales of the premium version. Both companies have considered advertisement revenues in addition to the production model, but so far haven't pursued them. However, they have launched a free of charge version with stripped functionalities and an advertisement banner pointing towards the premium software. This could be described also as a freemium business model (Andersson, 2009).

5.3.2 Case Hilmappi

5.3.2.1 Description

Gemilo's Hilmappi is a website that **offers** a better user interface to Finnish Government's procurement announcement service, named HILMA, by plotting them on a map and offering tools to manage and tag the announcements. According to a representative from Gemilo Oy, their user interface can save end-users' time remarkably. He said that before the service, their own employers spent 30 minutes daily just to browse the new announcements with the government's user interface. With help of Hilmappi, they can perform the same task in about 5 minutes. These 25 minutes saved every day add up in a substantial figure on annual level.

Hilmappi has been built with Gemilo's own **resources**, and according to the representative from the company, they spent altogether two weeks in developing the service. Since then it has required only some maintenance and administration work. Therefore, the service hasn't required large investment from Gemilo. The main data source of the service is an API to the HILMA database, which is operated by Ministry of Employment and the Economy in Finland. The database is comprehensive, because all public procurements over 30000e must be listed.

Revenue model of Gemilo Oy is to sell Hilmappi with a 50 €annual subscription fee. The service has also a one-month free trial period to attract users. However, the respondent stated that since Hilmappi is not their core service and they don't have much time to develop and market it, in future they might remove the subscription fee altogether and use the service only to generate public relations to the company's other services.

5.3.2.2 Analysis: Create a valuable user experience and monetize with subscription model

Hilmappi creates a valuable user experience on top of open data, just as the previous cases have done, but monetizes it with annual subscription fee. They have a one-month's free trial period after which they require a subscription. Although Gemilo is considering removing the subscription fee and making the service free altogether, it is still a case example of the subscription revenue model.

5.4 Monetize with crowd-funding

In Finland crowd-funding is still in its very early stages. This is largely due to a strict Finnish law about collecting funding from the crowds. The Money Collection Act 31.3.2006/255 dictates, that in order to arrange a money collection activity in which the money is collected by appealing to the public, a money collection permit needs to be acquired. The permit is granted only for non-profit purposes (Finlex 31.3.2006/255), thus completely outlining e.g. Kickstarter-type commercial crowd-funding activities in Finland.

5.4.1 Case Kansanmuisti.fi

5.4.1.1 Description

Kansanmuisti.fi is a journalistic website **offering** citizens an easier way to follow parliament activity with the help of public information sources in a non-partisan way. In their website they state that their mission is to:

"provide citizens with the opportunity to track parliamentary performance in an easily understandable and politically transparent fashion. Kamu collects information about the voting behavior of the MPs, members' statements made at plenary sessions (full-sitting sessions of parliament), as well as members' proposal of initiatives, and election funding." (KAMU Ry background, n.d.)

In their rules the association states that they collect donations, heritages, and grants to fund their activities (KAMU Ry rules, n.d.). However, in a discussion on 25.4.2012 with the Chairman of KAMU Ry, he stated that at the moment most of their income comes from speaker fees, not donations. Thus, they have still a long way ahead of them to reach a truly crowd-funded status, and only time will show if Kansanmuisti.fi will grow to be the first journalistic website funded by the crowds in Finland.

5.4.1.2 Analysis: Create a valuable user experience, and monetize with crowd-funding

Despite the strict Finnish laws, Kansanmuisti.fi is aiming to monetize valuable user experience with direct donations from the crowds. This revenue model requires transparent governance and a strong cause in which the end-users can relate to, so that they feel it is worth supporting.

6 Commercial open data publishers

Commercial open data publishers are especially interesting, because they bring a new horizon in the open data value network – instead of utilizing data from other open data publishers, a company can publish its own resources and achieve concrete business benefits in doing so. The commercial data publishers are portrayed within the open data sphere in *Figure 6* below. By releasing data, they join among other open data publishers, and enlarge the total open data offering available online. It is then up to the open data community to decide how this data will be utilized.
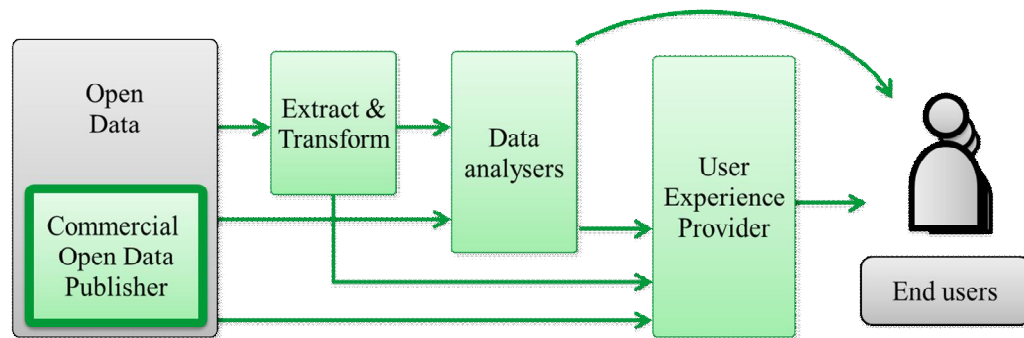
Figure 6 Commercial open data publisher

Opening up private corporate data should be, however, carefully thought and planned. For many companies data and knowledge is their core strategic asset, and releasing it might jeopardize the entire business. Nevertheless, in some cases opening the data can be proven to show remarkable business benefits. The argumentation is similar as with releasing open data in general: the owner of the data might not be its best exploiter. Thus, by opening up data, a third-party extractor & transformer, analyser, or user experience creator could innovate a valuable service that utilizes the data.

Aitamurto and Lewis have studied open APIs from four big news organizations in their (in press) article. They find that open APIs accelerate R&D process and generate new means of commercializing content, especially in niche segments otherwise difficult to serve. They call this process "extended product portfolio", in which the products are built on news organizations' content, but with external developer's user interface. According to Aitamurto and Lewis, this co-creation model generates the most value in open API ecosystem.

6.1 Co-creation under open license

A perfect example of the benefits of co-creation is Reittiopas API, which will be discussed below in more detail. They have gained remarkable savings in client development costs by releasing API to the public transportation schedule and routing data. They have encouraged tens of developers to create mobile clients based on the API. In fact, the clients are done so well that HSL doesn't need to allocate its own resources for mobile application development, and instead can focus on developing the web-interface and maintaining the API.

6.1.1 HSL Reittiopas API

6.1.1.1 Description

Reittiopas is a popular Finnish service, **offering** point-to-point public transport instructions within the Helsinki-region for over 150 000 daily users (HSL website news). Reittiopas is a free service offered by HSL (Helsinki Regional Transport Authority), which runs the commuter traffic service in the greater Helsinki region.

The service is officially available through a web-browser interface with both desktop and mobile instances, but without native mobile applications. According to the respondent from HSL, an application programming interface (API) to the Reittiopas service was built at the same time the service was launched in 2001, but it wasn't opened for the public until 2009. Before 2009, the API was used internally and in some partnership projects. In addition, it was given for third party developers on request, but according to the respondent from HSL, at the time it didn't raise much interest. In 2009, HSL decided to publish the API because at that point the amount of third party requests had risen, and the general awareness of open data possibilities had increased as well.

After opening the API, the developers have been very interested about it. According to HSL respondent, in May 2012 over 650 developers have already registered to get access to the APIs. HSL provides other APIs as well, but Reittiopas is the most popular amongst them. HSL is listing over 30 third-party applications utilizing the API in their webpage (HSL Palvelut muissa kanavissa, n.d.).

The respondent admits that developing and updating similar service offering for this amount of platforms would have been, in practise, an impossible task for HSL to do in-house. The respondent said that in the beginning they didn't think so much of cost savings, but were more interested in seeing what new could be achieved. Transparency of governance was also one of the arguments, that is, since the information was generated with taxpayers' money, taxpayers should have also a free access to it. (Project manager, HSL)

One of the most known applications is ReittiGPS by Essentia Solutions Oy, providing a native iPhone application for the journey planner service. The CEO of Essentia Solutions said the project was started in 2008 to satisfy the founder's personal need of checking the public transportation schedule easier on the road. At that time there wasn't yet a public API released from HSL, so the information had to be scraped from the HSL website. ReittiGPS and BusWatch were among the first applications to show journey planner information in a native mobile client combined with GPS coordinates from the mobile phone's GPS receiver. (CEO, Essentia Solutions Oy; Project manager, HSL)

The popularity and success of ReittiGPS was a strong indicator for HSL, that it might be worthwhile to release the API for the third-party developers. Quickly after HSL released the API, other similar applications started to emerge. Thus, although ReittiGPS sprang up without support from HSL, the official API release lowered the bar and encouraged several developers to create their own version of the mobile journey planner.

These new applications have increased competition and brought innovation to the marketplace. The newcomers forced the incumbent ReittiGPS to implement new features as well. A good example is when Reitit (previously Reitit for iPhone) by Fresh Bits integrated the Helsinki service guide interface in to its application, ReittiGPS had to implement it as well. (CEO, Essentia Solutions; Software developer, Fresh Bits)

The increased competition has even started a price war in the Apple's App Store. In the interview with Essentia Solutions, the CEO said that they had to answer the increased price pressure by dropping ReittiGPS price from 4 €to 3 € The respondent from Fresh Bits said that they purposely challenged the incumbent ReittiGPS by carefully pricing the client at approximately 2.5 €

The respondent from HSL said that after the initial release of the API in 2009, they have continuously improved it in order to better answer the needs of the third-party developers. In 2011, HSL organized their own developer challenge, HSL Mobiilikisa, which invited people to innovate new uses for the API. HSL received 63 submissions in their contest, out of which eight were rewarded (HSL Mobiilikisa).

6.1.1.2   Analysis: Crowd-sourced client development

A company can achieve remarkable savings by, in effect, outsourcing the client development to third-party developers. The core idea is that native mobile application development and updating is very expensive for a company whose core competence is somewhere else. Yet, a well working mobile application generates a lot of value for the company's customers. Thus, the company can either hire IT-professionals to do the development of the application in-house, source the client from a subcontractor, or, as is presented above, publish the necessary data and interfaces in the internet and let the third-party enthusiasts develop the application.

The benefits of crowd-sourced client development are obviously related to cost savings. In fact, since the only cost is publishing and updating the necessary data and supporting the developers, crowd-sourced clients can appear in unthinkable environments and platforms, where traditional outsourcing would not be feasible. However, initially HSL motivation was not cost savings or user interface crowdsourcing, but only to see what new could be achieved and to open the API since it was produced with taxpayer money.

What were the success factors behind crowd-sourcing Reittiopas user interface? Three observations can be made from the case; (1) importance of releasing an API, (2) personal need and motivation of the developers, and (3) facilitation of the developer community with competitions and support services.

The need for a good mobile user interface for HSL Reittiopas has existed a long time before HSL opened an official API for the developers. First application, ReittiGPS, was initially created without an API, only by scraping the necessary content from HSL website. The fact that someone made a client without an official API indicates that the demand for such clients was incredibly high. However, after opening the API, the amount of Reittiopas mobile clients increased dramatically. Thus, although ReittiGPS was available without the API, the API opening was necessary to ignite development in wider scale.

Another important observation is the personal need and motivation of the developers. Both of the interviewed developers had a strong personal need of the

route planner service they had created (see also *Chapter 5.3.1.2*). Therefore, motivating and getting the developers interested in the API is a key element in successful user interface crowd-sourcing.

Final observation is that competitions seem to increase the awareness and interest towards APIs substantially. Some good examples are HSL Mobiilikisa and Apps 4 Finland competitions. HSL Mobiilikisa was designed to encourage developers create more applications and for them to compete with each other. Since the competition received over 60 submissions, it has been a great success.

6.2 Co-creation under restricted license

Commercial companies have a risk of losing profits or customers if they release wrong data. They need to be very cautious on what exactly should be opened and with what licenses. At the same time, they need to inspire the developer community in order to induce activity around the data release. Thus, companies need to select the right data to open, and to put effort in encouraging crowds to utilize the data as well. In a way companies need to balance between what data can be opened in the first place, what would bring the most benefits for the company, and what would inspire the developer community.

6.2.1 Case HS Open

6.2.1.1 Description

HS Open **offers** an event that brings journalists, graphical designers and programmers under a same roof and encourages them to brainstorm and create new purposes for open data. The event is organized by Helsingin Sanomat, and first HS Open was held on 14.3.2011. It has been a very successful activity, producing tens of prototypes utilizing open data in visualizations and innovative user interfaces. It has been organized regularly, and the fifth HS Open was held on 21.5.2012.

HS Open events have encouraged a crowd of people to make data analysing on their spare time for free. These hobbyist analysers may use very advanced statistical methods, such as factor analysis, neuron networks, or self-organized maps, in their work. Some of the analysers work as researchers in universities or other research organizations, and use statistical analysing tools as a part of their daily work. Most of the analyses that were created during the HS Open event were published only in the HS Next blog or in their creators' private website or blog. However, some of the best visualizations inspired articles that were published in the paper version of the newspaper.

One example where these visualizations have been published comes from HS Vaalikone data opening. HS Vaalikone is a proprietary application aiding end-users to select a favourable representative in elections. HS decided no publish its entire Vaalikone data on 6.4.2011, which was a landmark data opening from HS because previously only pre-digested Vaalikone data has been given out. A week

after the data release, HS had received already 15 applications and visualizations utilizing the data (Mäkinen, 12.4.2011). These visualizations motivated at least two newspaper articles, which were published on 12.4.2011 and on 20.4.2011.

Another example is an article published on 24.6.2011 (see *Figure 7*) about electoral funding relations. The data analyser is a bioinformation technology student, who noticed that an algorithm made for the network analysis of genes could be utilized in electoral funding as well. The visualization inspired an article examining the power structure and connections behind electoral funding. (Mäkinen, 10.2.2012)

It is exactly these types of examples, which prove that the owner of data is not always the best interpreter of the data. These examples would not have been imaginable without the work done by the third-party data analysers, the facilitation done by HS Open, and in the Vaalikone case, the data opening by Helsingin Sanomat.
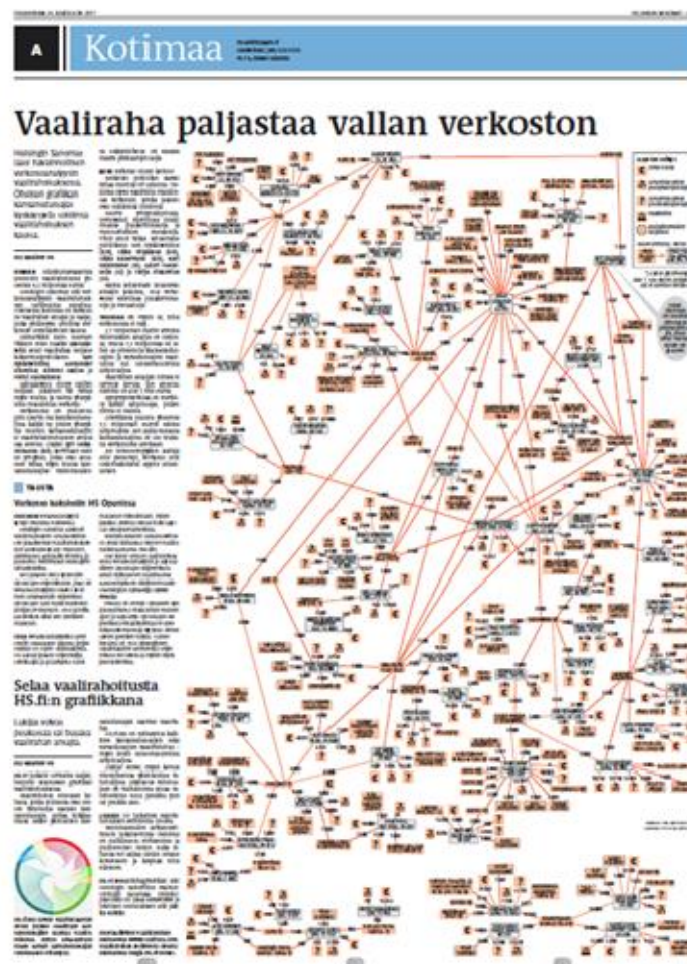

Figure 7 An example newspaper article based on crowd-sourced analyzing. Source: HS Next-blog post (Mäkinen 10.2.2012)

6.2.1.2 Analysis: Crowd-sourced data analysing

HS Vaalikone data opening and HS Open event are both good examples of well-done commercial co-creation. Starting with the HS Open events, in practise HS encourages crowds to analyse data and then use the results in newspaper articles and blog posts. The crowd has been facilitated by organizing an event with a certain theme, by bringing up interesting data to be analysed, and by inviting capable individuals from various backgrounds to come and do the work. Motivation for the crowd is the general interest towards transparency and data analysis, the opportunity to get their work published in a national newspaper, and possibly the complimentary beverages available. This concept has proven to be very successful, resulting in tens of visualizations and data analyses. For HS the cost of the events in terms of money or effort has been minimal.

HS Vaalikone data opening, on the other hand, is a good demonstration of the benefits of opening up corporate data. Endorsed together with a dedicated HS Open event, the data opening got enough publicity to catch the attention of the masses. As a result, the released data found surprisingly innovative usages, and spawned several full-page newspaper articles. This data opening represents a large shift in the management mind-set; a data, which during previous elections has been considered a private asset, was now released to public free of charge, albeit with some re-use limitations.

Helsingin Sanomat chose Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported (CC BY-NC-SA 3.0) license for its Vaalikone data, which restricts commercial re-usage and requires attributing the author (Mäkinen, 28.3.2011; Creative commons website, n.d.). Regardless of the restricted license, the data has been very popular amongst analysts resulting in tens of visualizations. Especially the non-commercial requirement in the license utilized by HS is in contradiction with the open data definition, which states that the data should be free of restrictions and permit unlimited commercial re-use (Open Definition; Poikola et al., 2010). To resolve the contradiction, a less strict open data definition was introduced back in Chapter 1.2. Thus, in this report, HS Vaalikone data is reckoned as open data because it is available free of charge, albeit it has commercial re-use restrictions.

What was the reason behind the shift in management thinking? For the public sector, the reason to open data is typically to advance citizens' participation in democratic decision-making process, increasing government transparency, and a general pressure from the crowds to open up data sources. For corporations, however, similar reasoning does not apply. Quite often, the data is a core competency of the company, which makes the companies understandably cautious when opening the datasets.

In the case of HS Vaalikone, the motivation was probably similar as with HS Open events; to use crowds to analyse the data at hand and give new insight. But this cannot explain the entire act, because surely HS had professional data analysers and reporters creating the first scoops out of the data. It is hard to imagine that HS would trust the entire data analysis solely on third-party hobbyist

analysers' competence and timetables. Something else must have been the decisive factor behind the data release.

This research believes that a two-fold motivation would explain the behaviour behind HS management. The first hand data analysis was still conducted by in-house professionals, but by releasing the raw data they (1) increased readers trust in their data collection and analysis process. At the same time, as people browse through the database and analyse the answers themselves, HS got (2) several "second opinions" in case HS missed something in their initial internal analyses. So, unlike with the HSL case where the aim was to outsource the entire mobile user experience development for third-party developers, HS didn't expect to outsource all the analysis. Instead, they planned to increase readers' trust into the data integrity and also to get a second opinion in to their internal analyses.

7   Support services and consultation

Support services and consultation are the fifth value network profile. Fitzgerald in in his (2006) paper describes that open source software has long established its business model on a combination of volunteer work, free-to-download software, support services, and consultation (Fitzgerald, 2006). Since open data is a close relative to open source software, similar business models can be applied. Support services and consultation is portrayed in the value network (*Figure 8*) as a separate entity, assisting all other profiles (1-4) in their business. Thus, this entity is not directly involved in the value chain from raw data to end user, but instead assists the players in it. Please note that the illustration in *Figure 8* is not entirely accurate, as for example, extract & transform (soRvi toolkit by Louhos) didn't receive any assistance from neither of the companies portrayed in this profile.
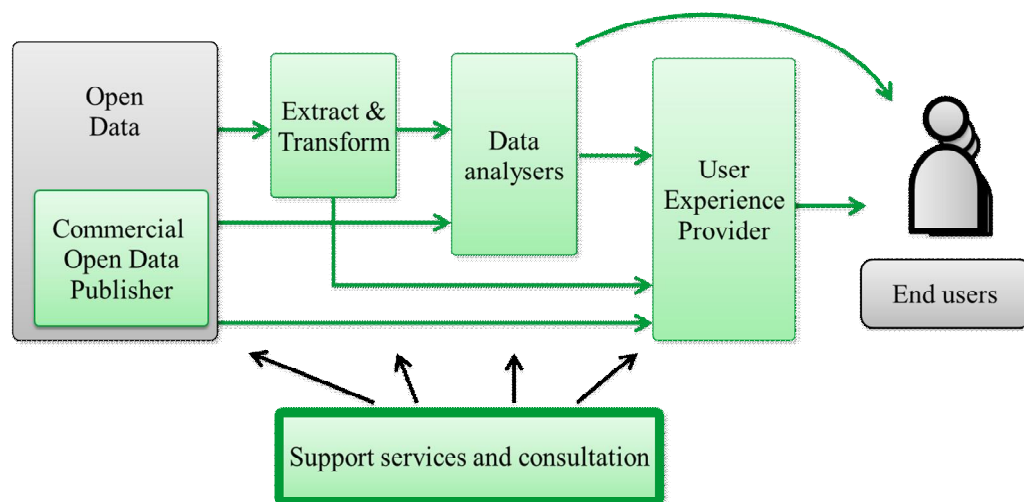

Figure 8 Support services and consultation

This study found two companies belonging in to this category, Flo Apps and Logica. They both have project-based business, where revenue is generated from tailored projects sold directly to corporate customers. They both perform consultation and subcontracting. Logica was more positioned towards machine-to-

machine (M2M) communication and system integrations, whereas Flo Apps was focusing more in visualizations. Flo Apps could be also portrayed in the data analyser box, but since they do not have their own products but instead perform project work for other clients, they are categorized here.

7.1 Case Flo Apps

7.1.1 Description

Flo Apps is a Finnish software company **offering** technical implementations of open data visualisations, thus helping their customers to present their data in a more visual and appealing form. The CEO of Flo Apps stated that the customers usually require tailored solutions, including information design, user interface design, and software implementation.

Resources are used efficiently, as Flo Apps produces only the technical implementation in-house, and outsources most of the graphical and information design to partners. This approach has kept the company lean and cost-effective, and able to engage in a variety of different projects. Open data resources have been used especially in their Apps 4 Finland submissions, but also in several customer projects.

Relationships. Flo Apps has been part of the open data scene from its early stages. They have participated to A4F competitions in order to build reputation for their company. Flo Apps has sent altogether six submissions to the competition, all with very good results. However, the competition rarely brings new customers directly. Instead, it positions the company in the spotlight of open data scene, and thus makes the subsequent sales work easier.

Revenue comes from projects sold directly for the customers, and in year 2010 approximately 30 % of the revenue came from open data related projects. The projects are priced based on estimated work, which holds the risk of the project being more complex than presumed in the calculations. The CEO says that the projects require substantial amount of tailored work, which rises the price tag easily so high that the customers are hesitating to commit. In order to close the deal, both parties often need to make compromises, which reduce profits.

According to the CEO, one solution to make the projects more profitable would be to document and productize the visualization process, thus making it more transparent and efficient. When the process is standardized, the cost can be estimated more accurately, the project can be completed faster, and the customer gets a more professional impression, all of which increase the profit.

Flo Aps has also investigated the possibility of creating its own information product, but the Finnish market is so small that creating and supporting it would be inefficient. The product-based business would require at least European wide distribution, which again would require standardized open data interfaces across European Union countries. At the moment such standardization does not exist,

therefore cross-border software would require extensive localization for each country.

Management mind-set is very entrepreneurial, and they were one of the first Finnish companies to commercially exploit the possibilities of open data. However, their business is not limited to open data. Instead, they are ready to engage in any software project as long as it is realizable with their expertise. This creates an opportunity cost for their open data related activities as well; they need to be at least as profitable as the rest of the business. This background has ensured that all their open data activities have strong business intention, and is probably one of the key factors why at the moment Flo Apps holds one of the biggest open data related turnovers in Finland.

7.1.2 Analysis: Consultation and software projects

The business of Flo Apps is based on tailored project work sold directly to business and government customers, similarly as with Hahmota Oy earlier. Open data is primarily used to gain visibility in Apps 4 Finland and similar application contests. The reputation, which Flo Apps has gained through these competitions, has eventually brought new contacts and new customers. In essence, it is the same marketing strategy, which also Hahmota Oy uses: Utilize visualizations created out of open data as a marketing tool for the company.

What differentiates Flo Apps from many other open data companies is their profit oriented managerial mind-set. Since they have started from general IT projects with no relation to open data, the activities concerning open data need to be profitable as well. This background in general software business has helped Flo Apps to focus only on the profitable open data projects, and discard the ones without potential to generate revenue.

Another observation is the importance of products and productization. Asiakastieto case demonstrated that easy reproducibility and transaction based pricing are key elements in information-based business. Flo Apps has experience with product-based business, does not yet have its own open data products. They have been considering the option to pursue to the product-based business, but so far haven't made the switch because they are concerned that the Finnish market is too small for open data product business to thrive.

If this statement true, does the Finnish market size affect other companies and types of products as well, or is it present only in some segments? That is, can the Flo Apps's experiences of insufficient market size be generalized in to entire open data business? To assess this question, counter-examples can be taken from the product-based case companies examined earlier in this report.

ReittiGPS and Reitit mobile clients, according to their creators (see *Chapter 5.2*), sell quite well, even though they are restricted to Finnish market only. However, this revenue hasn't been enough for the founders to quit their day jobs. Duunitori.fi (see *Chapter 5.1*) is generating revenue with advertisements and Mitamukaanlennolle.fi (see *Chapter 5.2.1*) with advertisements and licensing, but

they refuse to reveal their turnover figures. However, Asiakastieto (see *Chapter 3.2.1*) is probably one of the most profitable public data related companies in Finland, and it operates with product-based strategy. This one example alone is enough to revoke the too small markets statement, and thus it can be said that open data related product based business is conceivable in Finnish markets. However, since only one successful example was found, one should be careful not to generalize the result too much. In addition, Asiakastieto operates partly with public data (which is not digitally available over internet), and thus doesn't fully comply with the open data definition.

7.2 Case Logica

7.2.1 Description

Logica is a technology partner, which **offers** its customers help with all the levels of European Interoperability Framework (EIF); political, legislation, organizational, semantics, and technology. Relating to open data, they have been involved in HSL Reittiopas API project, Paikkatietoikkuna, and several other governmental open data projects. All these projects have had strong emphasis on machine-to-machine communication.

Resources of Logica as a big global ICT-player are vast, and they are easily the largest company interviewed in this report. Logica employs 3200 workers in Finland alone, and globally they are part of CGI Group with over 70000 employees.

Relationships are formed with both big and small players. They have searched small companies from A4F competition to do modules or components in their offering. In addition, they collaborate with Microsoft and Oracle, and utilize for example the Azure platform. They are also involved in the Data 2 Intelligence (D2I) program hosted by Tivit. In the D2I program the aim is put organizations first, and see how services and processes could be built in a new way.

Revenue model is moving towards service-based pricing. For example, transaction pricing using customers' environment as pricing point is getting more popular. So instead of having artificial pricing of CPU hours or license agreement, the pricing can be based on customers' business, which is easier for the customer to understand.

7.2.2 Analysis: Better services with machine-to-machine communication

Machine-to-machine (M2M) communication is related to open data through system integration and information systems' back-end architectures in general. Although M2M communication is not something end-users can directly see or relate to, it has an important role in service design through information logistics. The respondent from Logica emphasized this aspect of open data.

The interviewee stated that M2M communication is especially important when designing user centric services. For example, in government bureaus data is often stored in inaccessible and incompatible information systems. The consequence is, that the end-user needs to re-enter trivial data, such as his name and address, repeatedly in different electronic forms. This can be very frustrating. In addition, if the users' address changes, in some cases, the user needs to re-enter the data to each different system.

The respondent sees M2M communication as one answer to this problem. By opening the data internally within a bureau and between governmental bureaus, such problems could be avoided. Information sharing could bring other benefits as well. For example, by collecting data from several bureaus and utilizing advanced algorithms to analyse it, predictions from citizens' future behaviour and service needs could be made more accurate. If only permitted by privacy policies, predictions could be made even on an individual level. According to the respondent, a Social Services Department (Sosiaalivirasto) worker could, for example, proactively approach a long-term unemployed citizen if they would coordinate their datasets with Employment and Economic Development Office (Työ- ja elinkeinotoimisto). This could prevent social exclusion and bring inclusive service experience for the end-user. (Development director, Logica)

Youngin Yoo, a professor at Temple University Philadelphia, in his (2012) speech at Aalto School of Economics, took an example of Philadelphia's fire department not having access to the infrastructure information relating to water and gas pipelines, electrical cables etc., although this information is electrically available in another city's bureau. This is a good example where M2M communication could solve a problem relating to everyday life.

However, to be consistent with the previous analysis, M2M isn't a business model with revenue model, distribution channels etc., but more like an area of application for open data in general. Within the M2M area, there can be found several individual business models. Probably the most obvious is the one of Logica, which works as a systems integrator helping its customers to succeed in this sphere.

8   Summary of the value network analysis

The studied companies were categorized under five distinctive profiles in the value network analysis. These detailed profiles and their corresponding companies, offerings, resources, relationships, revenue models, and management mind-sets are recapitulated in *Table 1* below.

The companies are distributed very unevenly between the profiles. The extract and transform activities were conducted by all of the interviewed companies as part of their analysis, but only one entity focusing solely on extraction and transformation was found. Since the need for this activity is so high, there might be room for more players collecting and storing data, and offering it for third-party analysers. Furthermore, as the field matures, these players should offer data in standardized formats. Data analysers and user experience providers, on the

other hand, were very common. Half of the studied companies were categorized under user experience provision, and three under data analysis. This might be because these roles have a smaller barrier of entry, but also because they have more versatile revenue model possibilities.

Table 1 Recapitulation of the value network profiles

	Commercial open data publisher	Extract and transform	Data analyser	User experience provider	Support services and consultancy
Companies	• HSL Reittiopas • HS Open	• Louhos	• Hahmota • Cloud'n'Sci • Asiakastieto	• Duunitori.fi • Suomen turvaprojektit • Gemilo • ReittiGPS • Reitit for iPhone • Kansanmuisti • Pikkuparlamentti	• Flo Apps • Logica
Offering	• Data for others to analyse • Facilitate events to encourage activity	• Convert data in to an easier format for further analysis	• Data visualizations • Algorithm based data analysis	• User experience created with help of open data sources	• Consultation and subcontracting to help clients benefit from open data
Resources	• Maintenance of the data	• Data sources • Open source community developing the source code	• Data sources • Algorithms for visualization or numerical analysis	• Data sources • Development and maintenance of the user experience	• In-house work • Subcontracting
Relationships	• Open data scene • Support community of active developers to increase usage of the data	• Open source community • Open data scene • Governmental and private data publishers	• Open data scene • Finnish authorities • Global information providers	• Open data scene • Labour unions • Municipalities • Subcontractors	• Open data scene • Small open data companies as partners • Big technology companies
Revenue model	• Cost savings with crowd sourcing	• Pro-bono (open source)	• Project work • Product based transaction pricing • Modular-ecosystem • Pro-bono • Target segment: business customers	• Advertisement (two-sided market) • Production model • Subscription model • Crowd funded • Licensing • Freemium • Target segment: Consumers	• Project work • Service based pricing • Target segment: business customers
Management mindset	• Entrepreunial	• Pro-bono	• Entrepreunial	• Entrepreunial • Pro-bono	• Entrepreunial

Offering was used as a determining factor when grouping the companies together. Thus, the companies under one value network profile share similar offerings, but might differ in other aspects of their business models.

Resources usage differs quite a lot between the profiles; the commercial data publisher doesn't need data sources because it publishes its own, while extract & transform, data analyser, user experience provider all rely on external data to be analysed. Support services and consultancy focuses on its employee time management and in make or buy decisions much like any other consultancy or manufacturer company does.

Relationships were seen as; (1) a sales channel, (2) a way to increase public relations and knowledge, (3) a way to drive and encourage the developer and analyser community in co-creation, (4) a way to influence government to open up preferred data-sets, and (5) a method in managing big projects by utilizing help and resources from partners. The relationships ranged from independent hacker-like developers to small companies, governments, municipalities, labour unions and publicly funded open data projects.

Revenue models between the five profiles are very different, varying from pro-bono open source model to project work model. The revenue model selection is influenced by several aspects of the business operation, but mostly it correlates with the underlying customer base. For example, the user experience providers work directly with end-users, and hence apply consumer oriented revenue models such as advertising, subscription, or production model. The data analysers, on the other hand, work typically with business customers, thus applying business-to-business revenue models such as project work or transaction based pricing. The choice of the revenue model is affected also by the external market forces such as competitors' actions and the uniqueness of the offering.

Management mind-set affects strongly to the chosen revenue model and ambitions of the company. Some managers were quite happy with their open data product being a secondary income, whilst keeping their day job somewhere else. Others were pursuing start-up strategy with external investments and hiring employees. Louhos was purely pro-bono, having released their toolkit as source code free for anyone to use. Many other companies used free visualizations or applications as a marketing tool for their other products.

9 Conclusions for media companies

The general open data value network presented in *Figure 2* is redrawn from media perspective in *Figure 9* below. Media companies are positioned as user experience providers, because they have a direct connection with end users. The figure has been simplified in order to better visualize media companies' position in the value network:

- First, the support services and consultation has been omitted because they are not directly linked in the value chain.

- Second, the data flow between the value network profiles has been simplified by leaving out some connecting arrows. The general idea and relation between the value network entities is nevertheless kept the same.

The identified value network offers three opportunities for the media companies.

1. Raw open data resources can be used to create data journalistic content for the end users and to increase transparency in the articles. This activity is emphasized with "opportunity 1" box in the figure below, which encompasses the arrows coming from the open data source and from the extract & transform entity.

2. Media companies can utilize data analysis done by third-party analysers and create new articles or content based on those. This activity is emphasized with "opportunity 2" box below, and it encompasses the arrow coming from the data analysers-profile.

3. Media companies could publish their own data to be analysed and refined by the open data community. This activity is emphasized with "opportunity 3" box in the figure. These three opportunities will be described in more detail in the following sections.
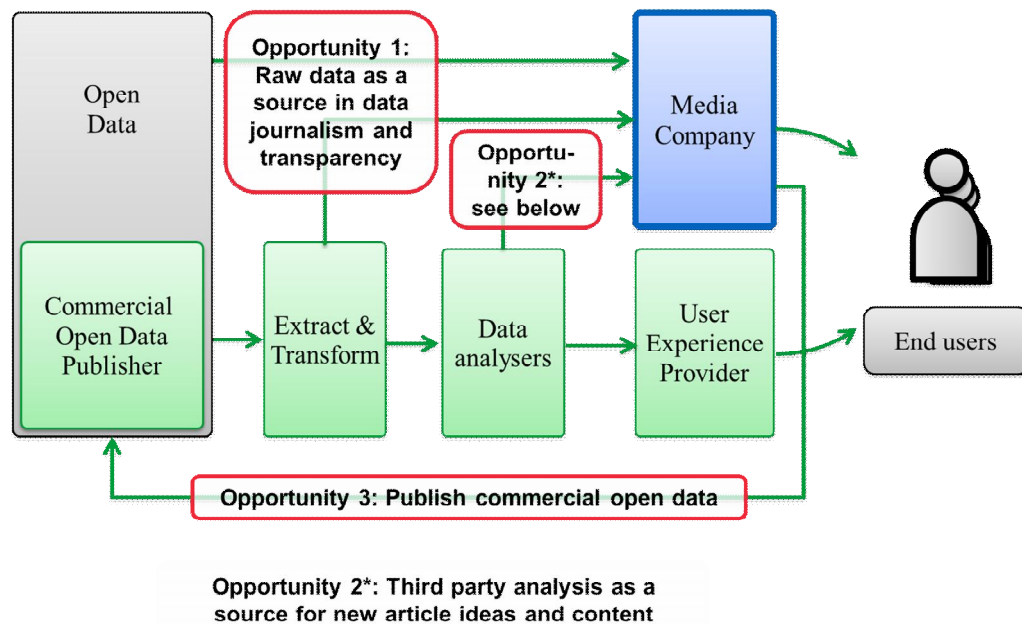

Figure 9 Open data value network from the media perspective

9.1 Opportunity 1: Raw data as a source in data journalism and transparency

One of the most palpable ways for media companies to utilize open data is to use it as a resource in data journalism. Data journalism is a form of journalism, which derives its article ideas from novel data interpretations, such as visualizations or numerical analysis. Typically, professional data journalists working on the media companies' payroll do the analysis. They retrieve the data from a data source, manipulate and analyse the data with various tools and finally make interpretation and write an article based on it. Data journalism can give more depth to the articles when utilized in this way. If open data has been used, and linked, as a source, the readers can also confirm the results and argumentation themselves, thus leveraging the transparency and credibility of the article. Alternatively, if the analysis is done from restricted data, the reader has no way of knowing whether the data journalist interpreted it correctly or not.

The business rationale for data journalistic stories is increased traffic and advertisement. For example, The Guardian has discovered that readers spend

more time with data-journalism stories than regular stories. However, this is not always the case, and often investment in time and resources to data-journalistic stories don't pay off. (Aitamurto, 2011, p. 14).

9.2 Opportunity 2: Third-party created analysis as a source for new content and article ideas

The analysis of the data can also be performed by professional or non-professional third-party freelance analysers instead of the media company's own staff. HS Open case demonstrated how the non-professional analysers can be nudged and encouraged to perform analysis on preferred topics. For example, as we saw in *Chapter 6.2.1*, a bioinformation technology student had utilized an algorithm, made for network analysis of genes, in generating a visualization of electoral funding. These types of multidisciplinary analyses would be very difficult to conduct with in-house personnel.

However, the Finnish freelance data analysers lack a marketplace where they could sell their work onwards to media clients. It is very cumbersome to contact and deal with several newspapers every time a new visualization needs to be sold. The lack of marketplace has led to a situation where most analysers work on their spare time for pro-bono causes, typically publishing their visualizations on their website or blog for free. The situation is very different with e.g. journalistic photographs, where photo agency STT-Lehtikuva connects freelance photographers to media companies. Similar intermediaries or exchanges could be beneficial for the data visualization ecosystem as well.

These circumstances might change in the future as Helsingin Sanomat considers launching a data visualization ecosystem, where they would compensate X euros per publication for the author (Mäkinen, 28.2.2012). Accompanied with inspiring events, such as HS Open, the rewarding ecosystem might increase the interest towards data visualizations in general. However, this would solve the issue only from the perspective of one newspaper; the market would still lack a marketplace connecting analysers to several different media companies.

9.3 Opportunity 3: Publish commercial open data

Third opportunity for media companies is to open up their own commercial data for others to analyse and utilize. Most media companies operate on a two-sided markets business model; on the one side are the readers who consume the articles and possibly pay for the content with subscriptions or pay-per-use, and on the other side are the advertisers who typically contribute majority of the media's revenue (Albarran, 2010). Opening up commercial data shouldn't conflict with either of these revenue sources, unless the company changes their entire business model on which they operate. Taking this into account, could there still be situations where publishing data would be beneficial?

This research found out that the companies have two basic options to publish the data: With or without restrictions for re-use. Only the option without restrictions

represents open data as defined by Open Knowledge Foundation (Open Definition). However, as stated in *Chapter 1.2*, this paper uses more relaxed definition for open data, thus allowing some restrictions in the data re-use as well.

In addition to licensing issues, the technical form of the released data needs to be decided as well. In its simplest form the data can be released, e.g., as a comma separated text file and placed on a web server. A more advanced option would be to offer a custom API, letting developers access the data faster and thus allowing more complex application areas. API is more suitable in large databases or in cases where the data is constantly changing, such as the HSL Live position tracking of busses and trams. However, many interviewees expressed that the technical form of the data is not important as long as the data is simple to access and utilize, and it has clear license allowing its re-usage.

The most difficult choice is to select whether the re-usability of data is somehow limited. The research done within this paper is not sufficient to offer an exhaustive answer on selecting the best option. However, some remarks can be made based on the case companies studied in this report. These findings are summarized in monetary value of data – limitations applied matrix in *Figure 10* below. The next two chapters will discuss both, limited and un-limited, approaches in order to explain *Figure 10* in more detail.
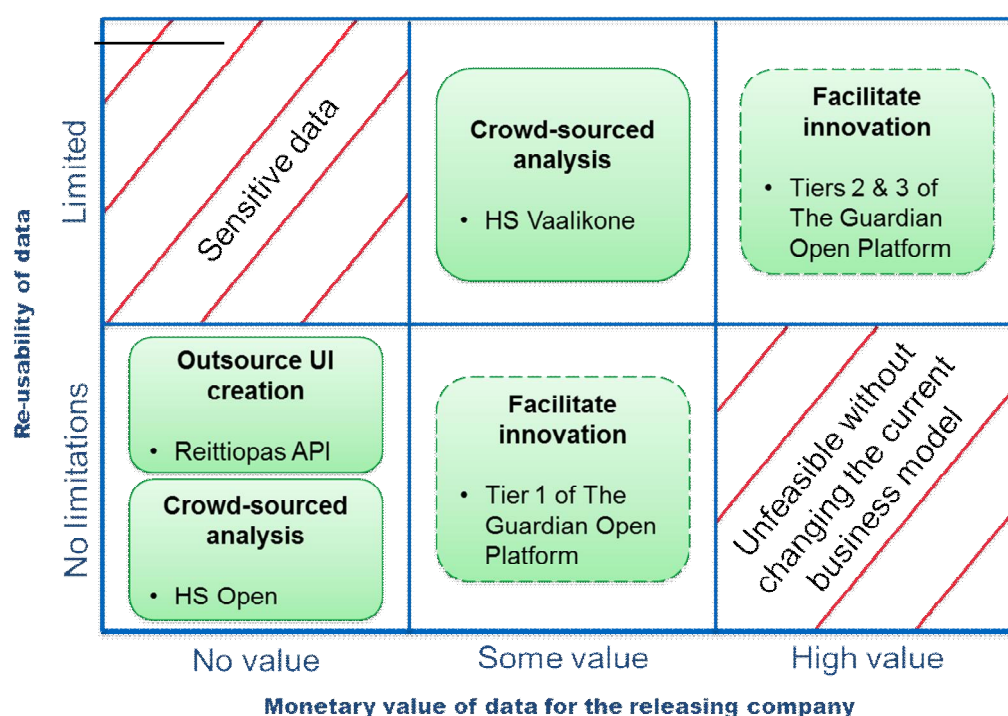

Figure 10 Monetary value of data – limitations applied matrix

9.3.1   Publish data with no limitations for re-use

No limitation for re-use is a practical policy in situations where the data doesn't offer direct monetizing opportunities for the company. Several open data cases

have demonstrated that often data, which seems worthless for its owner might find valuable application areas when released and someone else gets a look at it. In this situation it would be reasonable to release the data with no limitations, since it would be worthless for its owner to begin with. This option is represented by HS Open case in lower-left corner in *Figure 10*. By releasing the data, the publisher would gain general goodwill and possibly some third-party analyses to be used as content in new articles. Furthermore, in case the data has already been used as a source in a data journalistic article, releasing the raw data increases transparency and trust towards the journalistic process and the entire newspaper. While these benefits are vague, it is good to keep in mind that the data was worthless for the company to begin with, and thus, even if a little value is created this approach is vindicated.

Another situation when open license would be reasonable is when there is a need to create a new user interface for some information. This is because creating a new user interface, for example a new mobile client, is very expensive. Scaling the client to multiple platforms and handset models, and updating the client as new operating system versions come along is a cumbersome task. It requires investments, business case calculations, and large enough predicted user base for the project to be reasonable. Clearly not all potential projects fulfill these requirements, and thus many user experiences are left undone. For example, when HSL launched its application contest, it received 61 submissions utilizing Reittiopas API (lower-left corner in *Figure 10*) in new mobile client or web user interface (HSL Mobiilikisa, n.d.). Generalizing this in to media companies' circumstances, releasing data doesn't necessarily create direct revenues, but in the long run it might gravitate masses towards media house's other services and thus be beneficial.

Even if the data has some value, it might be justifiable to release it without restrictions, if it would be too expensive or too risky for the company to monetize it with in-house products. This option is represented by The Guardian Open Platform Tier 1 example in the lower-middle compartment in *Figure 10*. Developers trying, failing, and trying again different user interfaces and mashing up the data with other data sources creates an innovation environment which would be very difficult to replicate with in-house resources. The Guardian example will be described in more detail in the next chapter.

Finally, the lower-right corner represents a situation where the data is very valuable for the company. In this case, publishing data without any restrictions would be unfeasible without changing also the underlying business model. It is, however, possible to operate in this corner of the matrix. For example, typical open source companies, such as Canonical or Arduino, can give out the source code for free, because their revenue model is founded on, e.g., support services.

In general, any data which doesn't directly pose a threat to the existing revenue of the media company and which cannot directly be monetized with in-house products could and should be released as open data. The benefit for releasing the data generates goodwill for the media house, increased transparency, and engenders crowd-sourced analyses and user interface innovation.

9.3.2 Publish data with limited re-use

Another option is to release the data with some re-use limitations. For example, HS Vaalikone (see *Chapter 6.2* for details) made good results by publishing the data with commercial re-use prohibition, while still allowing developers to use the data for other purposes. In addition to several crowd-sourced analyses, releasing the Vaalikone data also increased transparency and trust towards the entire Helsingin Sanomat Vaalikone system. This strategy represents the upper-middle compartment in *Figure 10*. The approach is quite safe for the data publisher, because in order to go commercial and to make profit, the developers need to negotiate with the company. Therefore, the publisher maintains control over the data commercialization, while still leaving enough room for developers to experiment and innovate with the data.

In case the data is highly valuable for the company, the non-commercial limitation alone might not be sufficient to protect the core business. The research corpus examined within the empirical part of this paper didn't cover this situation, and thus cannot provide guidelines or best practises. However, a quick glance to the international market revealed The Guardian's Open Platform, which is a good example of very valuable *"open data"* released with license restrictions (upper-right corner in *Figure 10*). Open Platform lets developers access The Guardian's articles using a three-tier admission system. The first tier lets the developers access the headlines, but not the article body. The second tier, which requires registration and an API key provided by The Guardian, grants the developers a full access to the article body as well. In both cases, the usage of the data costs nothing for the developers, and they can even keep all the profits from their commercial activities. The only catch is that the developer is required to show The Guardian's advertisements in the article body. The third tier offers ad-free access to all The Guardian's content, but requires a contract with the newspaper. (The Guardian Open Platform, n.d.). Since The Guardian case was not in the main corpus of this research, it is drawn with dashed lines in the upper-right corner and the lower-middle compartment of *Figure 10*.

Open data is placed within quotation marks in the above example because the restrictions imposed on the data are in direct contradiction with open data's definition. Since the definition doesn't capitulate well in situations where companies are releasing essential data on which their entire business is relying on, a more relaxed definition of open data was introduced in *Chapter 1.2*. The point is not to practise terminology acrobatics, but to encourage enterprises to become part of the open data sphere while still guarding their intellectual properties. A multi-tier licensing model, such as the one employed by The Guardian, would let third-party developers to experiment and even run small-scale businesses, while still keeping the data in the hands of businesses. Without this exemption, it is hard to see how corporations could release data on which their core business is dependent.

Whether the limitations are prohibiting commercial re-use, requiring authentication from the developers, or something else, it is important to keep the data open enough to maintaining developers' interest towards it. The entire point of opening up data is to make it easily accessible for the developers, so they can

make quick mash-ups, visualizations or other analysis on it. On the other hand, the companies need to protect their intellectual properties in order to sustain their business. Therefore, opening up valuable data is balancing between restrictions and openness to benefit both needs.

Finally, the upper-left corner in *Figure 10* represents data which doesn't necessarily have direct monetizable value for the company, but which still might jeopardize its operations if released. This is typically sensitive data which needs to be kept within the company premises.

10   Summary

Open data companies in Finland operate in a value network depicted in *Figure 2*. The value network includes five roles that companies can take: commercial open data publisher, extract and transform, data analyser, user experience provider, and support services and consultancy. Media companies are positioned in a comparable position as user experience providers in the value network, that is, they have direct contact with end-users (see *Figure 9*). This observation leads to three opportunity avenues for media companies within the value network:

(1) Use raw data as a source in data journalism and transparency. This requires more effort and data analysing skills from the newsroom journalists, but also makes the stories more interesting and prolongs the time readers spend with the article. However, not always the data-journalistic articles become successful, and therefore, there is a risk in putting the extra effort to write a data-journalistic article.

(2) Use third-party analysis as source for new article ideas and content. This is closely related to the previous opportunity, with the exception that the analysis is done by third-party analysers. They can be motivated and guided with hackathon events, such as HS Open, but they can work on their own as well. The third-party analysers can be hobbyist, creating visualizations on their spare-time, or they can be professional freelance data journalists working for several newspapers. Utilizing third-party analyses reduces the risk associated in in-house data-journalistic articles.

(3) Pursue cost savings, transparency, and goodwill by publishing commercial data. Cost savings can be achieved by either crowd-sourcing data analysis or by letting developers innovate new user interfaces based on the data. When publishing commercial data, media companies should pay attention to re-usability restrictions of the data. Restrictions are necessary in situations where the data is essential for the company's core business, but if releasing the data doesn't inflict direct threat to revenue, applying limitations would be futile. Opening media's own resources, for example as was done in HS Vaalikone case, increases transparency and trust towards the newsroom because the users can replicate the analysis done in the paper or make their own ones. These third-party analyses give also a second opinion to the newsrooms internal analyse, and might even give a new perspective to the story.

References

Afuah, A. (2004). *Business Models: A Strategic Management Approach*. McGraw-Hill Higher Education. ISBN 978-0-07-288364-0

Aitamurto, T., & Lewis, S. C. (in press). Open Innovation in Digital Journalism: Examining the Impact of Open APIs at Four News Organizations. *New Media & Society*

Aitamurto, T., Sirkkunen, E., & Lehtonen, P. (2011). *Trends in Data Journalism.* Next Media, a Tivit programme.

Albarran, A. B. (2010). *The media economy.* New York: Routledge

Andersson, C. (2009). *Free: The future of a radical price.* New York: Hyperion.

Creative commons website. (n.d.). Retrieved on September 9, 2012 from http://creativecommons.org/licenses/by-nc-sa/3.0/

Finlex Money Collection Act 31.3.2006/255, http://www.finlex.fi/en/laki/kaannokset/2006/en20060255.pdf

Fitzgerald, B. (2006). The Transformation of Open Source Software. *MIS Quarterly, 30:4, 587-598.*

Hagiu, A., & Wright, J. (2011). Multi-Sided Platforms (Working Paper No. 12-024). Harvard Business School.

HSL Mobiilikisa. (n.d.). Mobiilikisa. Retrieved on May 28, 2012 from http://hslmobiilikisa.blogspot.com/

HSL Palvelut muissa kanavissa . (n.d.). Palvelut muissa kanavissa. Retrieved on May 28, 2012 from http://www.hsl.fi/FI/aikataulutjareitit/avoimentiedonpalvelut/Sivut/default.aspx

HSL website news. (2011). HSL:n Reittiopas 10 vuotta: Keskimaarin 150 000 kävijää päivässä. Retrieved on May 7, 2012 from http://www.hsl.fi/fi/mikaonhsl/uutiset/2011/Sivut/Page_20111107082427.aspx

KAMU Ry rules. (n.d.). Kansan muisti KAMU ry:n säännöt. Retrieved on May 25, 2012 from http://www.kansanmuisti.fi/about/rules/

KAMU Ry background. (n.d.) Mikä Kamu on?. Retrieved on May 25, 2012 from http://www.kansanmuisti.fi/about/background/

Kinnari, T. (2012). Open data business models for media industry – Finnish case study. (Unpublished master's thesis). Aalto University School of Business, Helsinki.

Kinnari, T., Lindman, J., Rossi, M. (2012). *New Value Network Profiles for Open Data*. Next Media, a Tivit programme.

Kuk, G., & Davies, T. (2011). The Roles of Agency and Artifacts in Assembling Open Data Complementarities. In Proceedings of Thirty Second International Conference on Information Systems, Shanghai, China.

Louhos website. (n.d.). Datawiki: ohjeet R-laskentaympäristölle. Retrieved on October 11, 2012 from https://github.com/louhos/sorvi/wiki/Data
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*, Sage Publications Inc. ISBN 0-8039-2431-3

Mäkinen, E. (2012, February 10). HS Open -tapahtumissa luodaan tietojournalismia. Retrieved on May 25, 2012 from http://blogit.hs.fi/hsnext/hs-open-tapahtumissa-luodaan-tietojournalismia

Mäkinen, E. (2012, February 28). Ehdotus datajournalismin bisnesmalliksi: X euroa per kertajulkaisu. Retrieved on May 25, 2012 from http://blogit.hs.fi/hsnext/ehdotus-datajournalismin-bisnesmalliksi-x-euroa-per-kertajulkaisu

Mäkinen, E. (2012, March 28). HS julkaisee vaalikoneensa avoimena tietona ennen vaaleja. Retrieved on October 10, 2012 from http://blogit.hs.fi/hsnext/hs-julkaisee-vaalikoneensa-avoimena-tietona-ennen-vaaleja

Mäkinen, E. (2012, April 12). 15 uusiokäyttöä HS:n vaalikonedatalle – viikossa. Retrieved on May 25, 2012 from http://blogit.hs.fi/hsnext/15-uusiokayttoa-hsn-vaalikonedatalle-viikossa

Open Definition. (n.d.). Defining the Open in Open Data, Open Content and Open Services. Retrieved on May 16, 2012 from http://opendefinition.org/okd/

Parker, G.G., & Van Alstyne, M. W. (2005). Two-sided Network Effects: A Theory of Information Product Design. *Management Science*, Vol. 51, No. 10, October 2005, pp. 1494–1504.

Parpola, A., & Kiljala, J. (2005). *Hyvä vai paha tieto?*. asiakastieto. ISBN 952-9708-13-0

Poikola, A., Kola, P., & Hintikka, K. A. (2010). *Julkinen data - johdatus tietoverkkojen avaamiseen* (pp. 1-96).

Rajala, R. (2009). Determinants of Business Model Performance in Software Firms. Retrieved from http://hsepubl.lib.hkkk.fi/pdf/diss/a357.pdf

Rochet, J.-C., & Tirole, J. (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, Vol. 1 (2003), pp. 990–1029.

soRvi GitHub website. (n.d.). Retrieved on Ocbtober 11, 2012 from
https://github.com/louhos/sorvi

The Guardian Open Platform website. (n.d.). Frequently asked questions.
Retrieved on May 24, 2012 from http://www.guardian.co.uk/open-platform/faq

Yoo, Y. Temple University Philadelphia professor, speech at Aalto University
School of Economics 11.5.2012