Towards open data based business: Survey on usage of open data in digital services

Anne Immonen[1], Marko Palviainen[2], Eila Ovaska[3]

[1]VTT Technical Research Centre of Finland, Oulu, Finland, Anne.Immonen@vtt.fi
[2]VTT Technical Research Centre of Finland, Espoo, Finland, Marko.Palviainen@vtt.fi
[3]VTT Technical Research Centre of Finland, Oulu, Finland, Eila.Ovaska@vtt.fi

Abstract- Now-a-days *open data is in high demand worldwide due to the possibility of creating innovative digital services and applications around open data. This paper discusses the open data from the business perspective; about the status of information usage in companies and motivation, opportunities and obstacles that relate to the open data based business. We carried out 11 interviews with company representatives to receive up-to-date information directly from the industry and to study the status of the information usage in industry to estimate how far the current business is from the open data based business. It seems that open data enables new business opportunities for actors providing data and for actors consuming data but also requires a new kind of business ecosystem that enables a win-win situation for all the actors in the open data ecosystem. The interviewed companies were highly interested in utilising open data in their own business but were afraid of opening their own data. In addition, the open data is often understood to be data that the actors of a public sector provide to the actors of a private sector. Thus, though there is high interest in open data, a lot of work must still be done to enable open data based business. In the end of the paper there are assembled research topics to which the attention should be paid in the future.*

General Terms- *Business development; Digital service development*

Keywords- *Open data; business ecosystem*

1. INTRODUCTION

Recently, the benefits of open data have been discovered widely around the world. Open data refers to the idea that certain data should be freely available for everyone to use and republish as they wish, without restrictions from copyright, patents or other mechanisms of control [1]. In United States and in Great Britain the opening of the administrative data is already a part of government's policies and strategies [2]. Also the European Union (EU) has a clear tendency to increase the utilisation of open data. For example, the INSPIRE directive[1] tends to create an environmental spatial data infrastructure across Europe and enable the public access to this data. EU also has an Open Data Portal[2] that provides access to a range of data produced by the institutions and other bodies of the EU. The tendency in many countries is that the data of the public sector collected along with tax revenues is obligated to be opened. However, there still exist some restrictions and different degree of publicity to this data, such as privacy rules, copyright laws and EU directives. Traditionally, the concept of open data has been linked to public administration; opening the administration data would benefit the business of the utilising companies and thus the national economy, and also provide valuable information to the citizens. Opening data can be beneficial to private companies, too. New ideas, new partners and even new kind of business models related to open data can provide new business opportunities for private companies. Applications and services will be more user-driven, since the users can create new ideas around the open data. The users can enrich their everyday life by digital services, for example, by exploiting services that can aid the person to monitor her/his health and well-being measurements and activities, merge information from his/her personal health care records, activity logs and surrounding environmental monitoring systems, thereby creating situation based personalised digital services to guide in health and well-being issues.

Opening and utilisation of data emerges new kinds of business ecosystems. A data-based business ecosystem is a dynamic structure of organisations each of them having their own part and know-how in the data-based business. The ecosystem's organisations, i.e. actors, affect and are affected by the creation and delivery of each other's offerings [3]. Inside the ecosystem different kinds of value networks are formed by several organisations that focus on fulfilling a certain purpose [4]. Each actor in the network has its role in the flows of information, material, money and influence relationships between each other. Each actor also has its own business model, i.e. how the company

---

[1] http://inspire.jrc.ec.europa.eu/
[2] http://open-data.europa.eu/en

delivers value to customers and how the payment is converted into profits [5].

The transformation to a new kind of business model requires a lot of investment. The business model elements, such as value proposition, key partners, cost structure, revenue stream and market segments must be newly assessed. Since the open data is a quite new concept, its utilisation is still at an early phase. This paper discusses the current status of the usage of information in Finnish companies and estimates how far their businesses are from the open data based business. We carried out 11 interviews with Finnish industry representatives. Our main goal was to untangle the following issues:

- How the companies use information at this moment and what are their development targets in the future?
- How the companies use open data at this moment, and what are the plans and possibilities in the future?
- What value does the information bring to business at this moment and what are the tendencies in the future?

Next section describes the background and starting point for our work. Section 3 describes how we implemented the research, and the results are represented in Section 4. In Section 5 we discuss the emerging issues of our research. Finally, Section 6 concludes our work.

2. BACKGROUND

According to Data-Information-Knowledge-Wisdom (DIKW) hierarchy [6], data is symbols, and information is that data which is processed to be useful, providing answers to the "who", "what", "where", and "when" questions. Data can be raw data or processed data. Raw data is a quantifiable, coarse-grained unit, collected through observations, monitoring, or with the help of questionnaires. Thus, raw data can be a single unit of measurement, such as weather temperature collected by sensors and monitored by context monitors. Data refinement processes analyse, align and aggregate data from different physical and digital sources and thereby increase the understanding of the data. This derived data is sorted for reasoning processes that are able to make decisions on the actions, the applications and services have to take, and moreover, how these actions should be performed [7].

The existing value-chains of data [2, 8-10] can be used as a starting point when identifying actors in open data based ecosystem. The actors' roles in the open data value chain are commonly divided into the data publishing roles and the data end-user's roles [2]. In addition to data providers and data users, the ecosystem also provides opportunities for other actors that provide tools and other utilities for the ecosystem actors.

The existing business models [5, 8, 11-13] and open data communities[3] provide building blocks for the business of open data. It is still obvious that the application of an open data business model requires the transformation from the proprietary side of the software industry to more open model.

Figure 1 describes an open data based ecosystem that has actors that can be organisations, institutes and companies representing one or more roles and having own motives and benefits when operating in the ecosystem. At least four kinds of actor types exist in the ecosystem:

- Data providers provide data for ecosystem, including raw data, derived data and data processing services.
- Application developers create applications based on open data. The applications can be developed for business, for organisation's internal use or for co-operation between partners.
- Application users use the data-based applications and services. These can be individuals or enterprise users that utilise the applications in their own business.
- Infrastructure and tool providers provide the necessary tools, such as the programming tools, configuration tools, and interface for the ecosystem.
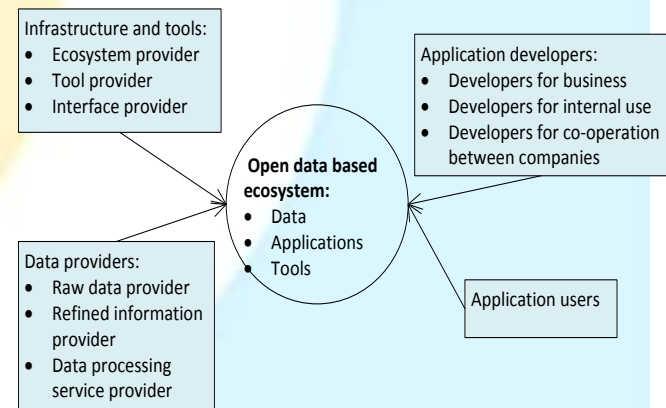

Figure 1: The actor roles in an open data based ecosystem

3. DESCRIPTION OF THE RESEARCH

The main goal of our research is to untangle the current status and future plans of the usage of information in companies. We examined these issues by interviewing the representatives of 11 Finnish companies in total. The interviews were conducted in Oulu, Helsinki and Tampere between June and August 2013. The companies were selected from different application domains to the interviews and interviewees were selected based on their knowledge about business viewpoint of their company. In the study there were general questions about the usage of information. The topics of the interview were divided into three main categories;

- The 'Current usage of information' category included questions about information importance, information content, publicity of data, challenges in management of data, usage of internal and external information, produced information outside company and development targets concerning information. The goal

[3] http://ckan.org

was to understand how companies use information, and to gather the main benefits, challenges and development targets regarding the usage of information. The usage of information was classified into internal and external information to examine what own and/or self-gathered information the company handles, and what information must be bought or otherwise acquired outside the company.

- The 'Open data in a company' category included questions about the current usage of open data, potential of open data in a company in the future and opening company's own data. The purpose was to examine how the open data is now used in the companies, and to find out the future visions and plans of the companies concerning open data.

- The 'Data in business' category included questions about information as a competitive edge, valuable information, usage of information in the future and challenges related to the usage of information. The purpose was to examine how companies use information in their business and what are their future visions.

The interviewed companies differed according to the company size, application domain and service types. The companies were categorised under the main roles based on how they would feel acting in open data based business ecosystem. The backgrounds of the companies are compared in Table 1. The size of the companies is defined according to [14]; micro enterprise <10, small enterprise <50, medium-size enterprise <250, and large enterprise >250 employees.

Table 1. Introduction of companies and their relation to information

Company ID	Company role(s)	Company size	Business type	Internationality of business	Usage of data
Company A	Data/service provider	Large	B2C	International	Information enabler products and services
Company B	Application developer, Tool provider	Large	B2B	International	Data based services
Company C	Application developer	Large	B2B	International	Information applying services
Company D	Data/service provider, Application developer	Large	B2B, B2C	International	Information enabler products and services
Company E	Data/service provider	Small	B2B	International	Data transfer services
Company F	Application user	Large	B2C, B2B	National	Information based products, information services
Company G	Data/service provider	Medium	B2C, B2B	National	Information applying services, information services
Company H	Application user, Data/service provider	Small	B2C	National	Data utilizing services
Company I	Application developer	Small	B2C, B2B	National	Information applying services, data utilizing services
Company J	Application user, Data/service provider	Micro	B2C, B2B	National	Information enabler services
Company K	Tool provider	Micro	B2B	National	Information representing services

4. THE RESULTS OF INTERVIEWS

4.1 The current usage of information

Table 2 summarises the highlights of the interviews concerning the 'current usage of information' category.

According to the interviews, information had an important role in all companies. The content of information was usually customer's content (i.e. the company handles customer owned data), and/or collected information about own customers or self-collected data for business purposes. The first two types of content are private, the

third type of content is used in company's business, e.g., the data is for sale. The following challenges were identified:

- Information storing, distribution and sharing of learnt knowledge. For example, the data structures prevented the data integration that a company wanted to perform.
- The encapsulation of great amount of complicated information to an understandable form.
- Extraction of knowledge or usable information from data. For example, one company found it hard to mine

the essential from the data masses and from social media, and to recognise opportunities of data masses.

- The recognition of the needs of customers and the opportunities of the collected data.
- Distributing the data to partners.
- The data fusion and algorithm development. These were seen both as an opportunity and as a challenge.
- The rapid growth in the use of the Internet
- The legacy information systems set limitations for the development of new products/ services.

Table 2. The usage of information in companies

	Application user	Data provider	Application developer	Infrastructure/ tool provider
Information importance	High	High	High	High
Information content	• Customer's information • Collected information	• Customer's information • Collected information	• Event information • Customer's information • Knowledge	Customer's information
Publicity of data	Private	• Private • Data for sale	Private	Private
Challenges in management of data	• Encapsulation of information • Growth in use of Internet	• Information storing and distribution • Data mining • Recognition of opportunities of data masses and the needs of customers • Data fusion • Algorithm development • Limitations of legacy information systems	Sharing of learnt knowledge	Put data to understandable form
Usage of internal information	• Analysis • Statistics • Forecasts	• Product and system development • Knowledge creation • Refinement and analysis	• Decision making • Management	Communication inside company
Usage of external information	• As part of products • Product planning	• As part of products • Public sector information is valuable	• Understanding current situation in business and technology	Open data is used as a reference in products
Produced information outside company	• Web pages • Reference material • Statistics	• Content for other stakeholders • Interface to data for the partners	• Market/ technology information • Reference material	Reference material
Development targets	Increase the data mining and analytics	• The usage of data • Algorithms development • Citizens as data producers in the future	• Information extraction, trend and information visualisation • Learning from the behaviour of customers • Recognising dependencies in customers' habits • Transfer of knowledge inside company	User registry

Internal information could be numeric, which is needed in decision making and management, or non-numeric, which is often silent information that requires communication between people. In three companies the usage percentage of internal information was mentioned to be low or too low. One company provided a system that was based on self-acquired data. Self-collected data was used to calculate the statistics, and making analysis and forecasts. Mathematical models, analysis and financial statement data were used to produce verbal evaluations for the credit rating of target companies or customers. The deduction of company's strengths and weaknesses were based on these analyses. One company created knowledge from information. This information must be adapted for its usage context. Data produced in research projects was also used in the product development.

The external information was used in one company when needed, whereas another company bought a lot of external information. Especially public sector information was seen valuable. The produced information outside the company concentrated mainly on Web pages or reference materials that were used to show company's competencies. Two companies provided content outside company for other stakeholders to utilise, but not for free. The strategy of another company was to become a data producer in specific areas. One company provided an interface on customer data to its partners with customer's permission, whereas two companies provided statistics about customers for advertisers.

Several companies had plans on increasing the use and utilisation of information. The lack of business ecosystem was mentioned as a problem that prevents the data utilisation. There must be a win-win situation in the ecosystem: all stakeholders should have a real role and benefit from the co-operation. It was also estimated that in the future there could be space for consulting that tries to improve the usage of open data in clients' businesses and to provide added-value with the help of open data. Data ecosystem could improve the usage of data. The data is often big data; it was found difficult to provide it for huge number of customers. There was also a need to increase the data mining and analytics. Two companies had a direct need for extracting significant information from bigger data, both from internal and external data. One interviewee said that different types and details of information must be recognized in the algorithm development.

Systems were required that are capable of obtaining the information from different data sources, visualising the essential information in an understandable way and visualising the trends. Furthermore, there was a need for systems that can learn from the behaviour of customers. It should also be possible to recognize dependencies between customer's habits and other phenomena related to place, situation and time. One company had a tendency to use citizens as data producers in the future. Also, the transfer of knowledge and other information inside a company

requires development. This involves, for example, searching of experts inside, information from previous projects and information about old software systems. One company had already developed a competence centre to assist the sharing of knowledge inside company. Providing non-numeric information for decision making was seen important in several companies. The automatic processing of knowledge and information inside a company is difficult, since this information is not machine readable but requires communication with people. This information is important because it gives explanations for numeric values, also assist in decision making. One interviewee estimated that to succeed, their development targets require a lot of inputs; it requires investments, organisational level support, and maintenance of a development community.

4.2 Open data in a company

Table 3 summarises the highlights of the interviews concerning the 'open data in a company' category. The interviewed companies were interested in the concept of open data. However, open data is a new thing for companies and is not yet commonly used in the interviewed companies. Only one company used open data when quality of data was found good enough. However, the potential of open data has been noted in the companies. One interviewee estimated that open data does not provide direct financial benefits, but it helps the company to maintain the current situation better, and it creates value to customers.

The applicability of data that will soon be opened must be inspected. The utilisation of open data was seen, among others, to improve the quality of the content and making the content more comprehensive. The open data could also add the value of companies' products. One company tended to provide open data to its customers in the near future. The integration of open data with company's own data was seen as a new possibility in one company. Both slowly and rapidly changing information, as well as refinement of information, was seen to create value for customers. However, in most of the companies the rapidly changing information was seen more important. Most of the companies could not provide their own data, since the collected data is personal and protected by privacy rules, or the data is customer's data (customer owns the rights to data). Three companies inspected the possibility of providing data. One of them considered providing open data to its customers. The second one would like the benefit financially of the data. The third one could also provide data, but since the collection of data causes costs, it is not possible to provide this data free for customers. Two other companies considered to provide data for research use. The published results should somehow benefit the company. Two companies considered to provide open interfaces/APIs for application developers to give applications access to data.

Table 3. Open data in a company

	Application user	**Data provider**	**Application developer**	**Infrastructure/ tool provider**
Current usage of open data	Not used	Not used or used when the quality is good enough	Not used	Not used
Potential of open data in a company in future	• Improving the logistics • Using data in company's operations • Data integration	• Providing data for customers • Using new data in products • Slowly and rapidly changing data	• Using data in decision making • Creating value for customers • Using data in products	Adding value to the products
Opening company's own data	Cannot be opened due to privacy rules	• Can sell data • Can provide open APIs for data • Only for research use	Part of data could be provided as an open data	Cannot be opened; the customer owns the rights to data

4.3 Information in business

Table 4 summarises the highlights of the interviews of the 'information in business' category. Two of the interviewed companies did not have direct competitors. One company had competitors but does not feel competing with them but instead concentrates on attracting customers with quality and services. The rest of the companies have many competitors. For some companies the information provides a competitive edge for many businesses, whereas the products' certain functionality or uniqueness are the strengths for some companies. In a traffic domain the real-time data was seen to provide competitive edge for actors. Information refinement, visualisation, statistics and reports were seen to assist in real-time and automatized or semi-automatized decision making. Processed and refined data was used in company's strategy planning. It also assisted in tailoring services to different customer segments. Especially predictions based on analysis of the behaviour of customer segments were seen important. In the future, the selling of data and data-based knowledge was seen to have potential, as well as the information about markets and technological possibilities. It was estimated that open data can both decrease and benefit the business. If the information is available for free, the customers may not want to use ready-made services that handle this data. On the other hand, when there is a lot of data available for free, many customers do not want to search data manually from different places but want to buy services that mine useful information from the available data sets. Open data can also provide some other new possibilities; a part of data can be provided as open and then additional chargeable services can complete the free service. Also, the selling of certificates can provide new business opportunities.

Generally, the companies were willing to pay for information if the customer pays for it. Companies must try to understand what information is important to customers, since all information that produces value for customers was seen as valuable information. Fresh information was estimated to be more valuable for customers. Three companies used chargeable information directly as a part of their service or product. There could be different contracts with different data providers, e.g., pay per use payments and monthly payments for used data. Information about line of business and information about internal situation was mentioned as valuable information. Public business information and internal information together assists in understanding the line of business (e.g., trends, leads) and the market situation in hand. Certification, validation and meta-data were seen to increase the value of information. There must be confidence of quality of data. Therefore, it is important to know the data source, to have descriptions of data and to know quality of data and its source. The quality of data is extremely important in risk management business.

The interviewees concluded that the importance of information is increasing. More and more services will be constructed around information. Information is important for business and a significant way to differentiate from competitors. There will also be a demand for data integration and analytics. Several interviewees believed that the popularity of open data will provide business opportunities for companies that produce processed information, for software integrators and for application developers. Three companies had a great demand for public data that is not yet made available. The interviewed companies had several information-related challenges at this moment. First of all, the legality of opening the data must be estimated in the case when the information considers individuals. Secondly, the personified information should transfer to more collective information. One company had data that could be opened but was not willing to do that since the competitors could take benefit of the data. The information sharing was seen to be important but challenging. Information sharing was estimated to be much easier in B2B than in B2C, since the latter lacks business models. The knowledge sharing inside a company was also seen as challenging. One interviewee

found it difficult to know what they already know inside their company. The integration and interoperability of data and representation of information was also considered challenging. Data is often in multiple sources, and it is also difficult to extract the important data from bigger data sets. There is a need for certification of data sources: data quality must be available, to all measurement information there should be attached meta-data about accuracy of measurement values. Furthermore, since the databases will be in clouds in the future, the ensuring of the data security, privacy etc. will be problematic.

The interviews revealed that APIs are more important than platforms. The data formats are often incompatible and open data's APIs do not often follow standards at this

moment. Thus, there is a great need for standardised APIs. Unfortunately, open data's APIs do not often follow standards at this moment. In addition, it seems that a data broker that enables the publishing of data and sharing of data is needed, too. The representation of information was found difficult, there should be ways to visualise information to different user groups. Some companies did not know what data is already available that could be marketed to customers. Thus, the available open data is not fully utilised at this moment. It is important to show (via concepts) how customer will benefit from information. Also, more real-time information was required for customers, such as targeted data, predictions, and information related to customer's location or situation.

Table 4. Information in business

	Application user	Data provider	Application developer	Infrastructure/ tool provider
Information as competitive edge	Concentrates on attracting customers	• Real-time data • Processed, refined data • Data-based decision support systems • Predictions • Data as part of product • Data used in product development • Selling of refined information • Strategy planning	• Information refinement and visualisation • Statistics • Used in decision making • Customer analysis • Predictions • Selling of data-based knowledge • Market information	To put data to a right form
Valuable information	• Product's content • Anything that customer is willing to pay	• Reliable/ high quality/ fresh data • Data in product's content • Anything that has value for customers • Certified, validated data • Meta-data	• Anything that produces value for clients • Information about line of business, • Information about internal situation and market situation • Information for decision making	Anything that customer is willing to pay
Usage of information in the future	• Data integration and analytics • More use of processed information • Attracting and guiding customers • Providing data for advertisers	• Making own information available • More accurate customer analysis • Providing an open interface • More real-time information • A great demand of data that is not yet available	• Integration of information from different sources • Passing information forward via APIs	• Data to design new kinds of products • Data to develop new ideas
Challenges related to the usage of information	The legality of passing data to the advertisers	• Information sharing, refinement and representation • Lack of business models • Development requires a lot of inputs • Data security/ privacy in clouds • A need for audited data sources • Data quality	• Incompatible data formats • Sharing of data and knowledge • Information refinement and representation • Collecting data	No information about what data there is already available

5. LESSONS LEARNT

5.1 Motivation

The interviews revealed several motives for using open data in business. The open data can be used to:

- *Add value of data, information and knowledge* – The interviews revealed that open data can provide additional value for (private) data and make the data more attractive for customers. Open data can be used for improving quality and accuracy of data and for providing meta-information for data. This improves the quality of information and knowledge produced from source data, too. Better information and knowledge will provide a competitive edge.

- *Improve usability of data, information and knowledge* – Open data can be used in the representation of data, information and knowledge. For example, open maps can be used in the visualisations of data.

- *Support management and decision making* – Fresh and reliable open data can assist the comprehension of the situation in hand. For example, statistics and identification of trends is important in management and decision making.

5.2 Opportunities

In accordance with the interviews, the open data was seen valuable both in business and in company's internal use. The following features summarises how data is now related to the business of the interviewed companies:

- *Refinement of data* – A company operates as an actor in a data flow: it produces content from data or integrates open data with own data and then sells the content to others that utilise and produce refined information and knowledge.

- *Service development* – A company operates as an actor in service development: it participates in a service co-creation process. It produces and shares data among a set of actors between organisations, and gets its share of the service fee. Open data enables new innovative and attractive data-based services. Data integration, especially the integration of open data with private data, is an interesting way to develop new applications and services. Open data also improves the services, bringing new ideas and content. Thus, open data enables the new business possibilities, and also assists in maintaining the current business better.

- *Information and knowledge based management and decision making* – A company uses data for process and business improvement: The company buys or collects (private) data or information, and makes statistics and analyses of data. Using the information the company can a) improve its internal knowledge sharing, b) target to innovative services or/and improve its existing services, and c) predict the future trends.

Open data was not widely used in the businesses of the interviewed companies. In addition, the companies were much more interested in utilising open data than acting as providers of open data. However, the business of the interviewed companies strongly relates to data. Thus, the results could be different if the information and information processing is not a part of the core business of a company but only a way to boost the core business. However, open data may provide several new and not yet widely identified business opportunities:

- *Certification of open data* – A company produces data for certain service developers that use the data in their services. The contracts (e.g., Service Level Agreements (SLAs)) and data fees ensure the quality and permanency of data for both parties.

- *Business-driven opening of public data* – There should be catalogues for possible but not yet opened public data and practices for companies to open this data for their purposes. This would be a win-win situation: the company-driven opening of public data provides new open data sources for all actors and at the same time directly enables the companies to get the data for their needs. In addition, this would ensure that the commercially most important data is opened first.

- *Out-sourcing of the development of information services and applications in the private sector* – A company provides an open interface to data for third-party application and service developers. Opening an interface requires a small investment from a company, but the company benefits from the development efforts of other actors: a) the selling of the product of the company increases as it becomes more attractive via applications, b) the different user groups are achieved since more applications are provided for different application domains, and c) the size of the ecosystem will expand since it becomes more interesting to developers. For example, a fun park provides an API to open data that shows how many customers are currently in the fun park. External service providers can use this API in their digital services that are exploited by different stakeholders; i) the customers can go to the fun park when there are shorter queues, ii) the utilisation rate of the fun park is smoother, and iii) external service providers will get more customers for their digital service.

Data provider can be anyone that can open its data without restrictions (e.g., privacy rules). Besides public administration, data providers are companies that open their data if identified beneficial for company's business. The open data based applications and digital services have potential in several areas. Accurate analysis services based on open and private data will be in demand on several application domains. Open data based business provides a great advantages most of all to those companies that can integrate and analyse a huge amount of (open) data at real-time.

In addition to service and application development, open data will provide big opportunities for algorithm providers for different purposes. Data mining, extracting and sorting

has a lot of potential in business and in company's internal use. The information and knowledge (produced of data) assists in business improvement, business creation and decision making. The customer analysis and predictions produces a lot of beneficial information to companies about customers' behaviour, preferences and future demands.

The application developers can benefit from the identified open data based business opportunities and vice versa; i) applications can be enhanced by the use of open data; ii) the business and organisation models may be used as such or adapted according to the situation; and iii) the crowd-based application development approach may be used in the both cases because the end-user itself can be the core in the development of applications and personalized digital services based on open data.

5.3 Obstacles

The interviews revealed obstacles that limit the use of open data and opening of data. The obstacles are discussed as under:

Unidentified customer needs – It is important to understand what data is valuable for customers and how the data should be delivered. This requires rich communication with the customers; the company must inform customers about possibilities of available data sources, inquire about customer needs related to data and understand how data is utilised in the business of the customers.

Lack of suitable open data – At this point only a minor part of public sector data is opened. In addition, the usage of open data requires suitable data. Open data will provide a lot of potential if it has certain properties. Valuable open data is characterised as consisting of following features:

- Dynamic; e.g., indicators/events from fast changes
- Reliable; reliability of data sources and data itself
- Real-time; continuously updated data
- Permanency; the guarantee that data is available in the future
- Integrability; the data can be used with private data

Trustworthiness – Before using a data, the companies must ensure that it is trustworthy. The trust must be established both for data provider and for the data. For the data, trust must include both security and reliability. Especially in the case of private data, unauthorised users may not access the data. Data reliability ensures that the data is accurate and complete; nobody has changed it or made own conclusions.

Data and APIs are not yet open – The interviewed companies were willing to use open data but were not in large scale willing to open their own data. The companies are technically able to open data and APIs, but the obstacles seem to be the lack of motivation and knowledge on benefits on opening data, lack of business models, and the lack of new operation models. New business models are required to motivate the companies to change their operation models. In addition, there is a need for an open

data based business ecosystem that enables the win-win business model for each ecosystem actor. Data interoperability issues should be taken into account at the very beginning; heterogeneous data and non-standard interfaces of open data complicates others for utilising the data.

Lack of business policies – To benefit from open data, requires training for using and opening data. The companies have to inspect carefully what data or what part of data could be opened. There might be a need to initiate new company policies, since it might be possible that opening data is against the company policies or against the company's business model. The data can benefit only competitors and ruin the company's own business. Thus, the open data based business model must verify that the data opening is beneficial to the company itself.

Insufficient marketing of open data – The logistics of open data delivery should be improved also; the user has to know what data/information is available, and where and when it is available. Up-to-date summary of relevant information should be provided. In the case of real-time data, the data updating issues must be taken into account, too.

Many existing technical solutions and practices, such as information and behaviour interoperability based on shared ontologies, run-time quality management, semantic web technologies and tools for developing interoperable applications, can be exploited and adapted for open data and crowd based digital service development. However, there are specific issues that concern open data based digital services; e.g., Who owns the data? Who is willing to pay for it? How to separate open and private data during service evolution? There still remains a lack of solutions and practises that hinder the utilisation of open data in digital services.

5.4 Future research topics

For business, the future research should concern a way the companies could ensure and analyse whether opening their data provides new business potential. If the companies could forecast and calculate their potential, they might be willing to open their data in the case when the cost of collecting data will be covered by the benefits achieved from opening data. The use of open data in digital services should be researched also from the customer's viewpoint; Is a customer willing to pay on an open data-based digital service?

The intellectual property rights (IPR) in the case of open data is still unclear. For example, if open data is used in knowledge creation, how the knowledge can be protected? The protection of immaterial products (i.e. data, information, knowledge) is difficult, and therefore instructions are required to clarify how the IPR issues are solved in the commercial use of open data.

The interviews revealed that several companies were interested in data that is not available at this moment. Especially the public administration had a lot of data that was demanded to be opened. Also, at this moment it is

hard to detect already available data. Some kind of public catalogue of available open data is required for service providers who use it. A catalogue may also assist to determine the most demanded, not yet open data.

The quality and the permanency of the open data must be ensured before starting business using it. If the open data is an important part of business, the companies could be willing to pay for data quality and permanency and thus ensure the permanency of their own business. Thus, new kinds of SLAs are needed for open data to guarantee the permanency of the business of service developers. SLAs could also motivate the data providers to maintain the quality of data or even to make the better versions of data.

6. CONCLUSION

Open data provides new kinds of business opportunities to data providers as well as data utilizers. Open data based business requires new kinds of business models and also an ecosystem in which each actor must define its role related to open data.

The study revealed that the interest towards utilising open data is high, and open data is seen as a new trend that has to be followed to keep the business up to date. However, most of the companies were not yet motivated to open their own data, mainly due to the lack of business and operational models. We identified several opportunities for open data, but also some obstacles limiting the usage of open data in business. A lot of work must still be done to overcome these obstacles and also to research the identified new topics to enable the open data based business. Once the obstacles have been overcome, the future applications will increasingly provide open data and utilise open data.

ACKNOWLEDGEMENT

REFERENCES

[1] Auer, S.R., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., Ives, Z., 2007. DBpedia: A nucleus for a web of open data, *Lecture Notes in Computer Science, vol. 4825*, pp. 722-735

[2] Poikola, A., Kola, P., Hintikka, K.A., 2011. *Public data - an introduction to opening information resources,* Ministry of Transport and Communications, Helsinki, Finland Available at: http://www.scribd.com/doc/57392397/Public-Data

[3] Iansiti, M., Levien, R., 2004. Creating value in your business ecosystem, *Harvard Business Review, March 2004*

[4] Lehto, I., Hermes, J., Ahokangas, P., Myllykoski, J., Collaboration in cloud businesses – value networks and ecosystems, *Communications of the Cloud Software*

[5] Teece, D.J., 2010. Business models, business strategy, and innovation, *Long Range Planning, vol. 43*(2-3), pp. 172-194

[6] Ackoff, R.L., 1989. From data to wisdom, *Journal of Applied Systems Analysis, vol. 16*, pp. 3-9

[7] Pantsar-Syväniemi, S., Simula, K., Ovaska, E., 2010. Context-awareness in smart spaces, *IEEE Symposium on Computers and Communications (ISCC)*, Riccione, Italy

[8] Tammisto, Y., Lindman, J., 2011. Open data business models, *The 34th Information Systems Seminar in Scandinavia*, Turku, Finland, pp. 762-777

[9] Kuk, G., Davies, T., 2011. The roles of agency and artifacts in assembling open data complementarities, *Proceedings of Thirty Second International Conference on Information Systems*, Shanghai, China, pp. 16

[10] Latif, A., Saeed, A., Hoefler, P., Stocker, A., 2009. The linked data value chain: A lightweight model for business engineers, *International Conference on Semantic Systems*, pp. 568-575

[11] Osterwalder, A., Parent, C., Pigneur, Y., 2004. Setting up an ontology of business models, *Proceedings of 16th International Conference on Advanced Information Systems Engineering (CAiSE03) Workshops, 3*, pp. 319-324

[12] Perr, J., Sullivan, P., Appleyard, M.M., 2006.*Open for business: Emerging business models for open source software companies,* working paper, Lab2Market. Portland State University

[13] Baden-Fuller, C., Morgan, M.S., 2010. Business models as models, *Long Range Planning, vol. 43*, pp. 156-171

[14] European Union Commission, 2003. Commission recommendation of 6 may 2003 concerning the definition of micro, small and medium-sized enterprises, *Official Journal of the European Union, L 124, vol. 46*, pp. 36-41

Author's Biography

Anne Immonen received her MSc degree in 2002 from the University of Oulu. Since 2002, she has been working as a research scientist at VTT Technical Research Centre of Finland. Her main research topics include reliable service engineering, quality modelling and analysis of composite services, and quality ontologies. Her current interest is business models and ecosystems in open data based application development.

Marko Palviainen received the M.Sc. (Tech.) degree from the Lappeenranta University of Technology in 1998 and the Ph.D. (Doctor of Technology) degree in Computer Science from the Tampere University of Technology in 2007. Since 1999 he has been working as a research scientist at VTT, Technical Research Centre of Finland. His current areas of research interest include mobile applications and application development methods, Ontology-Driven Software Engineering, and evaluation methods of parallel applications.

Eila Ovaska obtained the PhD degree in 2000 from the University of Oulu. Initially she worked for twenty years as a Software Engineer, a senior research scientist and the leader of the Software Architectures Group at VTT. Since 2001, she has been working as a Professor at VTT and an Adjunct Professor at the University of Oulu. Her current interest is in service architectures, especially, in self-adaptive digital service systems and services. She has acted as a workshop and conference organizer and a reviewer for scientific journals and conferences. She has co-authored over 130 scientific publications. She is a member of the IEEE and IEEE Computer Science.