IEEE Access
practical innovations : open solutions

Received December 2, 2013, accepted January 3, 2014, date of publication January 31, 2014, date of current version February 18, 2014.

Digital Object Identifier 10.1109/ACCESS.2014.2302872

Requirements of an Open Data Based Business Ecosystem

ANNE IMMONEN[1], MARKO PALVIAINEN[2], AND EILA OVASKA[1]
[1]VTT Technical Research Centre of Finland, P.O. Box 1100, FIN 90571 Oulu, Finland
[2]VTT Technical Research Centre of Finland, P.O. Box 1000, FIN 02044 Espoo, Finland

Corresponding author: A. Immonen (anne.immonen@vtt.fi)

This work was supported by the National Strategic Research Project, Open Data End-user Programming funded by the Finnish Funding Agency for Technology and Innovation and the VTT Technical Research Centre of Finland.

ABSTRACT Emerging opportunities for open data based business have been recognized around the world. Open data can provide new business opportunities for actors that provide data, for actors that consume data, and for actors that develop innovative services and applications around the data. Open data based business requires business models and a collaborative environment—called an ecosystem—to support businesses based on open data, services, and applications. This paper outlines the open data ecosystem (ODE) from the business viewpoint and then defines the requirements of such an ecosystem. The outline and requirements are based on the state-of-the-art knowledge explored from the literature and the state of the practice on data-based business in the industry collected through interviews. The interviews revealed several motives and advantages of the ODE. However, there are also obstacles that should be carefully considered and solved. This paper defines the actors of the ODE and their roles in the ecosystem as well as the business model elements and services that are needed in open data based business. According to the interviews, the interest in open data and open data ecosystems is high at this moment. However, further research work is required to establish and validate the ODE in the near future.

INDEX TERMS Business ecosystem, open data.

I. INTRODUCTION

In the future, an increasing amount of services and applications will be developed based on open data. Open data are data that are freely available to everyone to use and republish as they wish without restrictions of copyrights, patents or other mechanisms of control [1]. The benefits of open data have been widely recognized around the world, and there has been a tendency in many countries to open the data of the public sector.[1] In particular, the data that are collected on tax revenues are obligated to be opened in many countries.

Private companies could also open a part of their own data. In addition to earning direct profits from data sales, the opening of data can provide other benefits, such as new partners, new interests in the company's main products/services, new kinds of business activities and new customers for the product/service as a result of data-based applications. However, opening data requires a big change in a company's business. The lack of knowledge on the benefits of opening data, the lack of business models and the lack of new operation models are the main obstacles that explain why companies are not currently motivated to open their own data [2]. The developers of digital services and applications could greatly benefit from the business opportunities of open data. However, the lack of business ecosystems and business models has been identified as the main obstacle to data utilization in services and applications [2].

A business ecosystem is a dynamic structure of organizations that work together in a specific primary technological platform or core business [3]. In an ecosystem, value is not created in a chain but more in a network of actors. A data-based business ecosystem is formed by organizations that each has their own parts and know-how in the data-based business. The ecosystem's actors affect and are affected by the creation and delivery of the offerings of the other actors. Each actor also has a role in the flows of information, material, money and influence relationships between one another. Existing value chains [4]–[6], business models [7]–[11] and open data communities[2] provide building blocks for the business of open data. However, there is still a need for

[1]http://open-data.europa.eu/en

[2]http://ckan.org

an ecosystem to support both the technical perspective and business perspective of an open data based business.

The main contribution of this paper is to outline the Open Data Ecosystem (ODE) from the business perspective and to define the requirements of the ODE that must support different actors and business that are formed around various kinds of data, services and applications. This paper describes:

- The identified actors of the ODE.
- The services that the ecosystem should provide for the open data based business.
- The business model elements required in the open data based business and the description of how the ODE should support these elements.

The ODE outline is based on the state-of-the-art knowledge on business ecosystems, models and actors, and the state of the practice on data-based business and future visions in the industry on open data based business. This knowledge is based on a thorough literature survey carried out in the spring of 2013. The knowledge of the state of the practice is based on interviews conducted among Finnish companies in the summer of 2013. The ODE concept is novel: to our knowledge, there is no model published on open data ecosystems to date.

This paper is organized as follows. The next section compares our definitions in this work with the state of the art and indicates what parts of existing business models can be used in our work. Section III describes the initial outline for the ODE based on a literature analysis. Section IV describes the implementation and results of the company interviews and specifies a new version of the ODE that is modified according of the requirements collected in the interviews. Section V provides a discussion related to the characteristics of the ODE and the results of the company interviews. Finally, section VI makes concluding remarks.

II. THE STATE OF THE ART OF OPEN DATA BASED BUSINESS

A. TERMINOLOGY

The following terminology is used in this paper:

Raw data – Raw data are data that are produced by observing, monitoring, using questionnaires, etc. but have not yet been processed for any specific purpose.

Information – A refinement and processing of data will produce information from the data. A refinement of data can analyze, align and aggregate data from different physical and digital sources and thereby increase the understanding of the data. Raw data can be refined to increase the understanding of the data. This derived information is sorted for reasoning processes that are able to make decisions on the actions that the applications and services have to take and, moreover, how these actions should be performed.

Knowledge – Knowledge can refer to the theoretical or practical understanding of a subject. It can be implicit or explicit, and it is more or less systematic. Here, knowledge is used in both meanings: theoretical knowledge represents explicit knowledge on the meaning of data. Practical knowledge is implicit and less systematically collected, represented and shared. This knowledge is related to experiences of using (open) data in business.

Application – A combination of digital services that provides the data to the different end-users of data in their preferred representations when and where they are needed.

Service – A digital service that provides additional value for data processing and can for example, support data collection, analysis, sharing and/or representation. A service can mine and extract data from input data and produce relevant data for a particular context or domain.

Service chain – A set of services performing data refinement and processing steps and ultimately making the derived information available to users.

B. VALUE CHAINS OF DATA

Latif et al [6] define a linked data value chain that has four entities: a raw data provider, a linked data provider, a linked data application provider and an end-user. The chain supports multiple sources of data; i.e., the data may be acquired from several data providers and may be provided to several application providers. Kuk and Davies [5] introduce the assembly of complementarities involved in the chain from raw data to data-based services. There are parties that structure the raw data, make the data linkable, analyze or visualize the data, share the data within the source code of software and ultimately allow the developers to innovate services on top of the source code. Poikola et al. [4] define the roles in the open data value chain from the data publishing perspective and the end-user's perspective:

- The data publishing roles include: a Storer to collect and save raw material, a Developer to manage and process raw material, an Aggregator to combine and edit data from different sources, a Harmonizer to standardize and homogenize data from different sources, an Updater to update information, a Publisher to publish the data and a Register to maintain the administration of data resources.
- The data end-user roles include: an Application developer to utilize the data as part of the service, an Interpreter to interpret the data and a User of data-based services, e.g., an individual, company, or organization that uses open data applications and interpretations.

Tammisto et al. [11] have conducted research on the roles of linked-data developers and application developers in a Finnish context. The interviewed companies identified three developer roles: a consultant, a linked data developer and an application developer. The consultant role was seen as an additional source of revenue for the open data companies through consulting the raw data providers about the options and possibilities. Moreover, Chen et al. [12] identify two new roles related to data analytics; Data-as-a-Service (DaaS) providers collect, generate, and aggregate the content (i.e., data), and Analytics-as-a-Service (AaaS) providers deliver analytics services to analytics consumers. In addition, data value chains can include other non-profit roles, which support the finding, publishing and marketing of open

data sources, promoting the use of open data and networking-related data. For example, communities using the CKAN, an open-source data portal platform (http://ckan.org/), provide a huge number of applications and visualization components and libraries for the utilization of open data, as well as a regional open data network, the Helsinki Region Infoshare (HRI) (http://www.hri.fi/fi/), which intends to make relevant data easily available.

C. BUSINESS MODELS OF DATA

The business model must support the value proposition of a company in an ecosystem. According to [10], a business model defines how an enterprise delivers value to customers, entices customers to pay for value, and converts those payments to profit. Baden-Fuller and Morgan [9] summarize some common business models, stating that a business model has several characteristics other than value proposition, such as describing business behavior and organizing the company. The business model must describe how value is captured from the innovation. Traditional business models concentrate on gaining profits by overtaking competitors and keeping strict boundaries around the company. Open data force companies to re-think their business strategies and models because an open data based business cannot be shut down within the boundaries that surround the company. The transformation to an open data based business model requires a great deal of investment and newly assessed business model elements. The open data based business model can be based on the capabilities and features of open innovation strategies, open-source software business models, the business models of cloud services and the business models of analytics.

Open innovation breaks the boundaries around a company in the innovation phase; companies can create ideas by themselves, use external ideas or co-create ideas with other companies or with the actors of other communities. Two open innovation processes in service innovation have been identified by [13]: outside-in and inside-out. In the outside-in process the components of external knowledge and innovation are used in service development, whereas in the inside-out process, a company allows external parties to use its knowledge and innovation components in service development. Accordingly, the outside-in process can be applied by application developers, whereas the inside-out process is more appropriate for data/information providers.

Seven *open business models* identified within the context of open-source software (OSS) [8] can be classified into four categories based on how they capture value [14]: deployment, hybridization, complements and self-service. The deployment category includes support, subscription and professional services/consulting business models, which are similar to the proprietary side of the software industry. The hybridization category includes proprietary extensions and dual-license business models, attempting to attract customers by licensing to familiarize the customers with the product/service. This kind of proceeding could easily be applicable for open data providers as well as application developers. In the complements business model, open source software is provided with the vendor that sells and supports the hardware device or appliance. Finally, in a self-service business model, users with similar needs pool their resources and create applications for the community's needs. The self-service model can be applicable for application developers that do not sell applications but otherwise use them in their business (e.g., in communication between partners or inside an organization). According to [8], the modification of a company's business model is the trade-off between underlying value creation dynamics, IP ownership/license choice, community management, and target market/product categories.

The *cloud business model* is especially applicable in the case of large data sets, when storing, processing and analyzing require a great deal of resources. In the future, it can be assumed that large data applications will be the main driver of widespread cloud adoption [15]. A survey of cloud adopters reveals that decision makers currently implement public cloud applications and platforms mostly for business agility [16]. The same survey revealed that the latest technology and support for mobile workers are increasingly significant factors in the decision to move to cloud applications. The transformation towards the cloud business model has already been researched by [17], which reveals that the transformation affects all elements of the business model mainly due to customer-side characteristics, which include pay-per-use pricing, ubiquitous access and on-demand availability.

Data analysis (e.g., mining, extracting and sorting) has a high potential in business. In addition to being used in business, data has been found to be valuable in information- and knowledge-based management and decision making inside companies, helping in the understanding of the line of business and the market situation at hand [2]. The emerging *analytics business models* include the proprietary model, the shared data model, the shared analytics model, the shared value model, the co-development model and the new business development model [12]. The last model describes how DaaS and AaaS provide opportunities for application developers to create new business. In addition, the shared value model and the co-development model are suitable in open data based business; the end-users and partners create value together, and a set of companies participates in the development.

D. PRICING MODELS

There exist several pricing models that can be used in data- and information-related business. Services and applications can be priced commonly based on *features* [18] or *performance* [19], or the customer is charged a predefined price for customer-tailored services and applications usage [20].

The traditional pricing models, such as the *Value model* [21], *Portfolio pricing* [22] and *Market pricing* [23] are applicable for pricing the following three core elements: data, services and applications. In addition, the *Cost-based model* [19] is applicable when data, services or applications are added to the actual products or

customized according to the customer's needs. The service-related pricing models are applicable to all of the core elements; in *Pay-per-use pricing model* [24], the customer pays only for the data/service/application usage, whereas in the *Subscription model* [24], the client pays a fixed price for a certain time frame. Furthermore, information/Internet-related pricing models are also applicable. For example, the *Flickr multiple revenue stream model* [10] involves collecting subscription fees, charging advertisers for contextual advertising, and receiving sponsorship and revenue-sharing fees from partnerships. The *Freemium model* [10] is a free, limited-functionality version of the product/service offered to attract users, hoping that some users will pay a premium for advanced features, whereas the *Free Trial model* [25] offers a free trial of the service for 14 or 30 days to attract users, after which the users are required to pay to continue using the service.

III. THE FIRST OUTLINE OF THE ODE

The objective of the ODE is to promote open data based business by making the development of open data based applications and services easy and straightforward. The first outline of the ODE defines the core actors and elements of the open data based business. The actors have their own motives and benefits when operating in the value network of data. Each actor represents one or more roles in the ecosystem. We decided to use the existing value chains of data as a starting point for identifying the stakeholders of the ODE. It seems that there are three core elements in the ODE that businesses are formed around: i) data, ii) services and iii) applications. The applications use data and services and produce valuable information and knowledge for the user.

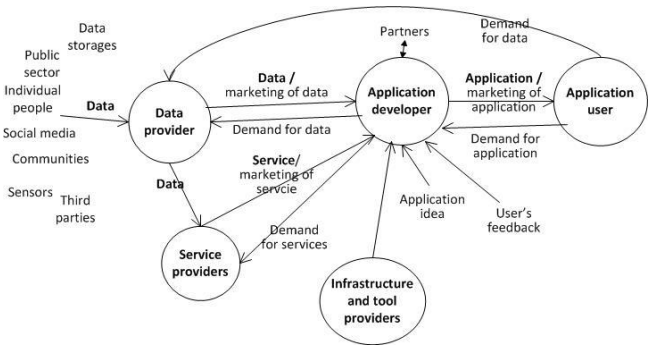

FIGURE 1. Initial actors and their relationships in an open data ecosystem.

A. THE ACTORS OF THE OPEN DATA ECOSYSTEM

There are (at least) five roles for the ODE's actors: i) data providers, ii) service providers, iii) application developers, iv) application users and v) infrastructure and tool providers. Fig. 1 depicts how data flow from the data providers to data consumers. Data providers make data available to other stakeholders, service providers produce services related to data, and application developers use the available data and services and develop applications for the data. Finally, the application users consume the data and services with the help of applications. The infrastructure and tool providers offer other utilities to the actors of the ODE. The following subsections describe the roles of the ODE, the services required for the ODE and the identified data-based business elements in greater detail.

1) DATA PROVIDERS

Data providers are organizations that provide data for the other actors of the ecosystem. The provided data can be raw data, refined data/information or analyzed information. Data providers can be divided into two groups according to their motive: organizations that provide data "for free" without any conditions or with some licenses that restrict the use of data and organizations that that do business from selling access to the data.

Organizations that provide free access to data are usually public administrations or other public entities that have a lot of data but no abilities or resources to use the data in the form of data refinement or the development of services with the data. These organizations provide data to improve the national economy, enabling enterprises and citizens to exploit the data. Several public licenses exist with which a licensor can provide access and copyright permissions to open the data, such as the Creative Commons License and Conformant Licenses. Originally, these licenses granted the "baseline rights" to distribute a copyrighted work without changes at no charge [26]. Most licenses currently contain some license elements that restrict the utilization of data, such as Attribution, Non-Commercial, No-Derivatives and Share Alike. These elements can be mixed and matched to produce a customized license.

The data providers who sell data can attract users to pay for data by providing only a subset of their data as open data, providing guarantees of availability only to paying users, limiting the frequency of access to the open data, providing open access only to stale information, using share-alike licensing, or requiring that people register to access the data [27]. Contracts such as Service-Level Agreements (SLAs) and data fees ensure the quality and permanency of data for the users of the data.

2) SERVICE PROVIDERS

Service providers offer services related to data and can earn income from the usage of services. Those organizations that do not have the abilities and resources to perform the data processing themselves can buy the data processing services from service providers or buy the processed data directly from data providers. Thus, the service provider must: a) identify the needs of customers, b) produce relevant data from input data to a particular context or domain and c) represent the produced data in a usable way. However, it is important to note that a service provider does not necessarily provide a complete service for the user but can simply provide a part of a service chain. These providers may provide ready-made service chains or these service chains may be composed at run-time. In order to make the usage of service chains

easy, these service chains can also be represented as a single service for the user even though the service is composed of several smaller services. In our vision, there are data drivers (described in more detail in [28]–[30]) that make the data and services available to application developers. The data driver services and easy-to-use tools facilitate the application development.

3) APPLICATION DEVELOPERS

Application developers cooperate with partners and innovate applications around open data or use open data as such or integrated with their own data in their applications. An application idea can be provided outside the organization, and the application users' feedback affects the continuous and iterative development of applications. The ODE provides tools and services for developing open data based applications. The applications are created by combining available data and services. Therefore, three kinds of stakeholders participate in application development: application developers create the applications, and data providers and service providers deliver data and services for the applications. Application developers pay for the use of the data driver services and possibly for the use of the data. The revenues are then shared between the application developers, data providers and service providers.

4) APPLICATION USERS

Application users consume data with the help of data-based applications and services. A user can be a consumer, citizen or an enterprise user. A consumer is a user that has bought a commercial application from an application store. A citizen can be user that uses the provided application as a citizen. For example, the application can enable a citizen to produce information from the environment and then consume information provided by the public administration. An enterprise user is a user who uses the applications in business.

5) INFRASTRUCTURE AND TOOL PROVIDERS

Infrastructure and tool providers offer the necessary tools for the ecosystem. The relevant roles are:

- ODE providers who maintain the ecosystem and receive income from the usage of services and applications of the ecosystem.
- Marketplace providers who provide a marketplace in which applications and data driver services can be bought.
- Tool providers who provide tools to develop applications, configure applications for different user needs and execute and control the application. The tools are used as services, and the provider receives income from the tool usage (pay-per-use).
- Cloud service providers who provide the physical facilities for the ecosystem and receive income from the facilities' "rent."

B. THE SERVICES OF THE OPEN DATA ECOSYSTEM

The ODE should provide and deliver the following services to its members:

- Data provider support: The ecosystem has to enable different open data providers to provide data to the ecosystem from heterogeneous data sources.
- Data adaptation: The ecosystem must enable different service providers to develop and provide driver software that is used for adapting the open data to be usable in different applications.
- Tool support: The ecosystem shall provide tools for application developers to create applications, services and new driver software.
- Diverse applications: The ecosystem has to enable and provide application deliveries for diverse application users.

C. THE BUSINESS MODEL ELEMENTS OF THE OPEN DATA ECOSYSTEM

There are several developed business models that are also applicable to data [7], [10], [31], [32]. Firstly, we decided to use the business elements of the Business Model Canvas by Alexander Osterwalder [7] as a basis because it is a well-known model and provides a stable template for developing new business models. The following business elements were identified to be needed in open data based business (summarized in Fig. 2):

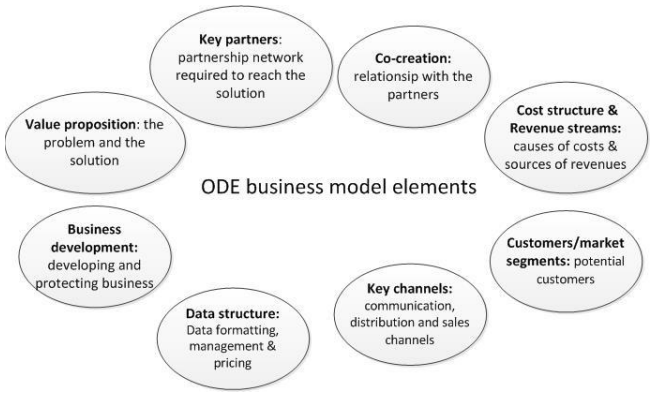

FIGURE 2. Business model elements for open data based business.

Value proposition – This concept deals with the problem and the suggested solution, the value of the solution, identifying a business opportunity based on a demand or innovation and creating a value network or a service chain to reach the solution, identifying the value of the solution to different network actors and identifying the obstacles to reaching the solution.

Key partners – The key partners in the value network include those that provide the solution to the problem and other ecosystem actors that are required to achieve the solution. This concept also deals with the types of relationships with the partners and the benefits/rationale/risks of partnership/cooperation.

TABLE 1. Summary of the interviewed companies.

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/information-applying 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

Co-creation – This concept involves forming cooperation with partners, identifying the key activities and resources for reaching the solution and the division of the key activities with the partners, and the role of the customer in reaching the solution (e.g., requirements or feedback).

Cost structure and Revenue stream – The main costs of reaching the solution can be calculated after defining the key resources, key activities and key partners. The revenue stream is defined as the way the income is made from each customer segment. The costs and revenue stream are defined for each actor in the value network.

Customer/market segments – These segments include the potential or targeted customers/market segments and the consumers interested in the provided data/information/application. This concept also deals with the type of relationship the company wants to create with its customers.

Key channels– Key channels include communication channels with partners/customers, delivery channels of the solution, and sales channels through which the service/application is delivered/made available to consumers.

Data structure – This concept deals with the uniform format of data, the separation of data between open data and private data, the management of the permissions to data, the management of the ownership of data, the pricing and charges for the use of data, data licenses and license management.

Business development – This concept involves a description of how to develop and protect a business, the definition and management of SLAs, the definition and management of license policies, charges for the use of the solution, the distribution of the revenues with all partners, and the ownership of the solution (IPRs).

IV. THE RESULTS OF INTERVIEWS

The main goal of company interviews was to collect the requirements of the ODE and to receive up-to-date information and views directly from Finnish industries. The theme interview format [33], [34] was selected because of the different backgrounds and contexts of the companies; it would have been difficult to define the exact questions that could be applicable to every interviewee. Each interviewee selected a role that represented his/her company's role in the ecosystem. The roles were defined based on the literature survey described in Section III-A. The interviewee inspected the themes from the viewpoint of this role. In the theme interview, there were no questions, but the themes were informally discussed within the scope of the sub-themes that were relevant to the interviewee.

We defined four main themes with sub-themes:

- Open data (sub-themes: the meaning of data, data as a competitive advantage, data privacy, data integration, data licenses, data sale and user groups): The purpose was to examine how the open data concept is understood and what kind of meaning and role it has in a company's business.
- Application (sub-themes: open data, data drivers, application domain, tool support, IPR, selling/marketing and users): The purpose was to inspect the open data concept from the application viewpoint, i.e., the key issues of open data based applications.
- Co-creation (sub-themes: application idea, partners, division of work, costs and profits, the distribution of profits and contract management): The purpose was to untangle how the cooperation between partners would be formed in the ODE.
- ODE-based business (sub-themes: motives, risks, requirements of the ecosystem, ensuring business and establishing cooperation): The purpose was to find out the opinions and thoughts about the ODE and to gather the requirements of the ecosystem.

Representatives of 11 Finnish companies participated in the interviews. The interviews were performed in Oulu, Helsinki and Tampere between June and August 2013. The companies were selected from different application domains, and interviewees were selected based on their knowledge about the business viewpoints of their companies. Table 1 depicts the backgrounds of the interviewed companies. As can be seen, the interviewed companies differed according to the

TABLE 2. The results of the "open data"-theme interviews.

	Application users	Data/service providers	Application developers	Infrastructure/tool providers
Meaning of data	Competitive advantage	Competitive advantage	Competitive advantage	Beginning of new ideas
Data as competitive advantages	• Several utilization domains • Providing information that competitors do not have • Selling knowledge • Integrating data • Enabling customer to achieve data	• Benefits for early adopters • (Almost) real-time data • Information for a company's core competence	• Rapidly changing and refined data • Market information • Knowledge • Predictions • Technological possibilities	Research data in an early phase about the potential demand for the product
Data privacy	• Private data: cannot be opened • Data utilization requires customers' approval	• The importance of customer analysis • Privacy management	• Private data cannot be shared • Open data are related to public sector	• Data are customers' data • Important how the data are used
Data integration	• Competitive benefit • Provides new information • Standards for content • A need for catalogues and schemas of open data, anonymous open data, and open data consultants		• Open data become private when integrating them with private data • Incompatible data formats • A need for a broker	
Data licenses	• Standards for the representation of open data • Reference to the data source • Data reliability verification	Licenses are required before data can be used	• Clear directives for data utilization • The correctness and permanency of data	• Customers are not willing to pay for information • Data licenses are restricted in contracts
Data sales	• Should consider only for processed data • Information about the customers cannot be sold	• B2B and B2C sales in an ecosystem • Collection of data causes costs	• A small fee for the data • SLA for ensuring the existence of open data sources	
User groups	Application developers, integrators, and end-users		More for hobbies	Research to find out the users' needs

company size, application domain and service type. In addition, a portion of the companies had more than one role in the ecosystem. The size of the companies is defined according to [35]: micro-enterprise <10, small enterprise <50, medium-size enterprise <250 and large enterprise >250 employees. The following subsections introduce the results of the theme interviews.

A. OPEN DATA THEME

Table 2 summarizes the results of the "Open data"-theme interviews.

Meaning of data – The importance of data was identified to be high in all companies, and open data were seen as a trend that companies should follow. Mostly, data were seen as a competitive advantage. Data were also seen as the beginning of new ideas; they help in innovation of new products and services.

Data as a competitive advantage – Information provides a significant competitive edge. The data providers thought that early adopters would benefit from open data. There is a need for consultants that market open data and help companies to understand how open data can improve their businesses. An application user stated that differences are made in services, i.e., what is made with information and how the services are created around information. The customer analysis has a great importance to several companies. Data mining from databases and from data flows and the combination of information will provide many benefits. Application developers did not consider raw data to be important; knowledge and refined information are most important for the business. Generally, almost real-time data was seen to provide a competitive edge. This information can improve safety and enable predictions. According to application developers, rapidly changing information and the refinement of information could create value for customers. In addition, providing information that competitors do not have, selling data-based knowledge and integrating data are seen as competitive benefits. The information can also be produced as a service to the customer. Thus, an application user company sees itself as an information enabler; the company's products/services assist the customer to produce information. A tool provider stated that the business potential and marketing of the products and services are required to be taken into account in the early phase. Market research is required to find out the potential demand for the product. Information for a core competence is achieved through scientific research. Predictions, market information, and technological possibilities also assist in this competition.

Data privacy – Private data are often the customer's own data or information collected about customers. The management of privacy is extremely important to the interviewed companies. For example, companies cannot typically give up or sell information about their customers. The statistics that are collected from the customers or from the customers' data can be delivered further only with the customers' approval. Open data will become private data when private elements are added to the open data set.

Data integration – Data integration is seen as a competitive benefit and provides new business possibilities. Open data can

TABLE 3. The results of the "application"-theme interviews.

	Application users	Data/service providers	Application developers	Infrastructure/tool providers
Open data	• Need for processed data • Need for intermediate data-processing layer	• Raw data • Need for real-time and global data • Need for reliable, secure and permanent data and data sources	• Standard APIs are required • Permanency of data sources • Changing data	
Data drivers	• Unify, analyze and visualize the data • Implementation should be done by third parties • Companies pay for the drivers used	• Implemented by integrators • Data adaptation for users • Need for data mining • Data are in a format that is required by customers • Ensuring data privacy	Handling of privacy	
Application domain		Applications for decision support	• Applications for operators and customers • Need for user-driven development	Not restricted
Tool support	The tools are created by third parties	Ready-made/standard solutions available	Need for standard tools	The tools are created by third parties
IPR		Ensure that the business is under own control	Need for clear IPR	The IPRs belong to a customer
Selling/ marketing	Small markets for applications in Finland	It is difficult to: • Find good applications • Find out what information is available	Applications can be provided as a service	
Users		• Achievement of new user groups • Usability testing; importance of UIs	Need for open data provided in standard formats	

provide additional value for other data sources. For example, the integration of open data with the information gathered by itself creates information that nobody offers at this moment. In addition, it should be possible to integrate one's own information with the advertisers' information. The application developers said that data integration should already be considered in the data-processing phase; the data should be in a form that can be integrated later. There should be schemas for open data. The produced content should follow the related standards. There is also a need for a data broker supporting the publishing of data and sharing of data.

Data licenses – Data processing and the collection of data causes costs, and thus, in many cases, it is not possible to provide data to customers for free. Application developers estimated that there could be a small fee for the used data; the producer will earn money from the data if it ensures the existence of the open data sources and provides some kind of SLA for the open data source. There should be standards for the representation of open data; at least, the government of different countries should obey standards. It should be clear how the data can be used in applications and services. The application users thought that there is an ethic viewpoint in the data licenses; it would be reasonable if the original source of the data would be referred. This assists in the estimation of the reliability of data: the application user could see that a reliable actor has provided the data and nobody has changed them or made their own conclusions as a result of the data. Application developers felt that it is unclear who has the responsibility for the correctness of open data and the permanency of APIs and data sources. According to a tool provider, customers often want everything to be free and do not want to pay for information. Data licenses and other payments are often defined or restricted in contracts. The data providers estimated that licensing rules depend on an application, i.e., how critical the application is. There will be different contracts for different data, e.g., for free data and for certified/audited data. Often, officials provide their own contracts that must be used.

Data sale – The data processing provides additional value for data. For example, processed data could be used for targeting advertisements for certain customers. Different shops and stores could be ready to pay for this kind of information. A data provider estimated that a company's sales in the ecosystem would come from B2B and B2C sales.

User groups – The data providers believed that there are several kinds of users, such as application developers, integrators and end-users for open data. These users' goal is to make all information available to the all actors. A tool provider revealed that the recognition of the user groups of a product/service is sometimes difficult. Discovering the problems and needs of users requires more user research. An application developer estimated that more applications will be required for hobbies in the future.

B. APPLICATION THEME

Table 3 summarizes the results of the "Application"-theme interviews.

Open data – Data providers' content is more or less raw data or information. Service providers can then produce refined information and knowledge from the content. In the future, the goal is to provide more open data for customers. Customers have a use for slowly changing data, but fresh and rapidly changing information is more important. Users can give feedback to data providers and thus improve the

correctness and quality of the data. The interviewees thought that it is more secure to use popular sources of open data in business-critical applications; the reliability and permanency of a popular open data source can be assumed to be better than a less-known open data source.

A data provider emphasized that due to operating a global business, it is not enough that certain (e.g., location-related) data are open in Finland; they should be open everywhere. When data are not open, certain services are only available in certain countries. Application developers emphasize that standard formats are required for open data. In addition, there should be a clear description of the API of an open data source. The data providers and application developers said that there should be a means to ensure the permanency of data sources and APIs and the reliability of data. The risk that data sources disappear can limit the usage of open data in applications. SLAs are needed for open data sources. Additionally, business should not be trusted in the hands of one data provider; there should always be a back-up plan.

Data drivers – Application users need processed data but cannot typically implement the processing itself. Data/service providers identify the need for data-processing services and for an intermediate data-processing layer. Data drivers were considered important, but third parties should implement these drivers because the companies find it difficult to implement the drivers by themselves. A driver should form and unify the data, analyze (or parse) the data and visualize the data. Drivers could be used, for example, for advertisement; the advertising companies could have some kind of link to the driver for which they would pay. The data must always be adapted for the user; they must be in a format that is needed in the customer's core business. The handling of users' privacy should be considered in the data processing of data drivers.

Application domain – Open data can make it possible to provide extra features to applications and will provide new opportunities for software integrators and application developers. The application domain would not be restricted to any specific domain. There could be applications for operators and applications for customers. Applications could also extend to existing big software systems by providing new features. Applications that support decision-making are needed in different application domains. For example, there is a need for applications that guide the user or provide a snapshot of the overall situation. Open data requires data-mining tools to extract the essential data from the data flows. Data security has an important role; there must be ways to recognize users and control the access to data.

Tool support – Tool support was considered self-evident. The tools and applications for processing are created by others. The tools should be standard, and they should have stable and compatible APIs. The challenge presented by these data is that often, situations are unique, and it is difficult to represent the data with the available tools.

IPR – Clear IPR is needed. It is important to make sure that the business is under its own control. In customer work, the IPRs belong to a customer.

Selling/marketing – Application users emphasized that there are small application markets in Finland. Several data providers stated that it is difficult to find out what kind of information is available, and it is difficult to find good applications from app stores. According to application developers, applications are often based on commercial agreements. Applications can be provided as a service or be sold to other operators (B2B). In addition, there are licensed applications, whereas a portion of these applications is a company's own property. Furthermore, applications are not always in the core, but there are markets for integrated solutions consisting of different kinds of devices and software systems.

Users – An application is often a way to provide a service to customers. The usability of applications is increasingly important in competition and easy-to-use and easy-to-learn user interfaces are needed. Thus, there is a great need for the usability testing of applications. In addition, application developers could tailor applications to different consumer groups of open data.

C. CO-CREATION THEME

Table 4 summarizes the results of the ''Co-creation''-theme interviews.

Application idea – The application idea can come from customers or from the company itself. In many cases, ready-made products or services can be applied, and standard APIs assist different stakeholders in integrating their solutions to SW systems.

Partners – Generally, the companies had quite compact and stabilized partner networks. Some companies contracted a great deal of consulting work, and some also use subcontractors in software development. In a project type of work, the customer sometimes selects the partners, and the company sometimes selects the partners. According to a tool provider, the selection of a partner was dependent on the case; the different partners are usually oriented to different technologies.

Division of work – Each actor must have a natural role in the ecosystem; otherwise, actions must be implemented by themselves. The division of work between partners must be strictly defined for each task. A tool provider stated that by cooperating with partners, a company can accept larger projects. Some companies saw themselves as independent actors in the value chain in which there was not much communication between different actors. Some companies saw themselves operating in a value network, as the information and money moved in many directions. The responsibilities should be clearly shared between actors; all actors do not do everything. In addition, clearly defined APIs, components, and the named responsible must be defined for each task.

A data provider company described itself as an actor in a data flow. All data are not necessarily delivered for the next actors in the data flow, but there can also be secondary flows in the data flows. It is important that customers participate in the early stage of the creation of application concepts. Pilot customers can provide feedback for development. One interviewee felt that the more partners there are, the more

TABLE 4. The results of the "co-creation"-theme interviews.

	Application users	Data/service providers	Application developers	Infrastructure/tool providers
Application idea		Importance of standards	The prioritization of application ideas/features	• The application idea comes from the customer • Styles for how to interview people for ideas
Partners	• Compact partner network • Stabilized partners • Consulting work contracted	• Several partners • An actor in a flow of data • Customer participation	Integrates software from sub-contractors	• Project work with customers and partners • Several stabilized partners • Partner selection based on technologies
Division of work	Operation in: • Value networks • Value chains	• Co-creation inside value network • Clear sharing of responsibilities		• Responsibilities are strictly defined • By cooperating, larger projects can be accepted
Costs and profits			Arrangement in each co-creation activity	
Distribution of profits		Mechanisms for agreements	A win-win situation for all stakeholders	
Contract management	Flexible contracts	A new consortium requires new contracts	Open-source solutions should be used in the development	

complicated the things are. However, when the company succeeds, it will be easy to find new partners.

Costs and profits – The earnings/benefits should be shared in the ratio of the work done. The sharing of tasks, costs and revenues is often difficult to agree upon beforehand. These specific assignments can be agreed upon later so that in each co-creation activity, an agreement is made with regard to these issues.

Distribution of profits – According to the data providers and application developers, cooperation does not work if an actor takes too large a share of the profits. It should be a win-win situation for all stakeholders. There should be fair mechanisms of the sharing of profits between partners.

Contract management – The contracts with partners are clear. Sometimes, the contracts must be flexible and to be able to be created quickly. An application user estimated that the contracts between partners will be functional only within each layer.

D. ODE-BASED BUSINESS THEME

Table 5 summarizes the results of the ''ODE-based business''-theme interviews.

Motives – All companies saw several motives and advantages of the ODE. The ODE can support the creation of ideas and the visualization of products and provides a new kind of basic function and new data-based content. Data can provide additional value for products, and the ecosystem enables the creation of more services/applications that do not compete against the company's own products. Thus, open data are seen as a way to improve competitiveness. The opening of data increases the utilization rate of data and makes the data available to a larger number of developers that create applications for special groups and for specific purposes.

Civil data improve the accuracy of the snapshot of the overall situation. However, there must be some kind of reward from produced data, and citizens must benefit from the data. The ecosystem enables an increased understanding of what could or should be done and with whom. The ecosystem also assists in achieving new partners and new customers and enables a company to serve its customers better. For example, the interviewees identified that the customers could use information more effectively in business. The ecosystem could also increase sales and facilitate sales and marketing efforts.

Risks – Open data can change the business environment, and companies also saw some risks in the ODE. Firstly, it is possible that a business would not be profitable in the ODE. Costs increase, but consumers do not use the services/products. Furthermore, there is a large amount of work in terms of the maintenance, support and marketing of the ecosystem. Secondly, because the situations in companies are changing continually, no company seeks to enter into very binding contracts. In this way, it is easy to break contracts. Thirdly, the concept of data drivers was seen as a risk: the application user stated that it would be more rational to use or acquire open data directly from a data provider than to be dependent on the providers of data drivers. There is also a great risk that the data driver service is not available when it is required. Fourthly, for open data, the identified risks included that the data could not be provided openly, the privacy rules change, and the data are so different that a single platform cannot manage all data. The data providers were worried about the quality of data and changes in the quality of data. The quality of data is not in their own hands. Furthermore, the easy access to data can bring new competitors. Open services can cause problems for chargeable services if another actor provides the same service for free. A tool provider was afraid that customers may not understand the concept and use of open data; the customers may be afraid that because the information is free, they must still pay for it according to the contract.

Requirements of an ecosystem – A business ecosystem cannot be created, but it should naturally emerge. There is no ecosystem if the ecosystem does not provide a role and

TABLE 5. The results of the "ODE-based business"-theme interviews.

	Application users	Data/service providers	Application developers	Infrastructure/tool providers
Motives	• New content, ideas and basic functions • Increased understanding in business • Easier service marketing and sales • Improved competitiveness • New customers	• Increased sales • More ecosystem users • Better offerings • Increased amount of application developers • Increased data utilization rate • Willingness of follow the open data trend	• Customers' demand for open data solutions • Improved brand • Available data that provide additional value to the products	• New customers • New partners
Risks	• ODE-based business is not profitable • Availability of data driver services • No binding contracts are made • The data cannot be provided as open	• The ecosystem requires a great amount of work or does not provide a role or benefits for its actors • Lack of open data operators • Low quality of data • New competitors	• Platforms do not work because data are so different • Business requires much work; 24/7	Customers do not understand the open data concept
Requirements of ecosystem	• Support in business know-how and finding new markets • Data validation • Definition of data/information creation • Standardization of data drivers	• Easy and clear joining of the ecosystem • Support for all members • Standard-based data and APIs • Operator services • Data validation	• A developer community for tools, APIs and frameworks • Trust that the framework works in real life	• Support for internationality, market research and contract making • Information about what data there are available
Ensuring business		• Standards • Business models • Replacing actors and solutions with others • Paying for the data quality		
Establishing cooperation	Will be settled practices for cooperation as the operation begins	It will be easier to get partners to join as the size of the ecosystem gets bigger		

benefits for all of its actors. If the ODE is established, it should be a proper investment from the beginning. Joining the ecosystem should be easy and clear. The ecosystem should provide proper support to all of its members and support the development of contracts between actors in the ecosystem. Often, companies are afraid that their ideas may be stolen. The contracts should clearly define the rights of each partner and how payment issues are dealt with. The growth rate of a business can be very fast in the ecosystem. A data provider estimated that there should be more operators of open data. Operator services are needed for a centralized payment mechanism and for a centralized mechanism to share profits inside the ecosystem. The ODE should support in terms of business know-how, finding new markets and assisting companies to see what kinds of products and services are in demand and what consumers really want. Often, Finnish companies let technologies lead the design. They use information but in a misguided way. The result is that they produce a complicated product for which there is no demand. The ODE should also consider internationality; cultural differences should be taken into account. The responsibilities of data verification should be defined in the ecosystem. It seems that there is a need for a data quality verification service in the ecosystem. The ecosystem should provide information about the available data. The ecosystem should also define how the data/information are created. There should be definitions and possibly standards of how to create measurable information, how to store it, etc. Local and unstandardized APIs will create risks for the whole ecosystem. The SLAs for data were also suggested. Furthermore, some form of standardization could facilitate or change the role of the data drivers. In this way, it would be easier to find another service that has the same functionality. There should be a developer community for tools, APIs and frameworks. In addition, there should be a mutual trust that development continues and that the framework works in real life.

Ensuring business – According to the data providers, to ensure business in the ODE, it is important to avoid monopolies. The network must be designed so that it is possible to replace the actors and used solutions with new ones. The usage of standard APIs helps in this regard. New kinds of business models are required when acting in the ecosystem, and the work performed must have value in the business. In addition, by knowing the customer's businesses well, services can be tailored for customers and their information systems. A data provider would be willing to pay to ensure the quality of data sources. Some interviewees saw that there could be an opportunity for the ecosystem and business opportunities for companies that produce processed information. There will also be a demand for data integration and analytics.

Establishing cooperation – As cooperation in an ecosystem begins, there will be settled practices. The size of the ecosystem is important; first, it will be small, but the bigger it gets, the easier it will be to find partners to join it. The larger the ecosystem is, the more interesting it is to application developers.

E. A REFINED OUTLINE OF THE ODE

We modified the first outline of the ODE based on the requirements obtained in the theme interviews. The theme interviews revealed that value networks are formed dynamically among actors to reach a solution. Thus, the ODE should support fast networking; once a demand has been identified, it should be possible to quickly establish a value network, ensuring that different actors can find needed partners, and agree on the division of work, distribution of costs and distribution of profits. Several value networks co-exist inside the ecosystem; they are formed and dissolved according to the situation at hand.

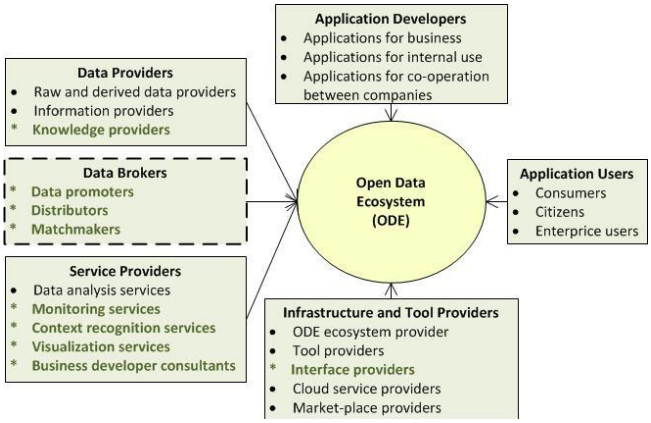

FIGURE 3. **The roles of actors in the refined outline of the ODE.**

The interviewees identified new kinds of actors and services that are needed in the ODE. Fig. 3 introduces the refined outline of the ODE through the actors' roles. A new actor group "Data brokers" is depicted in the box with dashed lines, and new roles within the previously defined actor groups are illustrated with the * mark. The following subsections discuss the identified ecosystem actors, roles and services.

1) NEW ACTORS AND ROLES IDENTIFIED

An entirely new actor class, Data Brokers, was identified. In addition, new roles were identified in the Service Provider, Data Provider and Infrastructure and Tool Provider classes.

Data Brokers–The broker actor includes the following roles:

- Data promoters: A data promoter finds out data and advertises them to the actors (e.g., application developers) of a certain domain or to matchmakers. It also maintains "a list" of available data in the ecosystem and the quality of data, price, applied licenses, etc. The promoter receives a fee from the data providers for the advertisement of their data and from end-users or consumers (when using the data directly) to mediate information about the data source.
- Distributors: These provide the communication channels and distribution channels of data and applications. A promoter receives a fee from the channel user depending on the channel and the usage type (e.g., communication, data transfer).
- Matchmakers: The ecosystem should uncover the demand for data and the quality of the demanded data from end-users and match the demand with the best available data source (received from the data promoter) and data transfer service. It should also inform if there is a demand for not-yet-available data. This matchmaking also enables the automation of data matching and selection in the digital environment.

Service Providers – There is a need for the following new roles of service providers:

- Monitoring service providers: These services are capable of monitoring data in physical and logical environments and detecting any changes in data. Based on monitoring services, the actors can create applications that can automatically or semi-automatically react to the changes (in real time) in the environment and, for example, notify the application users about changes in the environment.
- Context recognition service providers: Monitoring services require services that are capable of recognizing contexts from data streams.
- Visualization service providers: These services are capable of visualizing the essential information in an understandable way to different user groups.
- Business developer consultants: An open data consultant that assists companies in understanding the possibilities of open data in a company's business. It knows what kind of data a company owns and identifies customer needs and the utilization areas of the available data.

Data Providers – There is a need for knowledge providers that offer expert services based on knowledge refined from information.

Infrastructure and Tool Providers – There is a need for interface providers that develop user interfaces for consuming data with different kinds of devices.

2) IDENTIFIED NEW SERVICES

The interviews revealed that the ODE should provide services for:

- Contract-making: The ecosystem should assist the actors in developing contracts with each other. This contract-making should be clear and rapid.
- Finding partners: The ecosystem should maintain a catalogue of the ecosystem's actors and the expertise and reputation of each.
- Finding services: The ecosystem should assist in finding applicable services and candidate services.
- Finding information: The ecosystem should be aware of the available open data/information and assist the actors in finding relevant information.
- Finding markets: The ecosystem should have a means of market research and of finding new markets and customers.

- Data validation, and the definition and standardization of data and interfaces: The ecosystem should define and describe the used syntax and semantics of open data.
- Business models: The ecosystem should assist in defining the appropriate business model for the actors of the ecosystem.

V. DISCUSSION

The results show that an open data based business model can bring both direct and indirect benefits. In open data based business, data flows are created between data producers and data consumers. However, it seems that in many cases, a cash flow between data producers and data consumers is not created, but an open data business can indirectly benefit the business of the producer of the data. One goal can be outsourcing; there could be data providers that provide data for free and hope that there will be actors that develop useful applications and services around the data and will therefore support the business of the data provider. Another option is to sell data with a pay-per-use model or achieve profits in data-based application usage. In addition, a need for free open data and for licensed/premium open data can be identified. Openness is the key work in business models, as an open data based business moves ahead from the proprietary software business. The transformation to the cloud-based business model should also be considered. The cloud-based business model provides solutions to scalability, capacity and interoperability problems and enables customer-side characteristics, such as pay-per-use pricing, ubiquitous access and on-demand availability.

The selection of companies and interviewees affects the results of this research. Thus, the represented results/requirements cannot be generalized for all kinds of companies. For example, the requirements of the ecosystem could have been different for companies whose core business does not involve data/information. In addition, the selection software professionals instead of business professionals could have resulted in more technical outcomes of the interviews. The interviews were performed in Finnish companies. Thus, the results of the interviews cannot be directly globally generalized. However, because several interviewed companies conduct international business, it can be assumed that the results may also be valid in countries other than Finland.

The rest of this section summarizes the results of theme interviews from the perspective of the core elements, open data, services and applications. The following subsections discuss the most important challenges and opportunities that relate to these three perspectives, how the business model elements are supported in the ODE, and the feasibility of the ODE.

A. CHALLENGES AND OPPORTUNITIES OF OPEN DATA

Heterogeneous open data, non-standard APIs and varying licensing conditions complicate the use of open data sources in commercial applications and services. Several data utilization domains were identified in the interviews. In addition, several business opportunities were found, such as providing information that competitors do not have, selling data-based knowledge, integrating data and data mining. There is an increasing need for raw data, refined information, knowledge, rapidly changing information, predictions and market information. The reliability of data and data sources were seen as important in the interviews. Free data are often assumed to have low quality. Furthermore, the permanency of free data sources is considered uncertain. Contracts (e.g., SLAs) were seen as a way of ensuring the existence and quality of open data sources. In addition, there are business opportunities for data promoters validating the data before advertising them further to other ecosystem actors.

It seems that the companies currently do not have much knowledge about open data and the license conditions of open data sources: Firstly, the open data sources should be marketed for companies. For example, public and well-known catalogues are needed for open data. Secondly, the companies should identify the benefits of open data and recognize how open data can be used in their business in the future. There is a need for matchmakers that know what kinds of open data are available and also know what kinds of data company own and identify customer needs and the utilization areas of the available data.

Open data are now often available via non-standard APIs. As a result, there are strong dependencies between an application/service and open data sources and are difficult to create applications/services that work in different regions and in different countries. A standard way to describe data sources and APIs is required as well as a uniform format for the data. Because a large amount of data is private (e.g., the customer's own data, or data about individual persons), the management of data privacy is important. In addition, varying licensing conditions complicate the use of open data sources.

B. CHALLENGES AND OPPORTUNITIES OF THE SERVICES OF OPEN DATA

There is a great need for data processing, analysis, and integration services. The data processing provides additional value for data and the integration of open data with the private data creates information that nobody offers at this moment. Data mining from databases and data flows and the combination of information will provide many benefits. Generally, almost real-time data was seen to provide a competitive edge. The services of real-time data can improve safety and enable predictions. The huge amount of open data and the real-time property brings a challenge to data mining and semantics handling.

There are several technical/non-technical challenges related to the use of open data in business and in internal use. The data integration should be considered in the data-processing phase. The data processing is costly; there is a clear need for the intermediate layer of data processing. In addition, optimizing algorithms are required for the analysis of open data. An observer agent is needed to obtain relevant, new and accurate data for the service

provider. The utilization of silent data requires consideration of how the silent data can be extracted and opened. In decision-making, the exploitation of open data with internal information could assist in process improvement and automate decision-making. Decision-makers need non-numeric information and essential information that must be extracted often from larger data sets. The information should be finally visualized for decision-makers. In the customer interface, personalized digital services can be provided for open data. The challenge remains of how to exploit open data with customer-specific data.

C. CHALLENGES AND OPPORTUNITIES OF THE APPLICATIONS OF OPEN DATA

Applications can be developed for sale, for a company's internal use and for cooperation with partners. The developers of businesses receive incomes from application sale and usage. Developers for an organization's internal use create applications for the organization itself, for certain user groups, or to be used inside organizations, obtaining more efficient productivity and business. Developers for cooperation between partners develop applications to improve the cooperation between partners, obtaining more efficient cooperation with business partners.

The open data enable the creation of new features of the applications. In addition, the application can extend existing software systems. For example, there is a need for decision support and guidance applications. In addition, open data provide new ideas and totally new kinds of user-innovated applications. However, at the same time, the usability of applications must be ensured. Market research is required to find out users' needs. Thus, it is important to collect ideas for applications and feedback from the users and enable the co-creation of applications with different partners. There are small markets for applications in Finland. Thus, it should be possible to adapt applications for different regions and for different countries. For the cross-country open data applications, the standard APIs and globally opened data are prerequisites. A validation of the quality of data was required as well as assurance of the permanency of data sources.

D. HOW BUSINESS MODEL ELEMENTS ARE SUPPORTED IN THE ODE

The ODE supports business model elements as follows:

Value proposition – The ODE services provide assistance in finding what data/information/service is available for innovation and application development.

Key partners – The ODE provide assistance in finding partners, e.g., using a catalogue of classified, registered ecosystem actors. The data promoter ensures the trustworthiness of data providers in the ecosystem.

Co-creation – The ODE provide assistance in terms of clear, flexible and rapid contract-making with partners, e.g., with ready-made templates. These contracts define the responsibilities and earnings of each partner.

Key channels – Distributors provide the communication channels and distribution channels for data and applications. Network operators provide mobile connectivity. A marketplace enables the sales of data driver services and applications.

Cost structure and Revenue stream – The ODE provide assistance in making cost-sharing arrangements in each co-creation activity and mechanisms for agreements. Arrangements for profits-sharing should be made with regard to the ratio of work, enabling a win-win situation for all stakeholders.

Customer/market segments – The ODE services provide a means of market research and of finding new markets and customers, assistance in identifying different user groups, and assistance in conducting user studies, which can be used to find out the needs of certain user groups.

Data structure – Standard formats of open data define the syntax and semantics of open data and interfaces applicable to all data providers. The use of standard data formats guarantees that the data driver services produce output data in an interoperable way. Data reliability is supported by the validation of data through data promoters, which maintain "a list" of the quality of available data. Licenses provide defined and applicable license policies used in the ODE. Licenses are required before data can be used. Therefore, it should be clear to users how the data can be utilized. The ODE defines the practices of how privacy issues are handled and how to protect data from unauthorized users. Data fees specify the defined and agreed-upon fees for the data.

Business development – The ODE provide assistance in market research to find out the demand for data and data-based solutions and to provide information about new data for new ideas, content, products, etc. Matchmakers assist in matching the demand with the best available data source. Consulting services assist in understanding the possibilities for open data in a company's business. The ODE defines clear IPRs for the data, services and applications and also provides information about alternative data sources and services if the selected ones become unreliable or unavailable. The ODE also assists in defining the appropriate business model by providing a template and guidelines for elements of the business model.

E. FEASIBILITY OF THE ODE

In addition to technical challenges, we believe that the most difficult challenge is to obtain enough actors for the ODE. Firstly, there must be enough actors that see the benefits of the ODE and are motivated to actively participate in the development of the ecosystem. Secondly, the ODE should naturally emerge. All of the ecosystem actors are equal; monopolies should not exist. Each ecosystem actor should identify its role in the ecosystem and create the company's business model accordingly. The joining of the ecosystem should be fast and easy through registration. After registration, the actor has access to all data in the ecosystem. The actors in the ecosystem cooperate in value networks, which

are formed dynamically and rapidly to respond to a certain demand. Several networks may emerge simultaneously inside the ecosystem. In the beginning, the ecosystem is assumed to be small, but the bigger it gets, the more willing the companies are to join it.

There are at least two options for the development of the ecosystem: development from scratch or development by extending the existing technical solutions, such as CKAN, enabling straightforward service and application development, thus minimizing the amount of work needed to establish an ecosystem. This development can be started by a company, community or organization by itself, with partners or with other actors, for example, in the same domain. There are also two options for the content of the ecosystem: 1) the ecosystem emerges around a certain domain, being domain-dependent, or 2) there is a ''universal'' ecosystem applicable to all interested parties from different domains, being therefore domain-independent. It is obvious that there is a great amount of work needed to maintain the ecosystem. It must be clear from the beginning who is responsible for the maintenance, support and marketing of the ecosystem. There can be, for example, a separate ecosystem provider, or one of the ecosystem actors (e.g., a large data provider) takes on the role of an ecosystem maintainer.

VI. CONCLUSION

This paper defines an initial outline for the open data ecosystem based on a literature survey and describes the requirements of the ODE based on information collected by interviews performed in 11 Finnish companies. The interviews revealed the state of practice in data-based businesses as well as the future visions of the industry on open data based business. The results of the interviews assisted in refining the refined outline of the ODE and understanding the challenges that data-based business still embodies.

The study revealed several requirements of the open data ecosystem, including the roles of the actors and the required services that must be defined and implemented while establishing the ecosystem. The interviews helped to identify several motives for joining the open data ecosystem. However, there are still some obstacles and risks that have to be taken into account in the development of the ODE. A great deal of work is required for implementing the defined requirements and overcoming the identified obstacles in order to enable profitable business for the actors of the ecosystem. According to the interviews, the interest in open data based business is high. Thus, the ODE could provide great benefits for the involved actors and their businesses through open data and the services and applications around them.

REFERENCES

[1] S. R. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, *DBpedia: A Nucleus for a Web of Open Data* (Lecture Notes in Computer Science). Berlin, Germany: Springer-Verlag, 2007, pp. 722–735.

[2] A. Immonen, M. Palviainen, and E. Ovaska, ''Towards open data based business: Survey on usage of open data in digital services,'' *Int. J. Res. Bus. Technol.*, vol. 4, no. 1, pp. 286–295, 2014.

[3] M. Iansiti and R. Levien, *Creating Value in Your Business Ecosystem*. Boston, MA, USA: Harvard Bus. School Press, Mar. 2004.

[4] A. Poikola, P. Kola, and K. A. Hintikka. (2011). *Public Data—An Introduction to Opening Information Resources, Ministry of Transport and Communications* [Online]. Available: http://www.scribd.com/doc/57392397/Public-Data

[5] G. Kuk and T. Davies, ''The roles of agency and artifacts in assembling open data complementarities,'' in *Proc. 32nd Int. Conf. Inf. Syst.*, Shanghai, China, Dec. 2011, pp. 1–5.

[6] A. Latif, A. Saeed, P. Hoefler, and A. Stocker, ''The linked data value chain: A lightweight model for business engineers,'' in *Proc. 5th Int. Conf. Semantic Syst.*, 2009, pp. 568–575.

[7] A. Osterwalder, C. Parent, and Y. Pigneur, ''Setting up an ontology of business models,'' in *Proc. 16th Int. Conf. Adv. Inf. Syst. Eng. Workshops*, 2004, pp. 319–324.

[8] J. Perr, P. Sullivan, and M. M. Appleyard, *Open for Business: Emerging Business Models for Open Source Software Companies, Working Paper, Lab2Market*. Portland, OR, USA: Portland State Univ., 2006.

[9] C. Baden-Fuller and M. S. Morgan, ''Business models as models,'' *Long Range Planning*, vol. 43, no. 1, pp. 156–171, 2010.

[10] D. J. Teece, ''Business models, business strategy, and innovation,'' *Long Range Planning*, vol. 43, nos. 2–3, pp. 172–194, 2010.

[11] Y. Tammisto and J. Lindman, ''Open data business models,'' in *Proc. 34th Inf. Syst. Seminar*, Turku, Finland, 2011, pp. 762–777.

[12] Y. Chen, J. Kreulen, M. Campbell, and C. Abrams, ''Analytics ecosystem transformation: A force for business model innovation,'' in *Proc. Annu. SRII Global Conf.*, San Jose, CA, USA, 2011, pp. 11–20.

[13] C. M. L. Chan, ''From open data to open data innovation strategies: Creating E-services using open government data,'' in *Proc. 46th HICSS*, Wailea, HI, USA, 2013, pp. 1890–1899.

[14] H. W. Chesbrough and M. M. Appleyard, ''Open innovation and strategy,'' *California Manag. Rev.*, vol. 50, pp. 57–76, Nov. 2007.

[15] H. Liu, ''Big data drives cloud adoption in enterprise,'' *IEEE Internet Comput.*, vol. 17, no. 4, pp. 68–71, Jul./Aug. 2013.

[16] B. Narasimhan and R. Nichols, ''State of cloud applications and platforms: The cloud adopters view,'' *Computer*, vol. 3, pp. 24–28, Mar. 2011.

[17] J. Myllykoski and P. Ahokangas, *Transformation Towards a Cloud Business Model, Discussion*, New York, NY, USA: Commun. Cloud Softw., Apr. 2010.

[18] G. Tao, L. Yi-Jun, G. Jing, and G. Long, ''Research on the economic features and pricing of digital products,'' in *Proc. ICMSE*, 2006, pp. 152–156.

[19] R. Harmon, D. Raffo, and S. Faulk, ''Value-based pricing for new software products: Strategy insights for developers,'' in *Proc. Portland Int. Conf. Manag. Eng. Technol.*, 2004, pp. 1–24.

[20] A. Sundararajan, ''Nonlinear pricing of information goods,'' *Manag. Sci.*, vol. 50, no. 12, pp. 1660–1673, 2004.

[21] V. Allee, ''Value network analysis and value conversion of tangible and intangible assets,'' *J. Intell. Capital*, vol. 9, no. 1, pp. 5–24, 2008.

[22] L. Pastor, ''Portfolio selection and asset pricing models,'' *J. Finance*, vol. 1, pp. 179–233, Feb. 2000.

[23] V. Abhishek, I. A. Kash, and P. Key, ''Fixed and market pricing for cloud services,'' in *Proc. IEEE Conf. Comput. Commun. Workshops*, Orlando, FL, USA, Jan. 2012, pp. 157–162.

[24] C. Weinhardt, A. Anandasivam, B. Blau, and J. Stosser, ''Business models in the service world,'' *IT Prof.*, vol. 11, no. 2, pp. 28–33, 2009.

[25] A. Gaudeul, ''Software marketing on the Internet: The use of samples and repositories,'' *Econ. Innov. New Technol.*, vol. 19, no. 3, pp. 259–281, 2010.

[26] (2013, Nov. 21). *Creative Commons, Baseline rights, Creative Commons* [Online]. Available: http://wiki.creativecommons.org/Baseline_Rights

[27] (2013, Nov. 21). *Open Data Institute, What is Open Data? Online Guide. Open Data Institute* [Online]. Available: http://theodi.org/guides/what-open-data

[28] M. Palviainen, J. Kuusijärvi, and E. Ovaska, ''Semi-automatic end-user programming approach for smart space application development,'' *Pervasive Mobile Comput.*, vol. 2, pp. 1–32, May 2013.

[29] M. Palviainen, J. Kuusijärvi, and E. Ovaska, ''Framework for end-user programming of cross-smart space applications,'' *Sensors*, vol. 12, no. 11, pp. 14442–14466, 2012.

[30] M. Palviainen, J. Kuusijärvi, and E. Ovaska, ''Architecture for end-user programming of cross-smart space applications,'' in *Proc. 4rd Int. Workshop Sensor Netw. Ambient Intell.*, Lugano, Switzerland, Mar. 2012, pp. 823–824.

[31] H. Chesbrough and R. S. Rosenbloom, "The role of the business model in capturing value from innovation: Evidence from Xerox Corporation's technology spin-off companies," *Ind. Corporate Change*, vol. 11, no. 3, pp. 529–555, 2002.
[32] C. Zott and R. Amit, "Business model design: An activity system perspective," *Long Range Planning*, vol. 43, nos. 2–3, pp. 216–226, 2010.
[33] C. Livesey. (2013, Nov. 21). *Sociological Research Skills: Focused (Semi-Structured) Interviews, Sociology Central* [Online]. Available at: http://www.sociology.org.uk/methfi.pdf
[34] S. Hirsjärvi and H. Hurme, *Tutkimushaastattelu: Teemahaastattelun teoria ja käytäntö (in finnish)*. Helsinki, Finland: Yliopistopaino, Univ., 2001.
[35] European Union Commission, "Commission recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises," *Off. J. Eur. Union*, vol. 46, pp. 36–41, May 2003.

ANNE IMMONEN received the M.Sc. degree from the University of Oulu in 2002. Since 2002, she has been a Research Scientist with the VTT Technical Research Centre of Finland. Her main research interests include reliable service engineering, quality modeling and the analysis of composite services, and quality ontologies. Her current research interests include business models and ecosystems in data-based application development.

MARKO PALVIAINEN received the M.Sc. degree from the Lappeenranta University of Technology and the Ph.D. degree in computer science from the Tampere University of Technology in 1998 and 2007, respectively. Since 1999, he has been a Research Scientist with the VTT Technical Research Centre of Finland. His current areas of research interests include mobile applications and application development methods, ontology-driven software engineering, and the evaluation methods of parallel applications.

EILA OVASKA received the Ph.D. degree from the University of Oulu in 2000. Before 2000, she was a Software Engineer, a Senior Research Scientist, and the Leader with the Software Architectures Group, VTT Technical Research Centre of Finland. Since 2001, she has been a Professor with VTT and an Adjunct Professor with the University of Oulu. Her current area of interest is service architectures, particularly in self-adaptive digital service systems and services. She has acted as a workshop and conference organizer and as a reviewer for scientific journals and conferences. She has co-authored over 130 scientific publications. She is a member of the IEEE Computer Science.