Mind the Gap: Exploring Stakeholders’ Value with Open Data Assessment

Anders Hjalmarsson
Viktoria Swedish ICT &
University of Borås, Sweden
anders.hjalmarsson@viktoria.se

Niklas Johansson
Viktoria Swedish ICT, Sweden
niklas.johansson@viktoria.se

Daniel Rudmark
Viktoria Swedish ICT &
University of Borås, Sweden
daniel.rudmark@viktoria.se

Abstract
There is an ongoing movement in society to provide open data as an enabler for innovation and growth. As more data is passed on beyond organizational borders to trigger third party development of services, the expectations on what could be developed using open data increase. Developers have been observed to express frustration over different challenges in the wake of open data provision, such as localization issues, lack of quality in data etc. Knowledge and tools are missing to assess available open data. This absence risks ensuring continual improvement of open data markets and the management of hampering gaps between developers’ requests of data and data providers’ provision of data. This paper explores the benefits with open data assessment based on an exploratory single case research design.

1. Introduction

Open data service development is an increasing practice that calls for a greater need of, and access to, open data [1, 2, 3]. By granting public access to organizational data (whose purpose was originally internal usage), also labeled open architectures [4, 5], substantial societal value can be created. However, while there are currently high hopes that open data may serve as a foundation for creating increased economic growth in the digital service development sector [6], open data development has so far largely failed to live up to these expectations. Previous research studying the barriers to open data development has found that data owners are struggling with the publication of data. Since such publishing is an emerging practice, much uncertainty surrounds it [7].

This uncertainty involves several aspects of data publishing such as privacy, legal issues and technical formats, but also how to prioritize publication of datasets with higher value creation potential [1, 2, 3, 7, 8]. Moreover, as the distributed nature of open data development caters for few natural feedback mechanisms from service developers, data owners have at best limited information on the value and usage of published datasets [7].

Developers on the other hand have been noted to express frustration over how to find and use data. E.g. for an external innovator it is often difficult to find a particular dataset, since it is not obvious which public body owns and potentially could publish it [8]. However, even if the dataset is found, uncertainties concerning data quality, license restrictions and sustainability of the data delivery are hindering open data service development [7, 8, 9].

This paper addresses these uncertainties. While provision of open data changes the development processes of novel digital services, only a limited number of guidelines have emerged to support such development. Based on results from exploratory research, it explores the value that open data assessment brings to stakeholders who promotes, provides input to or performs open data development.

Our aim is to contribute to the on-going research into how sustainable open data markets, places where open data is supplied and consumed [10], are established, and specifically in this paper to analyze how can different open data stakeholders benefit from performing systematic open data assessment? We believe that systematic open data assessment is a key practice that will support actors to mind and manage the gap between available and sought data.

Next the theoretical framework is presented. This section is followed by a description of the case and the research method (sections 3 and 4). In section 5 challenges perceived by involved participants leading to the design of a systematic model for open data assessment is presented. The model is introduced in section 6, and practical and theoretical implications of the research are discussed in section 7. The paper ends with conclusions and future research.

2. Theoretical framework

The theoretical framework involves the notions of open data, open data markets, open service development, stakeholders and assessment.
2.1. Open data and open data markets

‘Open data’ in this study refers to internal data passed on beyond an organizational border, also labeled open architectures [4], often made available through open Application Programming Interfaces in machine readable formats [4, 5]. Open data with this definition is frequently intended to be used by external innovators in the development of novel digital services.

In a report from the Swedish Governmental Agency for Innovation Systems [11], four different motives are presented for why open data should be provided by governmental agencies and organizations: 1) Promote openness and transparency; 2) Stimulate innovation and growth; 3) Promote public involvement, citizen participation and quality assurance; 4) Enable data sharing between governmental agencies and organizations.

In Lindman et al. [1] a research agenda is introduced for open data services. The agenda is based on the knowledge gap of how to construct sustainable open data markets, namely a marketplace populated by interdependent stakeholders who supply and consume open data and digital services derived from this market [10].

In comparison to the motives put forward by Vinnova [11], Lindman et al.’s [1], research agenda primarily relates to the second motive for provision of open data – i.e. open data as stimulator of innovation and growth. This as the research agenda focuses on the innovation of open data services. However it could be argued that the complimentary three motives could be viewed as important rationales in establishing a sustainable open data market, as these provide arguments for why especially governmental open data providers should provide open data to a market, making it viable.

2.2. Open data development and stakeholders

In recent years, open data service development has been propelled through the provision of open data (e.g. traffic and environmental data) on open data markets. The rationale for distributing open data is to attract outside innovators to develop data services that goes beyond what existing services provide [3], increasing the value generated to end users or customers [1, 10]. Several challenges to open data service development have emerged which threatens the sustainability of open data markets: make localization of particular datasets easy [8], improve data quality, reduce license restrictions and ensure sustainability of data delivery [7, 8, 9].

Open data service development involves different stakeholders. Organizations, labelled as data providers [1], pass on data beyond their organizational borders, and often invite developers, also known as third-party developers [12], to pursue open data service development, driven either by non-profit grounds [3] or business models [13].

Lindman et al. [1] introduces data providers, developers, customers as key actors (i.e. stakeholders) in open data markets. We argue that a fourth important stakeholder is emerging, and that is the open data broker. As organizations adopt a distributed way of pursuing innovation by providing open data, it also means that they simultaneously lose significant control of the direction of the innovation work performed [14]. In order to retain some control of the distributed process and at the same time attract third party developers to use open data sources, open data brokers have emerged on the open data market. The open data broker, e.g. an organizer of a digital innovation contest [14] or a commercial or non-commercial data hub provider, brings together a vast number of data sources to organize transactions between the developers and data providers (e.g. http://data.gov.uk/ or http://info24.eu/).

<table>
<thead>
<tr>
<th>Data provider</th>
<th>Motive to be on open data market</th>
<th>Function on the open data market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulate innovation and growth</td>
<td>Provides open data</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Third-party developers</th>
<th>Innovate open data services based on commercial/non-commercial grounds</th>
<th>Build open data services</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Open data brokers</th>
<th>Increase usage of open data based on commercial or non-commercial grounds</th>
<th>Arrange links between open data and third-party developers</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Customers</th>
<th>Resolve needs in their everyday situation by using open data services</th>
<th>End-users of open data services</th>
</tr>
</thead>
</table>

Table 1: Key stakeholders’ motives and functions on an open data market

In table 1 we summarize the motives and functions that different stakeholders have on the open data market. In this paper we explore stakeholders’ value for using systematic open data assessment. This focus means that we have omitted customers in our study, as they primarily are interested in the end result from open data service development. In order to explore the benefits of open data assessment we instead focus on data providers, third-party developers and open data brokers.
2.3. Assessment

Assessment within the field of business process management is an integrated part of investigating process model maturity [15]; that is, either the assessment of the process maturity in an organization [16] or the maturity that the organization has in performing process modelling and managing the business through the outcome of this practice [17]. Assessment is in this setting performed using a set of criteria that is used by an assessor to create organizational awareness of maturity level that the organization has, given a pre-defined scale [15].

Transferred into the context of open data service development, there are both similarities and differences. One difference is that assessment in business process management often is an intra-organizational practice; that is, it is performed by the organization on the organization for the internal use of the organization. This is the case also if a data provider does such an assessment on the data it provides, whereas it is not the case if open data assessment is performed by a third-party developer or by an open data broker. In such cases it is an external stakeholder that performs the assessment on an external organization in order to create a base to make a decision about future actions. Consequently, it is more of an audit.

In order to explore the potential function with open data assessment, different stakeholder benefits of such assessments has to be explored. The unit of analysis in this study is therefore to explore stakeholders’ value with open data assessment and investigate how that assessment can be performed as a base for addressing challenges hampering open data service development [7, 8, 9].

3. Case description

Systematic open data assessment emerged as an important task in the early stages of the case that is in focus in this paper. The case is a research and innovation project that aims to improve personal transport and accessibility in rural areas in the North Sea Europe region through the development of, and research into, innovative transport and communication concepts. Contributing to the Digital Agenda of Europe 2020, the project focuses on the development and use of pilot services, developed by open data, to boost efficient, environmental- and user-friendly transport concepts in order to reverse the spiral of decline in remote areas. There is also an interest within the project to investigate the possibility to transfer digital services from one to region to others in order to catalyze the dissemination of results to other regions with the same needs as the one in which the development had its origin.

The project started in the spring of 2012 and will be completed in the first quarter of 2015. Five regions in the northern part of Europe are involved in the project and collaborate to develop digital service pilots for selected target groups in all participating regions, enhancing their access to public transportation. In addition to research and development tasks, pilot actions evaluating the services will be performed in the project. The project is organized in work packages that involve 16 partners (research, development, regional) from the five regions.

In order to gain knowledge about the needs in each region and develop pilot services that match these needs, “toolboxes” have been designed to assist in different tasks within the work packages. The aim of these toolboxes has been to provide a set of comprehensive and straightforward work models to be used by the partners. In total, four toolboxes were developed in the project: “assessment and identification of the transport needs of various target groups in remote areas”; “service innovation workshops”; “open data assessment model”; and “business innovation workshops”.

In this paper we will focus on the open data assessment model and its value to different stakeholders on the open data market. It was developed to support the identification and assessment of open data sources in the regions. This in order to prepare the transformation of service ideas, identified in other work packages, into service pilots improving personal transport and accessibility in rural areas in the North Sea Europe region.

4. Research process and method

Our study aims to explore what value open data assessment brings to open data stakeholders. As discussed in the introduction, there is currently an ongoing movement in providing open data; on the other hand there is at the same time limited research performed regarding guidelines to support open data service development. The aim with this study is to explore assessment of the data provided and the reasons (benefits, values) for why, how and when different open data stakeholders should perform such an assessment.

To address this knowledge gap, we selected the project above described as an empirical base for an exploratory single case study to answer our research question. The selection was based on access to a critical case that enables us to explore the research question. The value of a single case is that it allows
the researcher to investigate a phenomenon in depth, especially where research and theory are at their early formative stages [18, 19]. Given the nature of the research problem, we consequently found a single case study with an exploratory research design an appropriate way to address the research question [20]. We envisage that the results of this study will constitute a base that could be used to perform future confirmatory research steps involving additional cases (c.f. section 8).

4.1. Phase 1 – case study setup

The research process was triggered by the discovery of a need within the case to assess the open data sources available in the participating regions. The underlying aim was thus to understand if the data available in the regions were a sufficient base to use for developing the service ideas that emerged from other work packages within the project. This discovered need on the project level initiated on the theoretical level the design of the research question and consequently also the development of the theoretical framework to be used in the research process: 1) as a basis for an interview protocol to be used in interviews during data collection, and 2) as a basis to explore the unit of analysis by analyzing the data collected [20].

4.2. Phase 2 – data collection and analysis

The discovered need of assessment within the case, as well as the defined research question, initiated then on the project level the second phase in the research process. As a first activity, five interviews were held with representatives from the regional partners, in order to understand the challenges that they were facing regarding the open data situation in the different regions. The interviews were guided by an interview protocol consisting of identified challenges for open data service development from the theoretical framework together with questions to determine the capability to perform systematic analysis. Four interviews were made face to face and one was made using Skype. The interviews lasted 45 to 90 minutes, and was audio recorded and transcribed to facilitate a qualitative analysis of perceived challenges. Data analyses were performed following Miles & Huberman’s [21] advice to 1) reduce, focus and compare findings through continual probing of collected and transcribed data, and 2) display data using tables and models. This process was triggered by the analysis of the transcribed interviews, however continued throughout the case study then fuelled by observations and complementing document studies (e.g. reports from the projects) [20, 21].

Besides providing research data, the interviews raised a joint request from the involved regional partners to develop an assessment tool with the aim to increase organizational awareness of 1) what type of open data sources were available within a region, and 2) in detail assess the nature of the data sources available. As a second step on project level a model for open data assessment was developed and introduced to the participants.

4.3. Phase 3 – exploring stakeholders’ value

Based on experiences from the project level in designing and also using the open data assessment model, the case study about stakeholders’ value moved to the theoretical level in research phase 3. The purpose of this phase was to probe the unit of analysis and consequently explore the benefits open data stakeholders receive when open data assessment is performed.

5. Perceived challenges when assessing open data sources

Within the case a decision was made to systematically identify and assess the open data sources in each region. This assessment was performed by the regional partners acting as open data brokers. Their overall task in this role was to elicit service ideas appropriate for the project and match these ideas with third-party developers and open data available in the region. These teams had the task to develop service pilots using the open data available.
In order to understand the challenges that the regional partners had to perform open data assessment, research data was collected using participatory observation and interviews during the service innovation workshops that were organized in the fall of 2012. These were carried out in four of the five regions; the fifth interview was done via Skype.

When analyzing the data collected from the interviews it is notable to observe that the regional partners were: 1) strong advocates of open data service development; 2) eager to match ideas elicited from the project with third party developers and have pilot services developed; yet 3) tremendously unfamiliar with what open data actually is beyond the word “open data”. In one of the interviews the regional representative states: “Open data is a great opportunity to activate developers outside the PTA (Public Transportation Authority) and have services built which the PTA never will build themselves. However I do not really know what open data exists, how it technically works and what quality it has.” In another later interview a regional representative from another region states: “I really wonder if the service ideas identified during the SIW (Service Innovation Workshop) actually are possible to build with the data that exists. I am not myself certain what data exists and what data does not exist in our region.”

The analyses of the interviews resulted in the identification of two major challenges. The first challenge was that the knowledge among the intended assessors was low in terms of open data availability, where data was available and the quality of the data in relation to service ideas for the region. Secondly, was the capability low amongst the intended assessors to perform systematic assessment.

In order to improve the assessment capability, researchers, supported by external specialists in open data and API design, developed, based on the challenges, a comprehensive model for open data assessment. The assessors then used this model as a supporting toolbox to assess the available open data in the different regions.

6. A model for open data assessment

The main purpose of the model is to provide support to the user (the “assessor”) to develop organizational awareness of the open data situation in a specific context (e.g. a city, region, application area). In the case explored the “assessment context” constituted by a region and an application area (e.g. Värmland, Sweden (region) and Public Transportation (application area)) and the intended user of the model was representatives from the regional partners in each of the participating regions, thus acting as open data brokers. In that role the regional partner attempted to match service ideas improving public transportation and third party development teams with appropriate data in that region. The teams should then build novel digital service pilots that enhance public transportation in rural areas based on open data in that region.

In the following sub-sections the model for open data assessment is presented. The process model is divided into three phases (see figure 2): 1) scan available data sources (DS) and classify the identified sources into types; 2) perform detailed review and assessment of each open data source; 3) perform compilation and comparison.

![Figure 2. A three-phase assessment model](image)

6.1. First phase of open data assessment

The 1st phase in the model generates an overall assessment of the open data resources available in the demarcated context. Four broad categories for data type assessment is provided:

*What open static data is available?* As the name implies, this type of data has low volatility and is typically stored in databases or spreadsheets. This type of data – given that its content is explained thoroughly – often requires less effort to publish compared to other types: since data is rarely updated as it is a) easier to maintain data integrity and accuracy and b) expected traffic to such sources is low. One example of such data related to public transportation is bus stops (e.g. name and location).

*What open dynamic data is available?* In contrast to static data, dynamic data is changed frequently and the value-in-use is tightly connected to the age of the data. While fresh dynamic data can provide...
substantial value about a current situation, it also puts more requirements on the publishing organization. Dynamic real-time data needs to be updated frequently by service developers, which requires scalable IT infrastructure as well as possible authentication mechanisms (to impose data retrieval rate limits). E.g. in public transport such data could be the current location of a specific vehicle or an estimated time of arrival.

What open statistical data is available? This type of highly aggregated data for a specific purpose is often an openly published data source. One feature of statistical data is that its content often is updated at quite long intervals. A second feature is that the value-in-use often is created only when the type of statistical relationships sought is delivered. Since the number of such combinations is high and difficult to anticipate beforehand, publishing statistical data typically requires a query interface in order to extract the type of statistics sought by the user. In this sense it differs from static data and is treated as a separate category. E.g. in public transportation, on-time statistics is one example of statistical data.

What open services are provided? In some cases, the data necessary to develop a third-party service cannot be easily extracted from an existing data source. Instead a service must be provided, based on the preconditions of a situation that subsequently performs an array of data processing to deliver a correct answer. E.g. for travel planning, a suitable route can only be suggested when the user provides origin and destination alongside personal preferences such as maximum number of changes and/or necessary disability facilities.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Static</th>
<th>Real time</th>
<th>Statistics</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public transportation</td>
<td>Bus stop, Bus line (geographical), time table</td>
<td>Bus position, delays, disturbances</td>
<td>Line, departure density,</td>
<td>Travelplanner</td>
</tr>
<tr>
<td>Road administrator</td>
<td>Road network</td>
<td>Floating car data, accidents, Road usage, transport types,</td>
<td>Que-detection, parking advice</td>
<td></td>
</tr>
<tr>
<td>Taxi</td>
<td>Taxi stops, car register</td>
<td>Positions</td>
<td>Route statistics</td>
<td>Ordering, provide price est.</td>
</tr>
<tr>
<td>Public service</td>
<td>Location of schools, hospitals etc., school curriculum</td>
<td>Available healthcare appointments</td>
<td>Demographics</td>
<td>Doctor’s appointment booking</td>
</tr>
</tbody>
</table>

Figure 3. Example of output assessment phase 1

These four broad categories provide units of analysis in the 1st phase of assessment. They allow the assessor to identify the data sources available in a specific context, and consequently assess the types of resources available in that context. In figure 3 an example is provided from the case illustrating a result from the 1st phase of assessment in one of the participating regions.

6.2. Second phase of open data assessment

In the 2nd assessment phase each of the data sources identified through in the 1st phase is in-depth analyzed using six generic dimensions (c.f. figure 4). The six generic dimensions are: Access to the data source, Support available to use the data source, License terms in order to use the data source, Costs for using the data source, Technical format used, and Quality in the datasets.

![Figure 4. Example of output assessment phase 2](image)

6.2.1. Access. External innovators need to have access to the data in order to be able to develop open data services and the way data is made accessible can significantly affect how third parties will use and reuse an open data resource [22]. In the 2nd phase the assessment model therefore contain categories to help the assessor to evaluate and classify the type of access available for a specific data source: a) No access (where data exists within the organization but is not for external parties); b) Anonymous access (where data is available without registration); c) Online registration (data is available for users registered with an account online); and d) off-line registration (where re-users must sign a written contract in order to access the data).

6.2.2. Support. A specific feature of third-party development, especially among mobile application developers, is that their work situation is marked by hyper-competition where speed-to-market is essential [23, 24]. In this sense, as an open data provider, it is deemed essential to optimize time spent by third-party developers and ensure they make the most of their often scarce resources to add functionality to services, rather than dealing with technical issues and misunderstandings in interpreting the data sources. Hence, adding a formal support function such as an on-line forum helps developers increase speed-to-market but also may work as a “knowledge trading zone” where input on future improvements can be made by developers.

6.2.3. License terms. While the data itself represents what is possible to do in a technical sense, the license
terms stipulate what is allowed for external innovators to do with the data [25]. Advocates of open data often argue for liberal terms-of-use (e.g. OKFN [26]); however, organizational considerations (e.g. using data in direct conflict with organizational goals), legal constraints (e.g. storage and processing of personal data) or public/private partnership regulations may not allow such permissive license terms. To this end, the model includes a number of generic licensing issues (commercial restrictions, attribution restrictions, usage restrictions, access restrictions) to help raise organizational awareness of the license situation for each data source and how license terms may hamper open data service development.

6.2.4. Cost. The third generic assessment dimension is cost. The debate on open data has largely centered on the costs associated with publishing data openly [7]. While most public data is shared with external actors on equal terms, it may: a) be completely free; b) be free in some versions or under some conditions (after the data has reached a certain age, a part of the dataset and/or a highly aggregated version of the dataset); or c) the data will not be released unless upfront payment is provided. As costs have been identified as a barrier-to-entry, decisions on potential charging will affect the uptake amongst innovators.

6.2.5. Technical format. An important decision is the technical format in which data is released. Inspired by the widely recognized "Five Stars on Open Data" [27], the technical format dimension in the 2nd level of assessment supports the assessor to determine the technical implementation of the data source through a categorization: a) non-digital format (e.g. paper documents); b) digital off-line format (e.g. distribution on portable media); c) digital on-line, non-structured format (e.g. PDFs available for download); d) digital on-line, structured non-versioned format (e.g. CSVs available for download); e) digital on-line, structured versioned format (e.g. versioned API, URI-based schemas).

6.2.6. Quality. One of the major barriers to exploiting open data to its fullest concerns quality [7], and third party developers often struggle in maintaining and updating open datasets to be able to provide a good end user experience [3]. To raise awareness of quality issues with open data, the model thus also includes an assessment tool for quality issues in the open data source typically arising in the public transport sector: 1) update frequency (e.g. how often are ETAs updated?); 2) consistency (e.g. are bus stops identified in a uniform way?); 3) accuracy (e.g. are geographical data points correct?); 4) completeness (e.g. does the dataset include all bus stops?).

6.3. Third phase of open data assessment

By using the two first phases in the model, the assessor develops an overall understanding of 1) what data sources are available in a specific context, and 2) a detailed assessment of each data source in terms of access, available support, license terms, costs, technical format and dataset quality. The aim with the third phase is to compile the different assessments performed in the 2nd phase and supports a comparison between the data sources. The intention is to identify differences between datasets regarding data type, access, support, license terms, cost, technical format and quality. This compiled assessment constitutes the organizational awareness of the open data situation in the demarcated context.

For the project stakeholders in the researched case the regional assessments resulted in four overall implications: 1) the regional partners, acting as open data brokers, lowered expectations regarding what services actually could be built on the data available in the different regions; 2) third-party developers could systematically be coached in using the open data available. This was facilitated by introducing a helpdesk that served third party developers when using the open data sources; 3) differences between regions in terms of available open datasets with public traffic data, quality in data provided and regional licensing restrictions were identified, all acting as barriers both for developing services in the regional open data market as well as transferring services developed in one region to other regions (i.e. markets) – which was one of the overall purposes with the project as a transnational project; 4) in order to lower this barrier, an open data platform was constructed within the project that provided a single interface with APIs toward third party developers to use regional datasets (where available) to build services. The assessment resulted thus in 1) increased awareness among stakeholders; 2) improvement activities to manage some of the identified challenges; 3) modifications in expectations regarding the end results.

7. Discussion

In the following two sub-sections we will first discuss practical implications of the result derived from the case study in relation different stakeholders. This is followed by a discussion where we elaborate how this study adds to the evolving theoretical body
regarding open data development and open data markets.

7.1. Practical implications

In section 2.2, stakeholders were identified in relation to open data markets, of which three are focused in this paper: 1) the open data provider; 2) the third-party developers; 3) the open data broker. As discussed these roles have different motives in relation to open data, and thus also why, when and how to perform and use results from open data assessment as presented in chapter 6 (c.f. table 2).

7.1.1 Value for open data providers. Open data assessment is useful for open data providers as it provides the means for these stakeholders to systematically understand if a gap exists between the data provided by them and the actual need of data in a demarcated context, as depicted in figure 5.

Open data assessment can consequently be used by the data provider to trigger the process to make the decision whereas the data sources passed on beyond the organizational borders should either 1) be improved in terms of i.e. support, license terms, quality, cost, format or access; and/or 2) be complemented with additional datasets; or 3) no action at all should be taken.

7.1.2 Value for third party developers. Open data assessment is not only valuable for data providers. It is also valuable for developers prior to open data service development, as described in figure 6.

By assessing what data sources are available in a context, and comparing the result of that assessment with the service idea at hand and thus the data needed to transfer the idea into a open data service, third party developers can use assessment to validate if there is a gap between what they want to build and the resources available to do it. If there is no gap or a gap exists that is not that significant, then third party developers can decide to go ahead with the
development. However if a significant gap is present between idea and data available, then the assessment can be used as a base to change priority and focus on something else; that is, either dropping the idea or pausing development awaiting changes in the data provided. The latter alternatives are a major peril for the sustainability of a viable open data market, as the market requires consumption of data by developers to be viable.

7.2. Theoretical contributions

In Lindman et al. [1] a number of research questions are identified to stimulate research with theoretical contributions to the field of open data service development. The result from this study meet that “call-to-arms”. However, as this study uses an exploratory single case study, the theoretical contribution is restricted in terms of generalization especially when compared with a study design that uses multiple comparable cases [20]. However in relation to the research agenda, this study through the richness of its critical case provides two significant contributions.

First, in order to create a sustainable open data markets, the number of stakeholders must be increased from customer, data provider and third party developers to include additional stakeholders such as the open data broker. This is an important stakeholder that cuts in between developers and providers with the aim to enhance the marketplace in terms of sustainability based on either commercial or non-commercial grounds. This role aims to attract more actors from the other categories to the marketplace and, by stakeholders together to enable viable distribution and consumption of open data [10].

Secondly, this study shows that open data now has past the stage where open data publishing is the only crucial activity. Open data and open data markets now exists. Other challenges that need management surface instead [7, 8, 9]: e.g. open datasets are hard to locate, support varies, license restrictions hamper usage, and uncertain sustainability of data provision acts as barriers. These challenges inhibit open data service development and have to be managed [28] so that they do not danger the viability in open data markets.

In order to improve and support usage of open data, this paper 1) shows how assessment of existing open data sources in a market can be performed; 2) discusses the value that different stakeholders obtains in performing systematic assessments; and 3) elaborates how increased awareness can be used by different stakeholders to make decisions connected to open data service development and thus add to the sustainability of open data markets.

8. Conclusions and future research

By building on experiences from a single but critical case, benefits in using open data assessment have been explored in this paper, stating why, when and how stakeholders can use open data assessment.
A suggestion for future research is that the results from this study are used as a base for a confirmatory study with additional cases to evaluate and improve the contributions put forward here. One idea could be to investigate how the assessment model could be expanded with a pre-defined maturity level scale in order to grade a data provider’s maturity in providing open data as well as the data itself. Inspiration could here be used from the field of process modeling maturity assessment [15]. The aim then is to investigate if such a model could be an efficient instrument to establish continual improvement of data sources available on an open data market, reinforcing the market’s viability. In other words use open data assessment not only to mind present gaps, but also to remove and prevent future gaps between developers’ requests of data and providers’ provision of data.

9. References