Critical Factors for Open Data Publication and Use

A Comparison of City-level, Regional, and Transnational Cases

susha, Iryna; Zuiderwijk-van Eijk, Anneke; Charalabidis, Yannis; Parycek, Peter; Janssen, Marijn

Critical Factors for Open Data Publication and Use: A Comparison of City-level, Regional, and Transnational Cases

Iryna Susha*, Anneke Zuiderwijk**, Yannis Charalabidis***, Peter Parycek****, Marijn Janssen**

* Örebro University, Sweden, iryna.susha@oru.se
** Delft University of Technology, The Netherlands, a.m.g.zuiderwijk-vaneijk@tudelft.nl, m.f.w.h.a.janssen@tudelft.nl
*** University of the Aegean, Greece, yannisx@aegean.gr
**** Danube University Krems, Austria, peter.parycek@donau-uni.ac.at

Abstract: There is a lack of research concerning the factors influencing the success or failure of open data initiatives. Based on the results of two workshops, we provide a list of 47 success factors for open data publication and 18 success factors for open data use. We further use three case studies (ENGAGE, Open NY, and Open Vienna) to examine how the criticality of factors varies depending on the geographical level and other characteristics of the open data initiative. The cases, representing open data initiatives at city, regional, and transnational levels, point at different categories of critical success factors. Our key conclusions are that 1) the criticality of the factors depends considerably on the context of the open data initiative; 2) a number of success factors appear to be more universally applicable than others; 3) the factors that are critical to all three cases are derived from many different success factor categories, which suggests that open data initiatives should adopt an interdisciplinary approach, and 4) further work is needed to detail the success factors for open data publication and use in other contexts.

Keywords: open data, success, success factors, level, taxonomy, e-government, case study

1. Introduction

The Critical Success Factor (CSF) approach has been extensively used in the Information Systems discipline to study success of project implementations. CSFs are understood as the essential areas in which desired results lead to successful competitive performance (Borman & Janssen, 2013; Rockart, 1979). In our previous work we found that CSFs are a productive approach to understanding and

capitalizing on the success of open data initiatives. Thus this paper builds on and extends a paper that we presented at the Conference on E-democracy and Open Government (see Zuiderwijk, Susha, Charalabidis, Parycek, & Janssen, 2015).

Applied to the field of open data, we define critical success factors for open data as factors which are critical for the successful implementation of an open data initiative. An open data initiative is any activity that aims at improving the publication and/or use of open data, including initiatives on different levels (e.g., international, national, local), by different stakeholders (e.g., civil servants, citizens, universities) and concerning different types of open data (e.g., on different topics, such as social sciences, city construction or health). A successful open data initiative is one in which data supply stimulates use and generates value. While many open data initiatives have been set up in the last few years, the success of these initiatives cannot be guaranteed. Many authors have pointed at the impediments of open data initiatives that may block the derivation of value from the publication and use of open data (e.g., Barry & Bannister, 2014; Conradie & Choenni, 2014; Janssen, 2011). So far the potential and success of Open Government Data (OGD) appears to be supported mainly by anecdotal evidence (Jetzek, Avital, & Bjorn-Andersen, 2014). Various researchers have described factors which are important to make open data initiatives successful (e.g., Parycek, Höchtl, & Ginner, 2014; Zuiderwijk, Janssen, Choenni, & Meijer, 2014). Yet, what is perceived as successful depends on the context of the initiative.

The literature does not describe success factors for open data initiatives that take this context into account. Identifying context-dependent open data success factors may foster the publication of public data, and it may help policy makers, civil servants and other decision-makers who plan to start an open data initiative to consider whether they will participate in the initiative and under which conditions. This is expected to foster the publication of data, the successful use of published data and to stimulate its economic and societal applications.

To contribute to research on context-dependent open data success factors, this paper aims to answer the following question: Which factors are critical for the publication and use of open data in particular practical cases? This paper first obtains a broad overview of open data success factors, and secondly specifies which factors from the broad overview are critical in particular contexts.

2. Research Background

2.1. Research on Success Factors in General

Interpretations of what the concept of success encompasses might differ. In technology-driven project implementations, success is multidimensional and somewhat ambiguous. The expectations and evaluations of project performance might differ among project stakeholders. The Information Systems success model (DeLone & McLean, 2002, 2003) arguably provides the most comprehensive view and combines the key dimensions defining success. This model advocates that a system can be evaluated in terms of information, system, and service quality. These characteristics affect the subsequent use or intention to use a system and user satisfaction. According to the model, certain benefits will be achieved as a result of using the system. Following this view, in our study we assume

that two dimensions indicate success of an open data initiative: 1) the (focus on the) use of published open data, and 2) the (focus on) emerging impacts and benefits. We use this conceptualization of open data success as a backdrop for our critical success factors (CSFs) study.

2.2. Open Data Success Factors

The notion of success in existing open data literature can be related to a number of research themes, for instance, success in relation to the evaluation of open data implementations, maturity of open data initiatives, progress and development of open data infrastructures and benchmarking of open data efforts. It has been more common to investigate failure and barriers in open data initiatives than to define and measure success. Apart from the study that this article builds on (see Zuiderwijk et al., 2015), we are not aware of any other holistic and comprehensive CSF analysis that has been carried out in relation to open data. We found only a few studies that explicitly refer to the term *success factor* in relation to open data initiatives. A search in Scopus for *open data* in the keywords and *success factor* in all fields returned thirteen entries of which only three appeared relevant.

Table 1 summarizes the success factors mentioned in the three selected studies. One of these studies (Sayogo & Pardo, 2013) used the term *success factor* to investigate the motivation and the driving forces of adoption of an open data initiative; this is a specific interpretation of success factors expounded as *drivers*. Other studies use terms such as *enablers* (van Veenstra & van den Broek, 2013), *facilitators* (Cranefield, Robertson, & Oliver, 2014), or *organizational measures* (Susha, Grönlund, & Janssen, 2015). In sum, there is a research gap as a holistic framework of CSFs in relation to open data publication and use is missing. The studies that explicitly refer to the concept of success factors are mainly focused on the data publication dimension.

Table 1: Overview of Success Factors Mentioned in the Literature

Area in focus	Identified success factors	Authors
Evaluation of open data strategy and portal	Political will, decision and resources CIO in charge of the implementation process Clear definition of responsibilities Implementation using a process model Community Management	Parycek et al. (2014)
Design of open data publishing process	Think early about data publication Develop guidelines about privacy and policy sensitivity of data Provide decision support and liaise with other departments involved Make data publication an integral, well-defined, and standardized routine task Monitor how published data is reused	Zuiderwijk, Janssen, Choenni, et al. (2014)

Motivation to adopt a smart disclosure policy	Success factors understood as economic rationale, regulatory and policy incentives, technology incentives and mimetic forces Success attributed to the interaction between aforesaid factors and internal drivers of an organization, such as strategic fit and alignment and reputation risk	Sayogo and Pardo (2013)
Total: 3 articles		

3. Research Methodology

In this study we opted for combining multiple research methods, since we first aimed to obtain a rich overview of factors which are important for open data publication and use, and then to narrow down this overview to factors which were critical for the publication and use of open data in specific cases. First, during a brainstorming session four academic researchers with expertise in the field of open data each participant individually developed an initial list of factors which influence open data publication and use based on their previous open data research (e.g., Parycek et al., 2014; Zuiderwijk, Janssen, & Davis, 2014). The lists were developed from different perspectives (i.e., open publication and use) so that together they contained as many factors as possible.

Second, all four researchers presented the identified factors in two workshops:
- A 3,5 hour workshop at the Electronic Government and Electronic Participation Conference in September 2014 (see Susha, Zuiderwijk, Janssen, Parycek, & Loukis, 2014). The workshop was attended by 18 experts from the field of e-government and e-participation and involved especially experts concerned with open data research. About two-thirds of the participants were male and most participants were between 30 and 50 years old.
- A 1,5 hour workshop at the Conference on E-democracy and Open Government in May 2015 (see Susha, Zuiderwijk, Janssen, Parycek, & Charalabidis, 2015). The workshop was attended by 28 experts from the field of open government and open data. Most participants were aged between 30 and 50 years old with the majority being male.

The workshop participants on both occasions received an initial list of factors important for open data publication and use as a paper version, and they were asked if there were factors that they wanted to add to the list. Thereafter, additional factors were discussed with the group of participants and more factors were added.

The initial brainstorming session and the two interactive workshops resulted in a comprehensive list of factors important for open data publication and use. However, the overview was generic and did not account for the influence of contextual factors critical for open data publication and use. As a third research step, we conducted case studies, since this method allows for explorative research to investigate contextual factors. Case studies can help in "understanding the dynamics present within single settings" (Eisenhardt, 1989, p. 534). A case study can be defined as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin, 2003, p. 13). The

cases were selected based on theoretical sampling (Eisenhardt, 1989), i.e., they were chosen because they were expected to replicate previous cases, extend emergent theory, fill theoretical categories or provide examples. The following case study selection criteria were defined:

- The cases employ an open data initiative concerned with open data publication and open data use.
- The cases represent successful open data initiatives on a diversity of geographical levels and topics. In our case selection we used the diversity technique and chose the cases which represent examples of different categories of open data initiatives: one at international level, one at regional level, and one at city level. The national level was skipped in our inquiry because open data initiatives at supra- and sub-national levels have so far received much less attention than those at national levels. Hence, it is our intended contribution to address this gap and provide insights into the specific CSFs which apply at these *untypical* instances of open data initiatives.
- The cases allow for identifying factors which are critical to the publication and/or use. The data for the case studies was collected during the second workshop in May 2015 by asking the participants to fill in a questionnaire and provide an assessment of factors from the perspective of an open data initiative in which they were directly involved. The questionnaire contained all the factors identified in step 2, the criticality of which was to be ranked by the respondents on a five-point Likert scale based on their experience with a particular open data initiative.
- The case can demonstrate results of being a successful project. This means that the case should focus on the use of published open data, and on the emerging impacts and benefits of the open data initiative.
- Case study information should be available and accessible. Case study information was obtained through the second workshop, through desk research, and follow-up interviews with respective workshop participants as case study informants.

Based on these selection criteria, the following three cases were selected.

- The first selected case (international level) was the European FP7 ENGAGE open data initiative, which aimed to develop an infrastructure incorporating diverse governmental data resources, empowering researchers and citizens, and stimulating scientific collaboration and research (see www.engagedata.eu). The developed infrastructure provided tools for dataset processing and acquisition. It went beyond existing simple open data repositories by additionally offering enhanced rich metadata to allow for the improved search and utilization of datasets, and it provided a social and collaborative space for open data users. The initiative provided a one stop-shop to more than 52,000 datasets and has developed a community of over 700 registered users.
- The second selected case (regional level) was Open NY (Open Data State of New York) (see https://data.ny.gov/). This initiative aimed to make state government information more accessible to the public. It went beyond the release of datasets by also providing services to developers (e.g., offering an API endpoint for every dataset on data.ny.gov), by developing and providing interactive applications and by organizing public challenges to encourage the creation of apps to improve the lives of NY citizens. The initiative was launched in 2013, and

has grown to over 1,200 New York State data resources on topics ranging from farmers markets to broadband availability to health data.

- The third selected case (city level) was Open Government Vienna (see https://open.wien.gv.at/). This initiative mandated an external impact evaluation about the realized benefits through open data. The evaluation focused on the value of open data, and thus this initiative also focused on how the supply and use of open data can stimulate emerging impacts and benefits in society.

We believe that all three open data initiatives demonstrate results of success by focusing on the use, impact, and benefits of the initiatives. All three cases went beyond merely publishing open datasets, and in addition to the supply of open data they paid attention to the use and emerging impacts and benefits of the initiative.

4. Critical Success Factors for Open Data Publication and Use

4.1. Findings Regarding Open Data Publication

Table 2 provides an overview of the identified factors that are important for open data publication and shows to which extent each factor was critical in the case studies. Three categories of factors were most critical for the disclosure of open data in the ENGAGE initiative, namely 1) legislation, regulation and licenses, 6) sustainability of the open data initiative, and 9) accessibility, interoperability and standards. With regard to the Open NY initiative, the table shows three categories of factors critical for the publication of open data, namely 1) legislation, regulation and licenses, 4) training of and support for civil servants, and 8) open data platforms, tools and services. In the Open Vienna Initiative, the most critical factors were found in the categories of 2) strategy and political support, 3) management support and publication processes within governmental agencies, and 7) collaboration.

Table 2: Factors Important for the Publication of Open Data and Factors Identified as Critical in the Case Studies

Categories	Factors important for the publication of open data	Identified as critical (+) or very critical (++) in the case studies		
		ENGAGE	Open NY	Open Vienna
1 Legislation, regulation and licenses	Having in place a (national) legal framework for open data publication	++	+	
	Enforce publishing and curating of data on administrations (maybe even through penalties)	++	+	
	Provide information about data protection and privacy legislation and how open data can be published in compliance with this legislation	++	++	

	Develop a (national) guide on legal Intellectual Property Right (IPR) issues allowing organizations to pick the correct licensing form	++	++	++
2 Strategy and political support	Develop a strategy for open data publication at an (inter)national level	+		+
	Ensure that (top) management within governmental agencies supports publishing data		+	++
	Generate support of policy-makers for data publication		+	+
	Organize focus groups with heads of departments and open data policy implementers to give both proponents and opponents of open data an auditorium		++	++
	Introduce incentives schemes for public servants (e.g., explain why a data provider would release data, explain what kind of value is created for the data provider)			+
	Create consensus between open data publication and the organizational framework for publishing data	++	+	+
3 Management support and publication processes within governmental agencies	Define clear process steps for publishing data	++	++	++
	Determine which type of data is important to address societal issues and focus on the publication of these data	++	+	+
	Start with the publication of data which is interesting for users so that the users see the benefit of open data	+	+	++
	Determine which data and metadata will and will not be published		++	
	Determine which standards and vocabularies will be used for data publication			++
	Determine which personnel has the key responsibilities for publishing open data		+	++
	Determine where datasets will be published	+	+	++
	Release only data which is of high quality	++		
4 Training of and support for civil servants	Create a virtual competence center which assists in answering questions and helping out with administrative data publication processes	+	++	++

	Provide training on open data publication within governmental agencies (e.g. training on how datasets can be anonymized)		++	
	Develop information campaigns in which questions about open data publication are discussed		++	+
	Develop information campaigns in which success stories of internal and external open data use are discussed		+	++
5 Evaluation of the open data initiative	Develop metrics and success indicators for data publication by government departments		+	
	Evaluate the realization of metrics and success indicators as an integral part of the open data initiative	+	+	
6 Sustainability of the open data initiative	Identify the need for data	++		
	Create a strategy for maintaining published datasets	+	++	+
	Ensure data provision continuity, including timely and automatic updates of data	++	++	
	Be transparent towards open data users about the conditions under which data publication takes place	++		
7 Collaboration	Arrange meetings with open data users to find out what their needs are and how the data from the governmental agency are used		+	++
	Organize internal meetings to discuss the data publication processes and to evaluate them		++	++
	Organize inter-organizational collaboration about and management of open data initiatives			++
	Ensure agile and open cooperation with various other organizations (administration, universities, CSO, Open Knowledge Foundation)			++
	Organize inter-organizational collaboration (e.g. network meetings) to learn from the open data initiatives of other governmental agencies		+	++
8 Open data platforms, tools and services	Integrate the open data platform into existing Content Management Systems (CMS) to kick-start the progress			++
	Have one central portal which combines data from many different	++		++

	governmental organizations (federal level)			
	Implement advanced data search functionalities	+	++	
	Use complementary toolsets for performing additional curation tasks (cleaning, linking, visualizing, analyzing)			
	Use a "web 2.0" approach for open data, allowing citizens to post, rate, work with datasets and web services	+	++	+
	Integrate frameworks for assessing data quality and usability of data and platform, providing continuous feedback to developers and administrations	++	+	
	Provide a forum to discuss what can be learned from open data use			
	Develop a clear User Interface (logical symbols, clear setup of the web page, simple design)	++	++	+
9 Accessibility, interoperability and standards	Use standards for data, metadata, licenses, URIs and exchange protocols	++	+	++
	Use cloud infrastructures able to gather, manage and publish open data, interoperable with other sources within the country or region			
	Integrate metadata schemas and federated controlled vocabularies for properly categorizing information	++	+	++
	Provide various types of metadata, in line with metadata standards (e.g. CERIF, CKAN, DC, EGMS, DCAT)	++	++	
	Provide Application Programming Interfaces (API's) for open data provision in the form of service feeds (from open data to open services)	++	++	+
	Enable multilinguality of metadata and data, allowing for the reuse and integration of data from different countries/languages	++	++	

4.1.1. ENGAGE

As far as the ENGAGE initiative is concerned, merely letting governmental organizations release their data voluntarily and on their own initiative seemed not sufficient to stimulate the use of and value generation from these data. To make the ENGAGE open data initiative successful it was found that it is critical to develop a legal framework for open data publication, to enforce the publication

and curation of open data on administrations, to offer public administrations information on how they can comply with data protection and privacy legislation, and to develop guidelines on issues with legal Intellectual Property Rights (IPR) allowing organizations to pick the correct licensing form. The criticality of these factors may stem from ENGAGE's focus on sharing datasets with various communities, as well as extending datasets and sharing those.

Many critical factors were also mentioned in the sixth factor category concerning the sustainability of the ENGAGE open data initiative, which can be explained by the fact that ENGAGE was a temporal project that could receive funding for three years only. Critical factors were also found in the ninth category concerning accessibility, interoperability, and standards, which is rooted in ENGAGE's focus on providing homogenous access to heterogeneous datasets. Essential factors with regard to sustainability concerned identifying the need for data, ensuring the continuity of data supply (including timely and automated updates of data), and being transparent towards open data users about the conditions under which data publication takes place. Factors regarding accessibility, interoperability and standards that were critical for open data publication success referred to multilingual metadata and data, the use of standards (for data, metadata, licenses, URIs and exchange protocols), the integration of metadata schemas and federated controlled vocabularies, the provision of various types of metadata in line with metadata standards, and the supply of APIs for open data provision in the form of service feeds.

Another factor that was not part of the broad overview of open data success factors derived from the brainstorm session and workshop, but that was found to be critical in the case concerning data stewardship and the development of a management plan for this. Stewardship refers to the "careful and responsible management of something entrusted to one's care" (Dawes, 2010, p. 380). Dawes (2010, p. 380) writes that stewardship in the context of information refers to "assuring accuracy, validity, security, management, and preservation of information holdings". Our third factor category already partly covers management support, yet a number of these factors were assessed as "not critical" by the case study respondent. The case study showed that some other management factors related to assuring accuracy, validity, security, and preservation were critical for the success of ENGAGE. Management factors that were assessed as very critical were releasing only high quality data, defining clear process steps for publishing data, publishing data that are important to address societal issues. The criticality of these factors stems from the fact that ENGAGE was a new initiative which started without any users. It needed to attract a large user base from the beginning of the project, which required the provision of accurate, valid, high-quality, and secure datasets that OGD users could trust.

4.1.2. Open NY

Since 2013 the state of New York has released over 113 million of records and over 1100 data catalogue items across over 61 government entities. A lot of these data were however already available prior to the launch of the Open NY initiative, hence one of the objectives of the initiative was to stimulate further release. Overall the Open NY initiative was driven by the pressure from the state governor, as other states had already provided open data portals but not the state of New York. Besides, stimulating open data publication was seen as a strategy to reduce the number of requests

for information and thereby reduce the burden on the local governments and agencies and improve information sharing.

With these objectives in mind it was important to make civil servants across diverse government organizations in the state of New York comfortable about publishing more data (which they otherwise protect) by enabling them to check that they do so in full accordance with all applicable legislation (namely on data protection and privacy). Providing such information to the civil servants was also required because at the state level there is no statistics entity which otherwise might be competent in data publication issues. This circumstance made it critical for the initiative's success to focus on factors related to legislation, regulation and licenses for the release of open data. Moreover, the data resources could be reused by anyone through the Open NY platform, which led to the criticality of developing a guide on legal Intellectual Property Right (IPR) issues allowing organizations to pick the correct licensing form.

The objective to stimulate release of more data by local governments and agencies participating in the Open NY also led to the criticality of factors in the category of training of and support for civil servants. In the US public officials, the majority of whom typically have a degree in public administration, are not generally prepared to bear responsibility over data issues. Therefore, very critical steps were to create a virtual competence center which assists in answering questions and helping out with administrative data publication processes, to provide training on open data publication within governmental agencies (e.g., training on how datasets can be anonymized) and to develop information campaigns in which questions about open data publication are discussed. Providing training and support for civil servants was required in order to distribute the value of open data publication at the lower level and thereby realize the internal benefits for government organizations (such as improved information sharing). Moreover, factors such as creating a strategy for maintaining published datasets and ensuring the continuity of data provision were important.

The New York data were published on the Open NY platform. In the Open NY initiative, it was therefore critical to pay attention to factors in the category of open data platforms, tools, and services. More specifically, the implementation of advanced data search functionalities, using a "web 2.0" approach for open data, and the development of a clear user interface were seen as very critical for the Open NY initiative's success. The importance of the platform, its search functionalities and its user interface is linked to the fact that the Open NY initiative targeted all kinds of audiences including civil society, citizens, and developers. Increased accountability was a powerful driver behind the initiative, to a much larger extent than potential economic effects. While this was a young initiative aiming to create a presence as an initial effect, providing a top quality site using existing state of the art tools and technologies was considered important.

Besides the listed factors, an additional success factor was formulated to be very critical in the Open NY case – the underlying data governance and data management capability of the organizations. Data governance can be defined as a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods (The Data Governance Institute, 2015). An open data governance framework hence comprises decision rights and accountabilities regulating the data publication process and can

include issues concerning data quality, metadata management, data privacy, data ownership, data policies (Eckartz, Hofman, & Van Veenstra, 2014). In the Open NY case building organizational capabilities in terms of data governance and management was of critical importance given the objective to establish a sustainable data publication process which can generate internal benefits for the governmental organizations. In order to make open data part of business as usual a data governance framework and capability to enact it were essential.

4.1.3. Open Vienna

Many CSFs for open data publication in the Open Vienna initiative were found in the categories of strategy and political support, management support and publication processes within governmental agencies, and collaboration. The following factors were among those seen as very critical: define clear process steps for publishing data, determine which personnel has the key responsibilities for publishing open data, start with the publication of data which is interesting for users, determine which standards and vocabularies will be used, determine where datasets will be published, ensure top management support, organize focus groups with heads of departments, and create a virtual competence center.

These CSF are based on the overall Viennese initiative, which has been driven by political decision and a strong implementation mandate for the administration. With regard to the category of strategy and political support, the Chief Information Officer fulfilled this political initiative and build up organizational and operational structure, e.g., a virtual competence center (experts from IT department, statistics department, Geographical Information System (GIS) department, directorate). Every department had to nominate a person responsible for the OGD implementation at the department level, and a structured open data publication process was defined. This was aimed to achieve coordination and control, as well as effective prioritization and evaluation, including at the political level. These organizational and operational structures are still in place and used for an ongoing quarterly publication process.

Further CSF in the category of management was the determination of standards and vocabularies for data publication process. Therefore a cooperation between open data providers from the different federal levels was established in the early beginning of the OGD initiatives in Austria. Based on the quickly developed technical and legal standards, a nationwide metadata portal was developed, which offers a comprehensive overview of OGD in Austria (www.data.gv.at). The same standards have been used from the beginning for the OGD Vienna initiative. As mentioned the clear processes and well defined units and roles are CSF (determine which personnel has the key responsibilities for publishing data on the portal). The final main factor in the management category was the focus on relevant data, which has been asked by the citizens. This issue is closely linked to the cooperation activities, as described in the following paragraphs.

With regard to CSFs in the category of collaboration, community management, and training, support of civil servants played an important role from the outset of this initiative. Even before Open Vienna was officially launched, the open data community and representatives from the Viennese administration discussed potential data sets and forms of collaboration. Thus the administration was well prepared for the political mandate and pursued the realization of benefits of open data.

Various other very critical factors for open data publication were found in the category of collaboration. The case study showed that it was very critical to arrange meetings with open data users to find out what their needs are and how the data from the governmental agency are used. Moreover, it was important to ensure agile and open cooperation with various other organizations (administration, universities, CSO, Open Knowledge Foundation), and to organize inter-organizational collaboration (e.g., network meetings) to learn from the open data initiatives of other governmental agencies. These factors were critical for the success of open data publication in the Open Vienna initiative based on internal and external stakeholder interviews, which have been done during the evaluation of the implementation. OGD is still a new domain for the administration and therefore they needed to implement strong internal and external cooperation for developing a vibrant feedback loop and a steep learning curve. Besides building up experience, the external collaboration with companies and civil society is one of the most important factors, because without cooperation through external developers only minor value for the broad public is created. For creating social and business value for a broad audience these new intermediaries are needed.

Several additional factors for open data publication and use were proposed in the Open Vienna case. Having in place a timetable and data management were found to be important for open data publication. The 2012 evaluation (Parycek et al.) came to the result that the combination of technical aspects, internal organization and the strong collaboration with civil society, companies and cross federal cooperation was the main success factor – "The whole is more than the sum of its parts".

4.2.   Findings Regarding Open Data Use

The survey also asked for factors which were critical for open data use in the open data initiatives (see Table 3). The most critical factors for open data use in the ENGAGE case were: 10) legislation, regulation and licenses, 11) success stories, and 14) feedback and sustainability. The Open NY case points at CSFs in the categories of 11) success stories and 15) research and education. The CSFs in the Open Vienna case were found mainly in the categories of 10) legislation, regulation and licenses, 12) incentives for open data use, and 14) feedback and sustainability.

Table 3: Factors Important for the Use of Open Data and Factors Identified as Critical in the Case Studies

Categories	Factors important for the use of open data	Identified as critical (+) or very critical (++) in the case studies		
		ENGAGE	Open NY	Open Vienna
10 Legislation, regulation and licenses	Provide information on the meanings and implications of licenses	++	+	++
	Provide information about privacy legislation and how open data can be used in compliance with this legislation	++	+	++
11 Success stories	Provide readily available examples of open data use (e.g. apps) to non-experts	++	++	
	Develop stories of successful open data use	++	++	+

Category	Measure			
	Involve community key players to propagate success stories	+	++	+
12 Incentives for open data use	Provide incentive schemes to engage citizens in open data usage			+
	Stimulate the development of specialized, open-data driven start up incubators		+	++
	Stimulate the development of business models to allow enterprises to develop add-on services on top of open data platforms, at a cost	+		
	Support issue-oriented community building through participatory events		++	+
	Align events, competitions and hackathons with, for example, university curricula, awards, festivals and "direct marketing"			++
13 Training of and support for open data users	Ensure agile, dynamic, and professional support services and training for potential open data users	+		
	Organize events and ensure community building where the potential benefits of open data are communicated to users (e.g. by building scenarios for usage)			+
14 Feedback and sustainability	Provide mechanisms for governmental agencies to know how their data have been reused	+		++
	Provide mechanisms for governmental agencies to know what can be learned from the reuse of their data	+		+
	Provide mechanisms for governmental agencies to know how the publication of their data can be improved based on feedback that they received from open data users	+	+	+
15 Research and education	Develop university and continuous education curricula on open data		++	
	Develop and maintain research areas roadmaps on open data, in order to consolidate research efforts and address open issues		++	

4.2.1. ENGAGE

In the ENGAGE case, none of the factors mentioned in the survey about open data use was assessed as uncritical or very uncritical. The factors in Table 3 that were not assessed as critical or very critical were assessed neutrally as "neither uncritical, nor critical". Factors that were very critical for open data use success were found in the tenth category regarding legislation, regulation, and licenses,

which can be explained by ENGAGE's focus on freely sharing datasets, and in the eleventh category concerning success stories, which can be explained by the importance for the project to attract a large user base. With regard to legislation, regulation, and licenses, the provision of information on the meanings and implications of licenses, and on privacy legislation and how open data can be used in compliance with this legislation were critical. As far as the success story category was concerned, the provision of readily available examples of open data use (e.g., apps) to non-experts, as well as the development of stories about successful open data use were critical. Success stories were used to attract more people to the infrastructure. In addition, all the factors in the category of feedback and sustainability were assessed as critical for the use of open data in the ENGAGE initiative, including the provision of mechanisms for governmental agencies to know how their data have been reused, to know what can be learned from the reuse of their data, and to know how the publication of their data can be improved based on feedback that they received from open data users.

A factor that was not part of the broad overview of open data success factors derived from our brainstorm session and workshop, but that was seen as critical was that open data users should know precisely and in a scientific manner the methodology of how the data were produced. For instance, if the dataset was derived from a survey, it should be clear to the user which method was used, how the sample of respondents was selected, and under which conditions the data were created and can be reused. Since datasets could be reused and changed on the ENGAGE infrastructure, users needed to know how datasets had been produced, and how changed datasets differed from original datasets.

4.2.2. Open NY

The Open NY initiative stimulated the use of open data by offering interactive applications and by organizing public challenges to encourage the creation of apps. For the involvement and engagement of users, it was critical to address all three factors in the success stories category. The Open NY initiative required the provision of readily available examples of open data use (e.g., apps) to non-experts, the development of stories of successful open data use, and the involvement of community key players to propagate success stories. However, as mentioned in section 3.1.2, this initiative was mainly aimed at creating a presence and establishing data publication routines; boosting the use of data given the young age of the initiative was not the first priority. This circumstance explains why factors in the categories of training and support of end users and feedback from end users to data publishers were not considered critical in the Open NY case.

In the category of research and education, the involved case study participant mentioned that it was very critical for the Open NY initiative to develop university and continuous education curricula on open data and to develop and maintain research areas roadmaps on open data, in order to consolidate research efforts and address open issues. The case study participant was also involved in education and research in the domain of open data, which illustrates the importance of factors related to education and research in the Open NY initiative. No additional success factors for open data use were identified in this case.

4.2.3. Open Vienna

A first important category of critical success factors for open data use in the Open Vienna case concerns legislation, regulation, and licenses. Both the provision of information on the meaning and implications of licenses and the provision of information about privacy legislation and how open data can be used in compliance with this legislation were assessed as very critical. The OGD Austria cooperation agreed on Creative Commons license CC as the only excepted license on the nationwide OGD portal. This decision was supported by the City of Vienna and solved the problem of contradictory licenses and uncertainty in the developer community.

Many CSFs for open data use in the Open Vienna case were also found in the category of incentives for open data use. Very critical factors in this category were to stimulate the development of specialized, open-data driven start up incubators, and to align events, competitions and hackathons with, for example, university curricula, awards, festivals and "direct marketing". The City of Vienna offers different kinds of financial support programs for companies and civil society organizations. The impact was generated by including components of the open data initiative into existing funding programs and ongoing cooperation and collaboration, which has led to several open data driven projects. Besides, the open data community was supported by competitions like gov20camp and Apps for Austria.

Finally, the case showed that it was critical for the use of open data in the Open Vienna initiative to focus on factors related to feedback and sustainability. The most critical factor in this category was to provide mechanisms for governmental agencies to know how their data have been reused. The quarterly OGD meetings, organized by the office of the CIO, are the main platform for communication, discussion, and direct feedback from the data community to the data providers. The second minor aspect is the possibility of linking a developed app to the used data sets, which offers a quick overview who is using the data. The sustainability is ensured through the ongoing quarterly publication process. Organizing meetings with internal experts and community was found to be an additional success factors important for open data use in the Open Vienna case.

5. Discussion of Case Studies

5.1. Contextual Aspects Influencing the Criticality of Success Factors in the Cases

The findings in the previous section showed that the three open data initiatives that we examined (ENGAGE, Open NY and Open Vienna) differ with regard to the criticality of success factors for open data publication and use. The case study participants pointed at different categories of CSFs for open data publication and use, and most critical success factor categories were not the same among the three initiatives. Only the category of 1) legislation, regulation and licenses for open data publication was seen as critical for more than one initiative. As the use of open data is concerned, factors in the categories 10) legislation, regulation, and licenses, 11) success stories, and 14) feedback and sustainability were identified as critical in at least two cases. Beyond the categories, there were many factors that were critical to all three cases, yet these were divided among the categories and there was not one clear category of factors critical to all three cases (see section 5.2).

The many differences in CSFs for the three cases might be explained by our case selection criterion to select cases that represent successful open data initiatives on a diversity of geographical levels and topics. The diversity in the cases allowed for identifying contextual aspects that influenced the criticality of success factors for open data initiatives. The case studies showed that the criticality of success factors depends on the context of the open data initiative. One major contextual factor is the geographical level on which the initiatives are organized. Our study involved open data initiatives on international, regional and city level. For example, since the ENGAGE open data initiative was organized on an international level and involved datasets from different countries in various languages, the factor of multilinguality was critical. In the Open NY initiative, there was no statistics entity at the state level that could provide help in data publication issues, which made it critical for the initiative's success to focus on factors related to legislation, regulation, and licenses for the release of open data. Since the Open Vienna initiative functioned on a lower level than the national level, it was important in this initiative to consider the already existing national Austrian OGD initiative. This made it critical for the Open Vienna initiative to use similar standards and ways of publishing as the national initiative and to cooperate with open data providers at different federal levels.

Moreover, the type of end-user in an initiative was an important contextual factor than influenced the criticality of success factors. The three initiatives focused on different types of end-users. The ENGAGE case targeted mainly researchers and citizens, while the Open NY case focused broadly on all audiences, and the Open Vienna case targeted the intermediaries like companies, civil society organizations, and developers. The criticality of the factors needs to be viewed from the perspective of the end-user focus. For instance, the factor to "stimulate the development of specialized, open-data driven start up incubators" was seen as very critical in the Open Vienna case and critical in the Open NY case because of their particular end-user focus, but this factor was not critical in the ENGAGE case since it focused on other end-users.

Furthermore, the phase of the open data process that the initiative focuses on, i.e., the publication or the use of open data, also influences whether a factor is critical for an initiative's success, and it is important to make a distinction between factors critical for open data publication on the one hand, and factors important for open data use on the other hand. The case study showed that some factors were only critical for the success of data publication, while others were only critical for data use. If these categories are not separated, this may lead to unrealistic expectations of open data decision-makers.

The case studies showed that the criticality of factors was also determined by the presence or lack of certain expertise necessary for open data publication or use. In the Open NY case, many of the involved civil servants were lacking experience with data publication, and it was therefore critical to train them on how to deal with data publication issues. In the Open Vienna case, a virtual competence center was created with experts from the IT department, statistics department, Geographical Information System (GIS) department and the directorate. This helped in building up the organizational and operational structure. Each of the cases showed that certain expertise is required for open data initiatives, and if this expertise is lacking, it becomes critical to obtain the expertise.

Furthermore, the temporal dimension is a contextual factor that influences CSFs for open data initiatives. For example, the ENGAGE case concerned a three-year open data initiative, and therefore many factors related to feedback and sustainability were critical. The temporal dimension is also important since open data CSFs may vary over time, and they may depend on a long-run or short-term focus of the open data initiative. For instance, while the factors of collaboration for data publication were not found to be critical for success in the ENGAGE initiative, these factors may be critical for the sustainable provision of open data in the long-run.

Other contextual aspects that may influence the criticality of success factors for open data use are whether the data are reused for commercial or non-commercial purposes, since the need to make open data use profitable is likely to lead to different success factors than not-for-profit data use. A factor such as the development of a successful revenue model was not critical in the context of the three cases that we studied since they did not aim to make money with open data use, while this factor may be critical for commercial open data use initiatives.

5.2. Universally Applicable CSFs?

The findings of our study offer several interesting insights and implications. An important point of reflection concerns the extent to which the identified factors (or some of them) can be universally applied to different organizational settings. In the literature (Poon & Wagner, 2001) there is a notion of *meta-CSFs* – a smaller number of CSFs of utmost criticality which if managed correctly result in all other factors going right as well. In our study, twelve factors were found to be critical or very critical to all three studied cases.

1) Develop a (national) guide on legal Intellectual Property Right (IPR) issues allowing organizations to pick the correct licensing form;
2) Define clear process steps for publishing data;
3) Determine which type of data is important to address societal issues and focus on the publication of these data;
4) Start with the publication of data which is interesting for users so that the users see the benefit of open data;
5) Determine where datasets will be published;
6) Create a virtual competence center which assists in answering questions and helping out with administrative data publication processes;
7) Create a strategy for maintaining published datasets;
8) Use a "web 2.0" approach for open data, allowing citizens to post, rate, work with datasets and web services;
9) Develop a clear User Interface (logical symbols, clear setup of the web page, simple design);
10) Use standards for data, metadata, licenses, URIs, and exchange protocols;
11) Integrate metadata schemas and federated controlled vocabularies for properly categorizing information;
12) Provide Application Programming Interfaces (API's) for open data provision in the form of service feeds (from open data to open services).

These factors might be more universally applicable to open data initiatives. From our study it does not become clear to which extent these factors also apply to other open data initiatives, and further research in this area is needed. Nevertheless, the twelve critical success factors suggests that successful open data initiatives need to cover many different categories, since the twelve factors above were very divided among different success factor categories, and only few factors belonged to the same category. This also suggests that people involved in open data initiatives need to work interdisciplinary to reach a solution for a common problem, where they involve experts from all the categories of success factors, such as legislation experts, training experts and sustainability experts.

6. Conclusions

This paper aimed to answer the question: Which factors are critical for the publication and use of open data in particular practical cases? A brainstorming session and two interactive workshops were used to first identify a comprehensive list of factors which are important for open data publication and use. Subsequently, the findings from three case studies showed which of these generic factors were critical for open data publication and use in particular contexts. The open data initiatives that we studied were selected based on diversity, and they focused on different geographical levels, namely on city, regional and transnational level.

The case study participants pointed at different categories of success factors that were critical for open data publication and use. Hence, our first conclusion was that the criticality of the factors depends considerably on the context of the open data initiative. The second point we arrived at was that a number of success factors appear to be more universally applicable than others. Despite the differences between the categories of CSFs in the cases, twelve factors were found to be critical or very critical to all three studied cases. They were divided among different success factor categories and only few belonged to the same category. Thus our third conclusion was that open data initiatives should adopt an interdisciplinary approach and involve experts from all the categories of success factors, such as legislation experts, training experts, and sustainability experts. The broad overview of generic factors obtained from the brainstorming session and the interactive workshop may be used by other researchers as a framework for investigating the criticality of success factors in a certain context. The case studies allowed for identifying context-dependent open data success factors, which may foster the publication of public data, foster the use of published data, and stimulate its economic and societal applications. The combination of methods used in this study was found to be useful to identify CSFs for open data publication and use. We recommend further research to complement the three case studies with other cases from a variety of contexts. This could help to determine the criticality of success factors in other contexts, and to obtain more insight in whether a particular open data initiative may be successful or not. It may also offer guidance to decision-makers regarding whether they will participate in an open data initiative and under which conditions. Additionally, further research could investigate whether quantitative approaches towards measuring the success of open data initiative could be useful.

References

Barry, E., & Bannister, F. (2014). Barriers to Open Data Release: A View from the Top. Information Polity, 19(1), 129-152

Bertot, J. C., McDermott, P., & Smith, T. (2012). Measurement of Open Government: Metrics and Process. Paper presented at the 45th Hawaii International Conference on System Sciences, Hawaii, U.S.A.

Borman, M., & Janssen, M. (2013). Reconciling two approaches to critical success factors: The case of shared services in the public sector. International Journal of Information Management, 33(2), 390– 400.

Conradie, P., & Choenni, S. (2014). On the Barriers for Local Government Releasing Open Data. Government Information Quarterly, 31(supplement 1), S10–S17.

Cranefield, J., Robertson, O., & Oliver, G. (2014). Value in the mash: Exploring the benefits, barriers and enablers of open data apps. Paper presented at the 22nd European Conference on Information Systems, Tel Aviv, Israel.

Dawes, S. (2010). Stewardship and usefulness: Policy principles for Information-Based Transparency. Government Information Quarterly, 27(4), 377–383.

DeLone, W. H., & McLean, E. R. (2002). Information systems success revisited. Paper presented at the 35th Annual Hawaii International Conference on System Sciences, Hawaii.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. Journal of management information systems, 19(4), 9-30.

Eckartz, S. M., Hofman, W. J., & Van Veenstra, A. F. (2014) A decision model for data sharing. Vol. 8653 LNCS. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (pp. 253-264).

Eisenhardt, K. M. (1989). Building Theories from Case Study Research. Academy of Management Review, 14(4), 532-550.

Hellberg, A.-S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. Transforming Government: People, Process and Policy, 9(1), 35-51.

Janssen, K. (2011). The Influence of the PSI Directive on Open Government Data: An Overview of Recent Developments. Government Information Quarterly, 28(4), 446-456.

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2013). The generative mechanisms of open government data. Paper presented at the 21st European Conference on Information Systems, Utrecht, the Netherlands.

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2014). Data-Driven Innovation through Open Government Data. Journal of Theoretical and Applied Electronic Commerce Research, 9(2), 100-120.

Murillo, M.J. (2015). Evaluating the role of online data availability: The case of economic and institutional transparency in sixteen Latin American nations. International Political Science Review, 36(1), 42-59.

Parycek, P., Höchtl, J., & Ginner, M. (2014). Open Government Data Implementation Evaluation. Journal of Theoretical and Applied Electronic Commerce Research, 9(2), 80-99.

Poon, P., & Wagner, C. (2001). Critical success factors revisited: Success and failure cases of information systems for senior executives. Decision Support Systems, 30(4), 393-418.

Rockart, J. F. (1979). Chief executives define their own data needs. Harvard Business Review, 57(2), 81-93.

Sayogo, D. S., & Pardo, T. A. (2013). Understanding Smart Data Disclosure Policy Success: The Case of Green Button. Paper presented at the 14th Annual International Conference on Digital Government Research, Quebec City, Canada.

Susha, I., Grönlund, Å., & Janssen, M. (2015). Organizational Measures to Stimulate User Engagement with Open Data. Transforming Government: People, Process and Policy, 9(2), 181-206.

Susha, I., Zuiderwijk, A., Janssen, M., Parycek, P., & Loukis, E. (2014). Workshop on Critical Success Factors for Open Data – From Policy to Participation and Innovation. Paper presented at the Electronic Government and Electronic Participation. Joint Proceedings of Ongoing Research, Posters, Workshops and Projects of IFIP EGOV 2014 and ePart 2014, Dublin, Ireland.

Susha, I., Zuiderwijk, A., Janssen, M., Parycek, P., & Charalabidis, Y. (2015). Workshop on context-specific critical success factors for open data publication and use. Paper presented at the Conference for E-Democracy and Open Government, Krems an der Donau, Austria.

The Data Governance Institute. (2015). Definitions of Data Governance.   Retrieved 11 August, 2015, from http://www.datagovernance.com/adg_data_governance_definition/

van Veenstra, A. F., & van den Broek, T. A. (2013). Opening Moves – Drivers, Enablers and Barriers of Open Data in a Semi-public Organization. Paper presented at the Electronic Government Conference 2013, Koblenz, Germany.

Yin, R. K. (2003). Case study research. Design and Methods. Thoasand Oaks: SAGE publications.

Zuiderwijk, A., Janssen, M., Choenni, S., & Meijer, R. (2014). Design Principles for Improving the Process of Publishing Open data Transforming Government: People, Process and Policy, 8(2), 185 - 204.

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems Information Polity, 19(1-2), 17–33.

Zuiderwijk, A., Susha, I., Charalabidis, Y., Parycek, P., & Janssen, M. (2015). Open data disclosure and use: critical factors from a case study. Paper presented at the Conference for E-Democracy and Open Government, Krems an der Donau, Austria.

About the Authors

Iryna Susha

Iryna Susha is a researcher in the Department of Informatics at Örebro University in Sweden. Her research interests include open government, open data, citizen participation, and public sector innovation. For more information, see: http://www.oru.se/Personal/iryna_susha/

Anneke Zuiderwijk

Anneke Zuiderwijk is a researcher in the Information and Communication Technology section of the Faculty of Technology, Policy, and Management at Delft University of Technology. Her research focuses on open data, see: http://tinyurl.com/AnnekeZuiderwijk.

Yannis Charalabidis

Yannis Charalabidis is Assistant Professor in the University of Aegean, in the area of Governance Information Systems, while also heading eGovernment & eBusiness Research in the Decision Support Systems Laboratory of National Technical University of Athens (NTUA), planning and coordinating high-level policy making, research and pilot application projects for governments and enterprises worldwide.

Peter Parycek

Peter Parycek is a full professor in E-Governance and head of the Department for E-Governance and Administration at the Danube University Krems and Chairman of the ministerial working groups "E-Democracy & E-participation" at the Austrian Federal Chancellery. He is responsible for the CeDEM conference and the open access journal JeDEM.

Marijn Janssen

Marijn Janssen is a full professor in ICT and Governance and head of the Information and Communication Technology group of the Faculty of Technology, Policy, and Management at Delft University of Technology, the Netherlands. For more information, see: www.tbm.tudelft.nl/marijnj.