



**HAL**  
open science

## Open Data Innovation Capabilities: Towards a Framework of How to Innovate with Open Data

Silja Eckartz, Tijs van Den Broek, Merel Ooms

► **To cite this version:**

Silja Eckartz, Tijs van Den Broek, Merel Ooms. Open Data Innovation Capabilities: Towards a Framework of How to Innovate with Open Data. 5th International Conference on Electronic Government and the Information Systems Perspective (EGOV), Sep 2016, Porto, Portugal. pp.47-60, 10.1007/978-3-319-44421-5\_4. hal-01636444

**HAL Id: hal-01636444**

**<https://inria.hal.science/hal-01636444>**

Submitted on 16 Nov 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Open Data Innovation Capabilities: Towards a Framework of How to Innovate with Open Data

Silja ECKARTZ, Tijs VAN DEN BROEK, Merel OOMS

*TNO, The Netherlands*

**Abstract.** Innovation based on open data lags behind the high expectations of policy makers. Hence, open data researchers have investigated the barriers of open data publication and adoption. This paper contributes to this literature by taking a capabilities perspective on how successful open data re-users create value out of the available data sources. First, a framework of IT, organization and skills capabilities required to innovate with data is derived from literature. Second, a case study including a survey and interview with managers from 12 frontrunners in the Netherlands was conducted. The analysis reveals that skills are valued the highest closely followed by organizational capabilities. Setting up a multi-disciplinary team with motivated employees and giving this team the mandate to experiment with data, is essential when innovating with open data. Theoretically, this study contributes to open data research by offering a new capabilities perspective on the organizational level. Our results highlight the importance of entrepreneurship theories to explain value creation with open data. Practically, our study suggests that digital skills and start-ups are important to the open government data policies.

**Keywords.** Open data; Re-Use; Public Sector Innovation; Value of Data and Apps; Capabilities

## 1. Introduction

An increasing number of government agencies release open data to spur economic growth through the development of new digital services or increasing organizational productivity. McKinsey [1] estimates that open data may provide 900 billion US dollar additional annual economic growth in Europe compared to an economy without open data. To unlock this potential, more and more data sources, ranging from transport to educational data, are available in open data portals throughout Europe. However, the question remains how to further exploit the economic potential of open data. Currently, innovation based on open data sets seems to lag behind the high expectations of policy makers. Hence, previous research took stock of barriers to the publication and use of open data. This paper contributes by taking a capabilities perspective to study how successful open data re-users create both economic as well as social value out of the available data sources. What can we learn from these frontrunners? Hence, the central research question is: “*What capabilities are most important to create value out of open data?*”

The gap between open data provision and usage is reflected in IS research. To date, most attention has been paid to barriers at the side of open data providers, which are predominantly government agencies [2-5]. Recently, scholars have started to

examine the impediments that open data re-users experience [5]. However, most of these studies focus on the technical barriers related to opening, finding and using data that hinder usage [5, 6]. Several scholars urge for more research on the non-technical barriers, such as those related to business cases, management support and organizational culture, that open data re-users experience [7, 8]. Only a few studies examine the capabilities that open data re-users need to create economic value from data [9, 10]. Jetzek et al. [9] describe capabilities as “*the collective ability of individuals and organizations to use and re-use open data*” and focus on access to open data and data literacy. Complementing Jetzek et al. [9], we aim to study open data innovation capabilities related to data re-use focused on the organizational level.

Our paper contributes to open data research in two ways. First, we synthesize information systems (IS) and innovation management (IM) literature on capabilities for data-driven innovation to develop a conceptual framework. To complement current open data research, this framework includes IT, organizational and people-related capabilities. Second, we present the results of the study of 12 Dutch frontrunners in innovation based on open data to test the conceptual framework. The study includes a questionnaire and semi-structured interviews with the CIOs and/or technical managers of the 12 organizations.

The practical implications of this study are twofold. First, we formulate the capabilities required for organizations that want to start to innovate with open data, e.g. by developing new services or improving their current practices. Firms may use the insights to develop an open data innovation plan that includes technical and non-technical measures. Second, policy makers may use the insights of this paper to develop policy instruments that stimulate data-driven innovation. For example, our research may fuel policy debate about the digital skills that organizations need.

## **2. An Open Data Innovation Capabilities Framework**

A literature review identified IS and IM journal articles and conference papers that study capabilities needed for data-driven innovation. Keywords for our search included combinations of the words e-Skills, capabilities, innovation, “data-driven”, “digital skills”, “data re-users” and barriers. We found 19 relevant studies that examined capabilities for data-driven innovation. The studies are mentioned in the tables below.

We clustered the results of our literature review around three main types of capabilities: IT related capabilities, organizational capabilities and skills-. Each capability type was identified in literature as influencing the data innovation capacity of an organization. Within each concept we distinguish a number of categories and sub categories, presented in the sections below.

In contrast to previous research [10], the selected IS and IM studies point at the importance of non-technical capabilities to innovate with data. This reflects the general notion in innovation management that three quarter of innovations’ success is related to non-technical factors [11]. Furthermore, many capabilities are related to small or medium-sized enterprises and not to large enterprises, for example entrepreneurship and agility.

### 2.1. IT Capabilities

With respect to IT capabilities we found three categories that influence the data innovation potential of an organization: 1) Infrastructure and enabling technologies that include all hardware and software that is needed to collect, store, analyze and visualize open data, 2) An IT strategy that includes planning, data management and governance, and measures to secure data, and 3) Interoperability that includes capabilities on working with standards and ability to integrate IS. Table 1 provides an overview of the IT capabilities, how they are defined in literature and a list of studies that mention them.

**Table 1 Overview of IT capabilities identified**

| Category                             | Capabilities               | Definition  | Studies            |
|--------------------------------------|----------------------------|---|--------------------|
| Infrastructure & Enabling Technology | Hardware & Software        | General IT infrastructure (incl. tools and applications) needed to create value from data                 | [5, 9, 10, 12, 13] |
|                                      | Data analytics             | Technologies (software and tools) for data search, discoverability and mining                             | [5, 12, 14-16]     |
|                                      | Data output/ visualization | Technologies for showing insights from data analytics, including visualization and reporting technologies | [10, 14, 15]       |
| IT Strategy                          | Planning                   | Vision and IS portfolio management , knowledge management, architecture                                   | [9, 13, 17-19]     |
|                                      | Data Management            | Data lifecycle management, technical data manipulation  | [12, 14, 18]       |
|                                      | Data Governance            | Processes and policies including data quality and usability   | [7, 12, 15]        |
|                                      | Security                   | System and data security processes and mechanisms   | [20]               |
| Interoperability                     | Standards                  | Interoperability/ Standards   | [5, 15]            |
|                                      | System Integration         | Systems integration, IT integration,  | [14, 15]           |

### 2.2. Organizational Capabilities

With respect to organizational capabilities we found four categories that influence the data innovation potential of an organization: 1) Strategic capabilities that includes top management support for data-driven innovation and the ability to change decisions and policies based on open data, 2) Tactical capabilities that include the mixing of disciplines over organizational silos, allocation of resources, and ability to adapt organizational processes based on data, 3) Operational capabilities that include entrepreneurship and R&D, and 4) cultural capabilities that include a culture focused on innovation and agility. Table 2 provides an overview of the organizational capabilities, how they are defined in literature and a list of relevant IS and IM studies.

**Table 2 Overview of organizational capabilities identified**

| Category  | Capabilities         | Definition                                       | Studies          |
|-----------|----------------------|--|------------------|
| Strategic | Leadership           | Top management support                           | [14, 20]         |
|           | Governance           | Data-driven policies and decision-making         | [13, 14, 16, 17] |
| Tactical  | Multi-disciplinarity | Integration capability, multi-disciplinary teams | [12, 18, 21, 22] |

|             |                             |   |                        |
|-------------|-----------------------------|---|------------------------|
|             | Coordination capability     | Enhanced knowledge across boundaries, resource allocation, asset orchestration                        | [9, 16]                |
|             | Processes                   | Data-driven processes across entire organization, Performance management based on data                | [13, 15-17, 20]        |
| Operational | Entrepreneurial orientation | Ability to recognize value of information, imitate competitors and commercialize value of information | [10, 20-23]            |
|             | Develop new ideas           | Research and Development, creativity and ideas  | [10, 19, 23]           |
| Culture     | Innovation culture          | e-Awareness, Socialization (shared ideology and collective identity), communication                   | [7, 9, 14, 17, 22, 24] |
|             | Agility                     | Degree of adaptability of the firm, change management, improvisational organizational capability      | [14, 17, 18, 23]       |

### 2.3. Skills

With respect to skills we found two categories that influence the data innovation potential of an organization: 1) Hard skills, and 2) Soft skills. Where hard skills can be defined as more technical skills such as programming or data analytics skills and soft skills as more non-technical such as interdisciplinary cooperation and entrepreneurship. Table 3 provides an overview of the capabilities, how they are defined in literature, and a list of relevant IS studies.

**Table 3 Overview of skills identified**

| Category    | Capabilities                  | Definition   | Authors                |
|-------------|-------------------------------|--|------------------------|
| Hard skills | e-literacy                    | Computer science skills, utilization, Architecture skills, e-user skills         | [14, 20, 22, 24, 25]   |
|             | Data literacy                 | Interoperability, semantic webs, linked data skills                              | [5, 12, 22, 24]        |
|             | Programming skills            | Different programming languages, hacking machine learning, NLP, Mapreduce /      | [5, 12, 22, 25]        |
|             | Data analytics                | Knowledge of data models, cleaning, integration and analysis                     | [5, 12, 22, 24]        |
|             | Research skills               | Research and analysis skills / Statistics  | [5, 9, 12, 16, 25, 26] |
|             | Data output skills            | Communication skills: Story telling with data, data journalism and visualization | [12, 25]               |
| Soft skills | Management skills             | e-leadership skills, decision quality  | [16, 24]               |
|             | Customer and service focus    | Service provisioning and receiving, customer focus.                              | [26]                   |
|             | Interdisciplinary cooperation | People integration, knowledge management, Inter-disciplinary domain knowledge    | [5, 14, 15]            |
|             | Entrepreneurship              | Proactivity, e-Business, competitive actions                                     | [14, 16, 17, 20, 24]   |

### 3. Methods

To answer our research question, we studied 12 leading open data re-users in the Netherlands. We decided to sample cases from the Netherlands as the research team had access to Dutch cases and the Netherlands has a lively open data community with governments opening up more and more datasets. However, despite the Dutch central government wishes the Netherlands to be an open data frontrunner in Europe, policy evaluations indicate that this still requires significant effort.<sup>1</sup> The open data frontrunners in this study were selected based on their presence as showcase in national and European open data research (e.g. the former ePractice portal and the European Commission's join up community) or because they have won national and European app awards. Furthermore, we selected cases that vary in size and sector to avoid selection bias. The selection process resulted in a longlist of 17 Dutch frontrunners of which 12 respondents finally participated in our research. Frontrunner only participated in the survey. The response rate of 71% can be qualified as very high, which indicates that respondents were highly involved with the topic and therefore more inclined to participate. Our research targeted the frontrunners' technical managers, innovation managers, CIOs or CTOs. Usually, only one respondent per case was contacted.

We used a two-step approach to measure the capabilities in our conceptual framework. First, we operationalized the capabilities in the framework in a questionnaire that included seven questions on the organization, their application of open data and which capabilities they find most important. We tested and improved the questions on comprehensiveness. The questions in the survey asked respondents to rate the importance of each capability to the success of creating value out of open data on a five-point Likert scale (ranging from very unimportant to very important). In addition to the closed questions, we allowed respondents to add their own capabilities and explain why they are important. After we received the answers on the questionnaire, we contacted each respondent within one week after filling in the questionnaire for a semi-structured interview. In this interview, we asked respondents to explain their answers: why are certain capabilities important or not important at all in their view? Interviews were conducted by phone and varied in length from 15 to 45 minutes. The questionnaire (in Dutch) and interview protocol are available on request.

Table 4 below provides an overview of the frontrunners that participated in our research.

**Table 4** Description of cases

| Case | Description and goal of open data usage                                      | Respondent Function            | Size (nr of employees) | Type of organization | Years of experience with open data usage |
|------|--|--------------------------------|------------------------|----------------------|--|
| A    | Development and improvement of new products and services based on open data. | Chief concept creation         | 11-50                  | App developer        | More than 5 years                        |
| B    | Development and improvement of new products and services based on open data. | Big data consultant & engineer | 11-50                  | Data / IT consulting | About a year                             |
| C    | Find and compare all   | Owner                          | 1-10                   | App                  | 3-5 years                                |

<sup>1</sup> See the report by the Dutch Court of Audit (2016) on trends in open data, p. 42, available at <http://www.rekenkamer.nl/dsresource?objectid=23808&type=org>

|   |   |                                  |            |                                   |                   |
|---|---|----------------------------------|------------|-----------------------------------|-------------------|
|   | schools in the NL. Development of new products (App) in order to create societal value.   |                                  |            | developer                         |                   |
| D | Publishing data as well as using data to improve internal processes and develop new services.   | Enterprise architect             | 1001-5000  | Energy grid provider              | 1-3 years         |
| E | Development and improvement of new products and services based on open data. Create societal value.   | Co-founder and creative director | 11-50      | App provider                      | More than 5 years |
| F | Consulting and development of new products and services based on (open) data.   | Chief marketing                  | 11-50      | Solution provider big data        | More than 5 years |
| G | New services (app) based on open data.  | Director                         | 1-10       | App provider                      | 3-5 years         |
| H | Mainly data provider, using some other (e.g. geo) data to improve their services.   | Specialist data management       | 5001-10000 | Governmental organization         | More than 5 years |
| I | Use data to provide new services (air quality measures) and create societal value.  | Data architect                   | 1001-5000  | Governmental organization         | More than 5 years |
| J | Provide new services (expected need for roadside assistance) based on open weather data.  | Chief Technology Officer         | 1-10       | Deep-learning start-up (spin-off) | 1-3 years         |
| K | Consulting and development of new products and services based on (open) data. App gives shippers real-time information on berth availability. | Co-founder                       | 1-10       | App provider                      | 1-3 years         |
| L | Provides software, consultancy and data services based on open geographical data sets.  | Manager Consultancy & Content    | 1001-5000  | Solution provider big data        | More than 5 years |

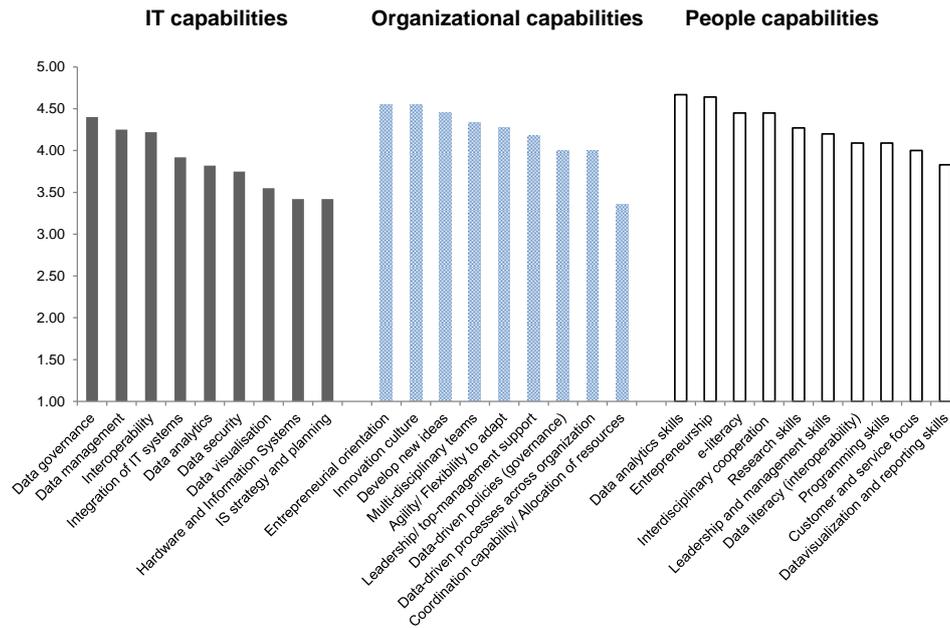
An analysis of the cases shows that two third of our respondents have more than 3 years' experience with open data. About half of our cases are organizations with less than 10 employees. Only 2 cases are large organizations.

Most organizations (91.7%) use open data to develop new products or services or improve existing products (58.3%). Half of the organizations aim to develop societal value with open data. Examples of societal value created by the organizations are insight in quality of schools or the plans of the local government in your neighborhood. Our respondents focus less on the optimization of internal processes and decision support with open data. This might be related to the relatively small average size of the organizations (mainly SMEs).

#### 4. Results

In this section we describe how respondents rated and explained the importance of IT, organizational and people-related capabilities for the use of open data. The focus in this

paper is more on the underlying reasons that organizations have for finding certain aspects important in relation to data innovation than to have representative scores. This is reflected in the modest amount of respondents. Therefore, the figure does not present representative numbers but a reflection of the average of our respondent group.



**Figure 1** Overview of how capabilities were rated on a scale from 1 to 5

#### 4.1. Information Technology Capabilities

The factors that were found to be most important are data governance (4.40, on a scale from 1 to 5), followed by data management (4.25) and interoperability (4.22). The factors that received the lowest scores were Hardware and Information Systems (3.42) and having an IS strategy and planning in place (3.42).

Looking at a higher level, we can conclude that IT capabilities related to the handling of open data are perceived as most important. First, data governance, which describes the processes and policies that ensures that important data assets are formally managed, is essential for data re-users as it ensures that data can be trusted and that data owners are accountable for risks, for example a low data quality. Specifically, the ability to judge data quality was mentioned by most respondents as being essential to evaluate if data can actually be used. Knowing who the data owner is, how the maintenance of the data is deployed and who is responsible for the quality of the data are success factors when innovating with open data. Another factor related to data governance is to define rules on how to deal with new technological developments related to the data, e.g. cloud storage. Second, respondents rated data management, which comprises all the disciplines related to managing data as a valuable resource, as a very important IT capability when innovating with open data. It is important to specify how data are stored, manipulated and processed, in particular when open data

are combined with proprietary data for new services. Third, innovation often means that data re-users need to be able to access different datasets and link them in new ways. Hence, interoperability, semantic standards and system integration are important capabilities, or as one of the respondents indicated: “*without standards there is no integration*”. The ability to handle different open data standards (both open source and proprietary) is important to ensure that data can actually be reused, e.g. data is readable, actual and relevant. The availability of standards makes the reuse of open data faster and more efficient. Data re-users need to be familiar with and their information systems need to be compatible with the latest data standards of the data providers. These standards are domain specific and require domain knowledge.

Data analytics technologies, such as software and tools were found to be important but as they are widely (commercially) available, this factor is not considered a challenge anymore. The same holds for hardware and software which are seen as a prerequisite for data innovation but are also considered as given and thus not a challenge anymore to invest in. Of all IT capabilities, having the right hardware and software is deemed the least important. However, just having the right hardware and software does not guarantee that they are used to their full potential. Hardware and software becomes an important capability when privacy, security and ethical questions of the use open data start to arise. With respect to data security it is important that the infrastructure used ensures that the data is stored securely and that data access is in line with access policies. Last, an IT strategy and planning seems to be less relevant to our respondents, which might be biased due to the relatively high amount of small and innovative firms in our sample.

#### *4.2. Organizational Capabilities*

The organizational capabilities that were found to be most important by our respondents are entrepreneurial orientation (4.55, on a scale from 1 to 5) and an innovative organizational culture (4.55), followed by the organizational capability to develop new ideas (4.45). The organizational capability that was perceived by far as least important for data-driven innovation is the organization’s capability to efficiently allocate resources (3.36).

We divided the organizational capabilities into strategic, tactical, operational and cultural capabilities. Based on the questionnaire and interviews we conclude that the operational and cultural organizational capabilities are considered to be most important when innovating with open data.

Two third of our respondents indicated that an entrepreneurial orientation is very important and that, similar to other types of innovation, it comes down to entrepreneurship. This result can be influenced by the fact that the majority of our respondents are smaller, innovative and often more entrepreneurial organizations. To create value out of open data, organizations need teams that are capable of making links between different subjects, develop a network in and outside the organization, recognize opportunities that open data sources offer and show willingness to exploit these ideas. While this capability will come back later as an important skill, it also requires an organizational culture which supports innovation and entrepreneurial thinking by e.g. handling short lines between different management levels and working in small teams. Consequently, respondents indicated that organizations, in order to innovate, need to provide an environment where experimenting with data is encouraged, which is a creative process that may take a while before it creates value. People in

organizations with such a mindset think in terms of possibilities and not in terms of barriers. They are not afraid of potential hurdles and challenges but just get started. These organizations are often data driven, have a long term vision and have a culture where innovative teams can play around with data and failure is acceptable. It is best to start an open data innovation project with a small team that has a strong mission and believes in open data innovation. In time this team can take small steps and expand with more and more people that share the same mindset. Moreover, an innovative culture implies that the costs of the innovation e.g. data applications are positioned as costs for the business, not as costs for the IT department. Thus, making the business responsible for data innovation is essential for its success. As the context and challenges around the data might constantly change it is essential to have a learning, agile organization that can quickly adapt to changes in the provision of data.

Multidisciplinary teams are crucial in the open data innovation process, as organizations need people with both technical knowledge and people that understand the domain context of the data and the needs of potential customers. Teams need to understand the bigger picture of the data and be able to talk across boundaries of systems and organizations. Multidisciplinary helps teams to understand and translate technical complex concepts in something that is visually appealing.

Our respondents indicated that central leadership in itself is less important to innovate, but having mandate and budget from top management is very relevant for the success of an open data innovation. Although innovation initiatives and ideas are most likely to develop on business level it is very important to have a data driven vision at board level to move from an experimental phase to a next phase where innovations are further deployed and implemented. Our interviewees further mentioned that organizations should use successful examples as inspiration to stimulate the re-use of open data in their own organization. Central coordination of resources and activities as well as asset orchestration was found to be the least important in supporting the innovation. Most of the small innovative and entrepreneurial organizations in our sample did not recognize this capability as being important. However, coordinated actions might be valuable for organizations to match the supply and demand of data and stimulate the dialog between data owners and re-users. The management of an organization might want to play a catalyst role in this.

We find that the larger organizations in our sample find it important that the overall organization is ready and willing to support processes and decision-making based on open data. If an organization really wants to innovate with data it should let go of its classical decision making process and use the results of its data innovation to form its own decisions and policies.

Relationship management is a capability that was not initially considered based on our literature analysis and questionnaire, but based on the interviews the ability to engage with external stakeholders (e.g. end-users) is found to be important to be able to innovate. Often, the availability of open data sets is a first step, and close collaboration between the data provider and user is necessary to take the next steps and create value. As one of the respondents said: *“Do not think in organizational silos but look across organizational boundaries when innovating with data.”*

#### 4.3. Skills

The factors that got the highest scores and are found to be the most important skills are data analytics skills (4.67 on a scale from 1 to 5) and entrepreneurship (4.64) followed

by general e-literacy (4.45) and the capability of people to cooperate between disciplines (4.45). The factor that received the lowest score and was found to be less but still important when innovating with open data are data visualization and reporting skills (3.83). This shows that a successful organization or team consists of people with data analytics skills in combination with entrepreneurship skills and who are able to combine knowledge from several disciplines.

Regarding technical or hard skills, our respondents foremost stressed that it is important to *“employ people who can detect patterns in data, who are creative and have a high quantitative skills”*. Even though respondents valued a team with a diverse set of skills, most team members should have a basic understanding of data (e.g. experience with working with large databases) and good analytical skills to actually make technical sense out of the data. One respondent noted that *“it is more important to have mathematicians than having only programmers. The latter you can simply hire”*. General IT skills, such as programming, are still important when you innovate with open data: *“if you do not have your IT Skills in order, then you are not part of the game anymore”*. However, IT skills just qualify, but a deep understanding and creativity with data gives a competitive advantage. Consequently, open data re-users need employees that have R&D skills to experiment (or ‘play’ as respondents qualified this skills) and test how different data sources can provide value or insights about the problems they want to address. More specifically, employees need to think in large distributed systems and make complex data combinations (from both open and closed data source).

Regarding non-technical or soft skills, having employees who are team players is key. Respondents valued employers that are open to the ideas of others and who are willing to learn. These quick learners need to be good networkers with a mindset to create new ecosystem, that are able to set up new collaborations with new parties and defend and their business decisions taken based on data. Additionally, it is important to have people with the capacity to think in terms of customer needs. Business models based on open data are often still a challenge. For example, one of the respondents see applications based on open data as a showcase to get to know new customers and sell additional data services. Hence, it is important to start early in the process with the identification of customer needs and business potential.

## **5. Discussion**

Comparing the three types of capabilities we see that all were found to be important but accents differ. On average skills are valued the highest (4.27) closely followed by organizational capabilities (4.19). IT capabilities are with a score of 3.86 rated the lowest by our respondents. This may indicate that investing in technology and other IT related capabilities that allow you to process and analyze data may not be enough when you want to create value from the re-use of open data. It seems even more important to focus on skills and organizational capabilities. For example, a lack of employees with the right skills is the biggest barrier to open data innovation.

Although we address the three capabilities in isolation there are many interdependencies between the capabilities. Future quantitative studies with a higher number of respondents may conduct factor analysis to cluster capabilities into a more coherent set of open data innovation capabilities. In addition, some capabilities are present on multiple levels. For example, entrepreneurial orientation (organizational) and an entrepreneurial mindset (people) are very similar and probably overlap.

Respondents stress the importance of an emergent innovation strategy: you need to keep your goal in mind, but how and what IT resources are needed may change rapidly. Organizations need to know where they are going, however this does not need to be planned completely in an IT strategy. This finding may contradict with IS Strategy literature that emphasizes the importance of a well-planned IT strategy that are more important to large organizations. SMEs employ more emergent and flexible strategies. Furthermore, we find that frontrunners focus strongly on open innovation: networks are important in the innovation process to make optimal use of partners' competences and be able to complement each other in the open data innovation process.

While open data is in some aspects different from internal, proprietary organizational data some experts expect that the barriers related open data are largely the same. However, open data may be less sensitive to privacy risks than proprietary data. The value of open data may be the highest when an open data sets are combined with proprietary datasets. Such a combination can bring new insights which would not have been possible without open data.

The small number of organizations studied and the focus on the Netherlands is a limitation of our research which does not allow us to generalize the results to a larger amount of organizations in different national contexts. The surveys and interviews, however, provide interesting insights for future research on the re-use of open data. Future research may expand the number of cases in different countries over a longer period of time . A longitudinal study is required to study how dynamic capabilities may change over time, for example along the maturity of the data innovator.

## **6. Conclusion**

This paper aims to study which capabilities are required to innovate successfully with open data. We present an open data innovation capabilities framework based on a literature review on capabilities for data-driven innovation. Our analysis of 12 Dutch open data front runners reveals that for innovation with data to happen it is not enough to just make more and more open data available by government agencies and focus on IT capabilities at the side of data re-users.

Theoretically, we contribute by offering a capabilities perspective on open data innovation that complements earlier studies on the barriers of data availability and re-use [2-5]. As data-driven innovation seems to require entrepreneurship, open data research may benefit from theories in entrepreneurship literature, such as dynamics capabilities or causation/effectuation, to better explain the value creation process [27]

Practically, digital skills and digital entrepreneurship policies are as important as persuading governments to open-up data. Policy makers may need to set up stimulating programs that are aimed at educating more employees and entrepreneurs with the right set of skills. Organizations aiming to create more value out of open data may follow the following lessons from frontrunners: 1) Set up a multi-disciplinary team with motivated and creative employees, give this team the mandate to experiment with data, and let them formulate an emergent data strategy and get top management support to create a stable innovation environment. 2) Take an entrepreneurial approach: Recognize and exploit opportunities and strive for an open and experimental culture. When successful, the start-up team may motivate and engage more teams within the organization. 3) Think outside organizational boundaries and silos. Set up new inter-organizational networks with complementing skills and experience. The social interaction between

employees from a variety of organizations and domains may inspire innovation. Furthermore it is essential that data users and suppliers keep discussing their data needs and ensure that open data is more reliable and usable for external stakeholders, decreasing the risks for data innovators regarding data quality and availability.

## References

- [1] Manyika, J., *Open data: Unlocking innovation and performance with liquid information*. 2013: McKinsey.
- [2] Huijboom, N. and T. Van den Broek, *Open data: an international comparison of strategies*. European journal of ePractice, 2011. **12**(1): p. 4-16.
- [3] van Veenstra, A.F. and T.A. van den Broek, *Opening moves—drivers, enablers and barriers of open data in a semi-public organization*, in *Electronic Government*. 2013, Springer. p. 50-61.
- [4] Zhang, J., S.S. Dawes, and J. Sarkis, *Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing*. Journal of Enterprise Information Management, 2005. **18**(5): p. 548-567.
- [5] Zuiderwijk, A., et al., *Socio-technical impediments of open data*. Electronic Journal of e-Government, 2012. **10**(2): p. 156-172.
- [6] IODC, *Enabling the Data Revolution: An International Open Data Roadmap*, 2015.
- [7] Attard, J., et al., *A systematic review of open government data initiatives*. Government Information Quarterly, 2015. **32**(4): p. 399-418.
- [8] LaValle, S., et al., *Big data, analytics and the path from insights to value*. MIT sloan management review, 2013. **21**.
- [9] Jetzek, T., M. Avital, and N. Bjorn-Andersen, *Data-driven innovation through open government data*. Journal of theoretical and applied electronic commerce research, 2014. **9**(2): p. 100-120.
- [10] Venkatraman, N.V., et al., *Theorizing Digital Business Innovation: Platforms and Capabilities in Ecosystems*. Available at SSRN 2510111, 2014.
- [11] Volberda, H., *Concurrentie en Innovatie Monitor*, 2005, Erasmus University Rotterdam.
- [12] Hemerly, J., *Public policy considerations for data-driven innovation*. Computer, 2013. **46**(6): p. 25-31.
- [13] Norris, D.M. and L.L. Baer, *Building organizational capacity for analytics*. Retrieved from EDUCAUSE website: <http://www.educause.edu/library/resources/building-organizationalcapacity-analytics>, 2013.
- [14] Cosic, R., G. Shanks, and S.B. Maynard, *A business analytics capability framework*. Australasian Journal of Information Systems, 2015. **19**.
- [15] Jia, L., D. Hall, and J. Song, *The Conceptualization of Data-driven Decision Making Capability*. 2015.
- [16] Sharma, R., S. Mithas, and A. Kankanhalli, *Transforming decision-making processes: a research agenda for understanding the impact of business analytics on organisations*. European Journal of Information Systems, 2014. **23**(4): p. 433-441.
- [17] Knabke, T. and S. Olbrich, *Exploring the Future Shape of Business Intelligence: Mapping Dynamic Capabilities of Information Systems to Business Intelligence Agility*. 2015.
- [18] Kung, L., et al., *Managing Big Data for Firm Performance: a Configurational Approach*. 2015.
- [19] Kusiak, A., *Innovation: A data-driven approach*. International Journal of Production Economics, 2009. **122**(1): p. 440-448.
- [20] Kamioka, T. and T. Tapanainen, *ORGANIZATIONAL USE OF BIG DATA AND COMPETITIVE ADVANTAGE—EXPLORATION OF ANTECEDENTS*. 2014.
- [21] Chang, Y.-C., et al., *How do established firms improve radical innovation performance? The organizational capabilities view*. Technovation, 2012. **32**(7): p. 441-451.
- [22] Kabir, N. and E. Carayannis. *Big Data, Tacit Knowledge and Organizational Competitiveness*. in *Proceedings of the 10th International Conference on Intellectual Capital, knowledge Management and Organisational Learning: ICICKM 2013*. 2013. Academic Conferences Limited.
- [23] Knight, G.A. and S.T. Cavusgil, *Innovation, organizational capabilities, and the born-global firm*. Journal of International Business Studies, 2004. **35**(2): p. 124-141.
- [24] Mitrovic, Z., *Building Open Data Capacity through e-Skills Acquisition*.
- [25] Davenport, T.H. and D. Patil, *Data scientist*. Harvard business review, 2012. **90**: p. 70-76.
- [26] Maruyama, H., et al., *Developing Data Analytics Skills in Japan: Status and Challenge*. 日本経営工学会論文誌, 2015. **65**(4E): p. 334-339.
- [27] Bharadwaj, A., et al., *Digital business strategy: toward a next generation of insights*. Mis Quarterly, 2013. **37**(2): p. 471-482.