REVIEWED

ELEMENTS OF A SUCCESSFUL BIG DATA HACKATHON IN A SMART CITY CONTEXT

Thorhildur Jetzek
Department of IT Management,
Copenhagen Business School
tj.itm@cbs.dk

This paper presents a case study of a Smart City initiative in Lyngby-Taarbæk municipality, which has successfully applied the triple helix model to create an informal collaboration between academia, government and private industry. The study recounts how a group of university students, participating in a big data hackathon, managed to create a Smart City solution prototype based on open data in only 48 hours. The solution offers to make the municipality more cost efficient and improve citizen services, while simultaneously contributing to reduced CO_2 emissions, thus addressing a difficult societal challenge. A special attention is paid to how the Smart City vision, based on the triple helix model, is used to align interests and enable an informal collaboration between heterogeneous stakeholders. This collaboration represents an underlying value network, where value generation is moving beyond the simple profit-driven mechanisms of the markets. The paper identifies three main roles in the triple helix based value network: The Influencer, the Facilitator and The Implementer.

Keywords: Smart city, big data hackathon, innovation contest, energy efficiency, spatial data

INTRODUCTION

For the past two decades, information and communication technologies (ICTs) have been exerting a growing influence on the nature, structure and enactment of urban infrastructure, management, economic activity and everyday life (Kitchin, 2014). This has led to a growing interest in the concept of Smart City. The Smart City concept can be viewed as an overarching concept that describes a city's ability to use data and technology for improving the livability and wellbeing of its citizens[1]. Concurrently, there has been an increasing focus on societal challenges that are reflected in our societies' inability to sufficiently address complex problems, such as the refugee crisis and climate change (OECD, 2011). I propose that Smart City initiatives based on informal collaboration between stakeholders in different sectors offer a new model for solving these grand challenges. The key to success is a Smart City project's ability to encourage and activate more members of society to collectively address societal challenges. History tells us that silo structures, which oftentimes characterize governmental organization, are poorly suited to tackling complex problems across sectors. Moreover, the market by itself lacks the incentive structure and appropriate business models needed to solve societal challenges. And stakeholders with interest and drive, such as civil society or universities, might lack the capital, skills and resources to take promising ideas to scale (Murray et al., 2010).

In order to successfully address societal challenges, it is necessary for all of these stakeholders to leverage their individual strengths and capabilities. However, in order to incentivize a diverse group of people to collaborate on finding and implementing solutions, it must be acknowledged that their motivations and goals may vary widely. In this paper, I study a case where a loosely organized collaboration between different stakeholders and sectors has succeeded in enabling individual participants to create and capture value, while simultaneously addressing a societal challenge, namely climate change. The case data is based on 5 interviews with participants from different sectors, including follow-up; as well as analysis of online content and documents provided by the interviewees. A list over organizations interviewed is provided in Appendix A. The case context is that of a Smart City, however the case includes many other new and interesting concepts such as big data, innovation contests and open government data. The case offers insight into how different motivations can be aligned through the triple helix model, i.e. how to motivate and enable heterogeneous stakeholders to collectively contribute to a common goal. Moreover, I discuss how value can be created in a value network, moving beyond the simple profit-driven mechanisms of the markets towards a complex network of aligned interests.

URBANIZATION AND THE SMART CITY

Urbanization, the demographic transition from rural to urban, is associated with shifts from an agriculture-based economy to mass industry and more recently, technology and service. If these trends continue as projected, the percentage of people living in urban areas will increase to 70% before 2050[2]. The trajectory of the rapid urban population growth is not just an interesting fact but also requires a demanding imperative for sustainable development and better livability (Nam and Pardo, 2011). As an example, although cities currently occupy less than two percent of the landmass of the earth, urban residents consume over 75% of the world's natural resources and are primarily responsible for green-house gas emissions (Marceau, 2008). Urbanization is also changing how we need to approach problems. Multiple diverse stakeholders are now sharing a physical space, which results in high levels of interdependence, competing values, and social and political complexity (Dawes et al., 2009; Weber and Khademian, 2008).

[1] http://tti.tamu.edu/group/transit-mobility/files/2013/05/3-Definitions-of-livability-handout.pdf
[2] http://www.who.int/gho/urban_health/situation_trends/urban_population_growth_text/en/

Making a city smart is a novel way to approach such challenges (Nam and Pardo, 2011). But what is a Smart City? Bolici and Mora (2015) define Smart Cities as urban areas in which information and communication technologies (ICTs) are used to solve their specific problems and support their sustainable development in social, economic and/or environmental terms. The Lyngby-Taarbæk City of Knowledge initiative defines Smart Cities as digital and inclusive cities that seek to optimize how the city functions by creating synergies between the physical and the social in the digital space. According to their definition, a Smart City should support relationships between authorities, businesses, organizations and citizens, mainly through sharing of data and information across organizational boundaries. Wikipedia[3] offers this definition: A Smart City uses digital technologies or information and communication technologies (ICT) to enhance quality and performance of urban services, to reduce costs and resource consumption, and to engage more effectively and actively with its citizens. It is safe to say that a Smart City is an emerging phenomenon and as such has no precise definition. However, what all of these definitions have in common is a focus on the digital space and how new technologies and new means of collaboration can facilitate and accelerate how we address many of the societal challenges that result from increased urbanization.

THE CITY OF KNOWLEDGE: LYNGBY-TAARBÆK'S SMART CITY VISION

There are a number of Smart City initiatives in Denmark. Perhaps the most prominent one is Copenhagen Smart City Initiative which has won awards like the World Smart Cities Award in 2014. Additionally, various smaller municipalities have started their own Smart City initiatives, although some of them might not explicitly use the Smart City concept. One of them is Lyngby-Taarbæk municipality. Lyngby-Taarbæk is a host to many technology and information driven companies as well as one of the most respected technical universities in Europe, The Technical University of Denmark, DTU[4]. The municipality has identified Lyngby-Taarbæk as a City of Knowledge & Urban Development. Their City of Knowledge vision includes attracting and retaining knowledge-based businesses, developing Lyngby-Taarbæk into a university town, creating urban life, forming networks, furthering social innovation and internationalization, inspiring entrepreneurship, and broadening municipal services to the businesses and citizens in Lyngby-Taarbæk[5].

The City of Knowledge initiative is designed as a triple helix model (Etzkowitz, 1993; Etzkowitz and Leydesdorff, 1995; Ranga and Etzkowitz, 2013). The Triple Helix thesis is that the potential for innovation and economic development in a Knowledge Society lies in the hybridisation of elements from academia, industry and government to accelerate production, transfer and application of knowledge. The City of Knowledge & Urban Development includes stakeholders from all three sectors and is governed by an independent organization that is jointly funded by all of the sectors. The participating stakeholders all agree on the common vision for the City of Knowledge & Urban Development, and presumably expect to benefit from this collaboration. However, their motivations for collaborating vary considerably. The key to success in the triple helix model is to create a win-win-win situation where each of the partners can focus on their own benefits while their individual contributions will add value to the larger ecosystem in which they operate. Figure 1 shows an example of a triple helix model. Academia mainly contributes through knowledge creation which is disseminated through teaching and research. Government contributes to a healthy environment for innovative collaboration, creating

[3] https://en.wikipedia.org/wiki/Smart_city

[4] DTU is listed number 43 in Thomson Reuter's list of the World's most innovative Universities, and counts number 7 of all the European Universities on the list. See: http://www.reuters.com/article/2015/09/15/idUSL1N11K16Q20150915

[5] http://www.vidensby.dk/English.aspx

policy and supplying necessary services, e.g. through funding organizations or open data platforms. Industry contributes through wealth generation, and provides the capital and work processes, necessary for scaling up promising ideas and introducing them to markets.

The Big Data Hackathon

The independent City of Knowledge & Urban Development organization governs and facilitates different networks where members develop ideas and common projects. One of these networks has a focus on climate and green technology. Network members showed an interest in gaining improved access to Lyngby-Taarbæk's data for supporting the development of data-driven smart city solutions, such as Intelligent Energy Systems. The idea to support an innovation contest, or big data hackathon, was originally raised by DTU Compute department but the City of Knowledge agreed to partner in the organization of the event, together with representatives from Lyngby-Taarbæk municipality and IBM, which provided IT tools to the participants. Lyngby-Taarbæk municipality agreed to give the hackathon participants access to some of their data, as well as providing a description of some of the problems or challenges the municipality was faced with, in a hope for a potential (partial) solution. The hackathon was hosted by DTU Compute in the new DTU Skylab building on the18th and 19th of November 2014. Simultaneously, DTU hosted a big data conference where the prizes were to be awarded. The first three winning solutions were to get prizes of a total of DKK 55.000, which were sponsored by Danske Bank, a private company in the municipality. Moreover, the EU climate innovation initiative, Climate-KIC, contributed a special prize of DKK 10.000 for the idea providing the most climate friendly solution.

Invitations were extended to university students in various Danish universities, mainly through Facebook sites and student organizations. In short, the hackathon was a success with 65 participants

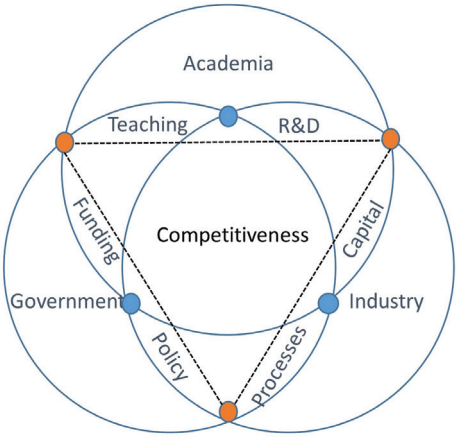

Figure 1. The Triple helix model. Loosely adapted from Farinho and Ferreira, 2013.

and provided many interesting solutions. Interestingly, an emerging literature on innovation contests in the open data literature has shown that such contests are in many cases poorly attended and do not produce sustainable solutions (Hjalmarsson et al., 2014). However, for this hackathon, the results were considered as a huge success by all participating stakeholders. Thus, I have attempted to extract the potential success factors of this hackathon from the interview data:

- There was an introductory meeting where students could show up and form teams. A positive result of this event was that the meeting gave the students a chance to meet others with complementary qualifications and the resulting teams offered more diversity of knowledge and skill.
- The municipality not only contributed data but also formalized some questions or problems they were facing. An overarching theme was to create a solution which would make the lives of the citizens in Lyngby-Taarbæk easier and contribute to a more sustainable environment (special prize). This gave the participating students, who had little or no prior knowledge of the societal challenges faced by municipalities, a starting point from where to develop their solutions.

- The technical and business related requests for the solutions also helped the students think more broadly in terms of future applications. The solutions were required to make use of big data, have a novelty value, be user-friendly, scalable and have commercial potential.
- The students were somewhat motivated by the cash prize but even more motivated by the fact that prominent members of industry were a part of the panel of judges. Other industry stakeholders were supporting the hackathon with IT solutions and prizes. As many of the participating students were just about to finish their studies, they needed industry contacts.

From Hackathon to a Startup Company

The winning team consisted of 6 individuals, 2 with computer science skills, 3 with mathematical modelling and machine learning skills (all five from DTU) and one student from CBS with a business administration background. They had an opportunity to meet once before the hackathon to brainstorm, but all of the real work happened in the 48 hours of the hackathon itself. In the following, the solution itself is described.

Lyngby-Taarbæk provided a number of data sources from different departments in the municipality. The choice of data was more or less ad hoc, based on which data could easily be provided. The winning team arrived to the contest with a semi-structured idea in mind from the brainstorming meeting. They started by looking at data on buildings owned by the municipality and thermographic images of houses in Lyngby-Taarbæk. Next, they created a program that could link the addresses of buildings owned by the municipality (provided in an excel spreadsheet) to a geolocation. From this geolocation they could link the addresses to the map of thermographic images and see which houses were losing most heat. They could also use the geolocation to connect these data to the Danish elevation model, which is provided as open data by the Geodata Agency (Geodatastyrelsen). The geolocation thus has a very important function as a key identifier, making diverse sets of data interoperable.

Having access to the property data gave them information about the age of the building and from that they extrapolated the type of insulation in different houses. From the thermal images they could draw conclusions on the relationship between the insulation and how well the house retained heat. Based on (openly available) data from several providers of insulation material they could calculate the potential cost of insulating an old house to a modern standard. They looked at (provided but closed) data on heating sources and expenses for the properties owned by Lyngby-Taarbæk, and from combining all these data, they could deduce how cost-efficient it would be to insulate different houses and the magnitude of possible environmental effects (reduced CO_2). The interesting thing about how these students approached this task is that they did not only utilize a single dataset provided by the municipality as has been shown to be the case in many open data applications, but rather combined the datasets provided with openly available data from other sources.

Afterwards, the winning team calculated for each property whether or not it would be cost-efficient to implement solar panels. For this they used the elevation model to find the angle and orientation of the roof, information about yield based on angle and orientation (from various sources), open data on yearly solar radiation from Danish Meteorological Institution (DMI) and available information from different solar panel vendors (prices pr. m2, efficiency pr. m2, efficiency guarantees etc.). From their knowledge about roof sizes (provided open data), energy costs (open data) and composition of energy sources (provided closed data), they could also calculate the eco footprint for individual houses. As the group had access to data on

Figure 2. The winning team (Source: www.DTU.dk)

current energy sources for the municipality´s own buildings, they could present a solution that could make the municipality more cost- and energy efficient. While they did not have such detailed data for all the privately owned property in the municipality, they could calculate the energy efficiency of solar panels based on the roof size and direction and then calculate estimated energy savings. Hence, the solution delivers openly available content, which can help the citizens of Lyngby-Taarbæk municipality make informed decisions about how to influence their own energy costs and eco-footprint. The solution was simultaneously addressing the need for more cost efficient municipality, the need for improved citizen services and the ability to improve energy efficiency and reduce CO_2 emissions.

One of the sponsors of this contest was Climate-KIC, EU's main climate innovation initiative. Climate-KIC has an acceleration program for entrepreneurs in Denmark and their representative encouraged the winning team to apply for funding so they could develop their ideas further. This is a very important element for further development of data-driven products. The open data literature shows that many of the solutions that have been developed in open data innovation contests are not sustained, in the sense that they fail when it comes to scaling up and developing the solution for the market. A suggested reason for this is that the public sector participants that often plan and execute such contests do not have the capabilities or the funds required to function as innovation incubators and/or accelerators. While Lyngby-Taarbæk municipality has committed some funding to further the development of the product for use in the municipality, the winning team's dream was to develop the solution further and make it ready for general marketing. However, such development requires funding and support. The winning team founded a company, Picodat, and continue to develop their solution. They are currently working on a more general solution which can be marketed for other municipalities in Denmark and hopefully later in Europe as well.

DISCUSSION AND FINDINGS

In this section, I discuss some of the main findings that emerged from the case data analysis.

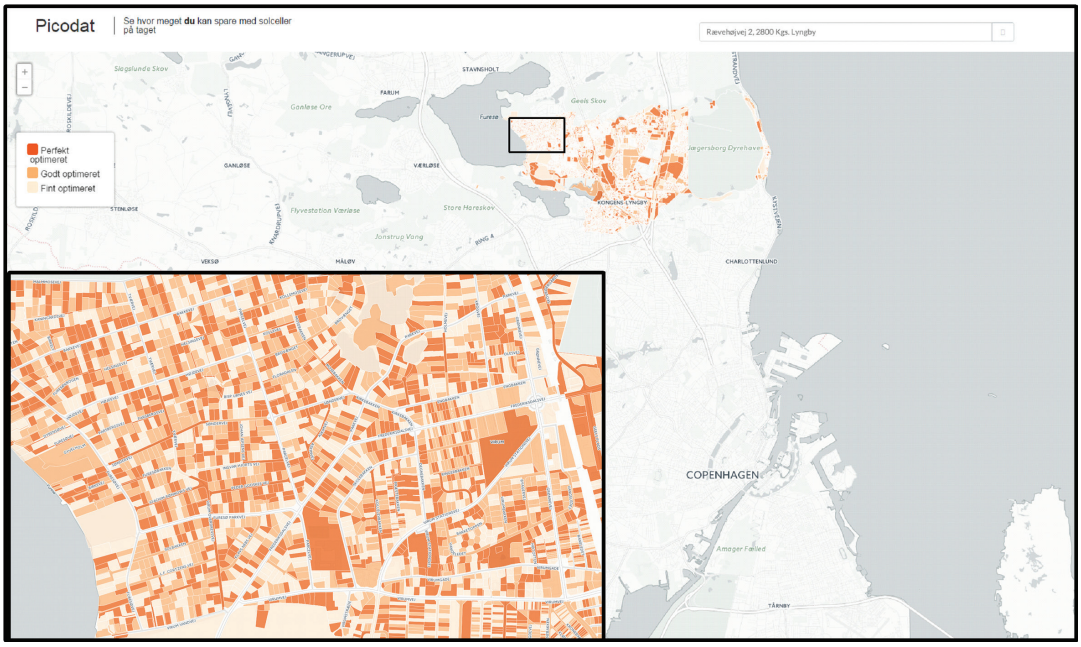

Figure 3. A screenshot of the winning solution. (Source: Picodat)

Different roles of stakeholders in the triple helix model

The idea that the university, industry and government are relatively equal interdependent and interacting institutional spheres is the basis of a triple helix society (Etzkowitz et al., 2007). However, these spheres are not only autonomous but overlapping, not entirely distinct but not completely merged either (ibid). Thus, I conceptualize three stakeholder roles, moving away from the institutions themselves and their roles in society, and towards the individual stakeholders that represent these organizations within a triple helix collaborative initiative. Doing so, I propose, will provide an extra layer to the triple helix model, representing a value network of aligned interests, where roles can be switched or spheres can provide more than one role, which can explain how they overlap.

Academia

The academic partner in this particular triple helix model was the stakeholder that originally came up with the idea of a big data hackathon. The University's organizational role is to do research and to disseminate knowledge to society. Accordingly, they like to test some of their new ideas and methods with real data and applications. DTU acted as the thought leader or the driver behind the hackathon, mostly to raise awareness of how data and data science could contribute to society. While the other partners (from the municipality and industry) did not previously have any structured data-related initiatives, they were happy to go along with ideas and initiatives leading in this direction. Thus, in this triple-helix constellation, the academic stakeholder has the role of *Influencer*.

Government

Lyngby-Taarbæk does not yet have an open data strategy or a specific open data initiative, so they did not function as influencers in this particular triple helix setup. However, the municipality was willing to experiment and provided access to data in the hackathon and information on problems in need of solving. Moreover, the role of Lyngby-Taarbæk's City of Knowledge organization as a coordinator between the different stakeholders was very important for keeping all the partners aligned. For this case, I

propose that government acted as the *Facilitator* as their contribution was important for creating the right environment, including aligning the incentives of different participants.

Industry

The presence of industry in the panel of judges in the hackathon itself, as well as industry's contribution to winning prizes, clearly created an incentive for the students to participate. However, industry played a larger role in the development of the final product. The panel of judges contributed important knowledge regarding commercial potential and scalability of the solutions presented. Moreover, Climate-KIC ultimately provided the funding necessary to take the idea to the next level, contributing to a sustainable solution[6]. I propose that industry played the role of *Implementer*. Hjalmarsson et al. (2014) argue that only a limited number of results from contests successfully reach the end user market. Having implementers on board increases the chance of promising ideas being implemented in practice, thus, this role might have been missing in some earlier open data hackathons that did not provide sustainable solutions. Moreover, after a company is founded (in this case Picodat), the participants in the hackathon become Implementers themselves.

Value network

Through this case study I want to contribute to knowledge on how a constellation of heterogeneous partners in a Smart City context can collectively generate new value from existing data. One of the findings is that different stakeholders in a triple-helix constellation not only have different roles, they are also differently motivated. The academic stakeholders were interested in stimulating interest in big data, in order to further research, develop new methods and contribute to knowledge. Moreover, they were interested in getting access to more open government data, and perceived the hackathon as a potential venue to raise awareness to this issue. The stakeholders from the municipality were interested in seeing a practical example that could demonstrate how their own (siloed) data could be used more effectively, for increased efficiency and improved services. The stakeholder from Climate-KIC was primarily motivated by the prospect of supporting solutions that could contribute to reducing CO_2. Other industry sponsors were motivated by having access to future talent or present their products and/or services. The participants in the hackathon were mainly university students. While cash prizes and just having fun were most likely strong motivational factors, some of them were motivated by the prospect of getting industry contacts and others by their wish to start their own company. The members of the City of Knowledge & Urban Development were motivated by the potential of a successful outcome, which could also promote Lyngby-Taarbæk as a Smart City.

Interestingly, while different stakeholders exhibited different motivations and drivers, they collectively addressed a societal challenge through the hackathon, i.e. climate change. As this wasn't the primary goal of any of the stakeholders besides Climate-KIC, this finding is presented as evidence of the usefulness of such a triple helix setup for creating an environment where complex societal challenges can addressed through synergies that arise when strengths of individual sectors are combined. The City of Knowledge and Urban Development has created an environment where the interests of different stakeholders with different motivations are successfully aligned, ultimately creating a win-win-win situation, which made the resulting outcome possible. Ultimately, all of the interviewed stakeholders shared the notion that the success of the winning team, Picodat, equaled their own success.

[6] It might be controversial to include Climate-KIC in the Industry category as they are a PPP which include industry partners, academic partners and public/not-for-profit organizations, thus representing a triple helix setup on their own. However, as they are 50% business, 30% academic and 20% public and not-for-profit, they are included with industry. http://www.climate-kic.org/about/how-we-are-organised/

Other findings related to use of open/big data

Research has pointed out that there are five main dimensions that contribute to the state of openness of individual datasets (Jetzek, 2015). These are: strategic dimension (availability), economic dimension (affordability), legal dimension (reusability), conceptual dimension (interoperability) and technical dimension (usability, accessibility and discoverability). In the case of Picodat, the availability dimensions is quite important as the team could find a number of available datasets online that were not provided by the Hackathon. The same goes for affordability, it would have been a barrier if they had been forced to pay for access to these data. As for the other dimensions, Picodat did not comment on open licenses or lack thereof. However, their dependence on open licenses might increase when they start to commercialize their solution. Interoperability between heterogeneous datasets did not seem to be a barrier in this case either, as all the different datasets were linked through the geolocation, which functioned as a common identifier. The technical dimension did not seem to be very important for this prototype work, although some of the data that were used were discovered through web searches and therefore depended on the discoverability of the respective data sources. Some of the data weren't provided or available in very user friendly formats, however this did not discourage Picodat from using them. Admittedly, this sentiment might change when they try to scale up their solution and make it more re-usable across different municipalities. Moreover, it should be noted that Picodat had direct access to the custodians of most of the data, which might not be the case when they develop the solution for other countries.

Picodat made good use of all the talent in their team and their ability to use complex mathematical modelling is considered as an important factor in the success of their solution. Moving beyond the mashing-up of different sets of data, they utilized calculations that require some in-depth knowledge of concepts such as energy efficiency. Moreover, they used a variety of available information to draw conclusions (increase their knowledge) about the cost-efficiency of different approaches. The team needed to pitch their solution to the committee of judges and explain why it had potential to generate value for the municipality. Besides technical skills, they also needed, and made use of, business perspectives. It is encouraging to see how the students managed to capitalize on the diversity of their group and utilize this diversity in their efforts to generate a solution that is both easy to understand but at the same time quite sophisticated. Hopefully their solution will not only help reduce CO_2 emissions in Lyngby-Taarbæk, but all around Europe in the future. I personally hope that this will be one of many initiatives that will drive more open access to an increasing number of data sources, which can later be used to address societal challenges through improved information dissemination and scientific knowledge, as well as commercial products and services.

CONCLUSION

The case of Picodat is a case of a successful hackathon that resulted in a new big data startup company and a solution that offers a potential for Lyngby-Taarbæk to increase their own energy efficiency and improve citizen service. Moreover, the solution contributes to the important goal of addressing climate change by reducing CO_2 emissions. The City of Knowledge and Urban Development managed to align the interests of different stakeholders through use of the triple-helix model, despite quite different motivations and goals. In this case I have identified three stakeholder roles for the triple helix model: The Influencer, the Facilitator and the Implementer. For future research, it could be interesting to analyze and compare successful and unsuccessful big data hackathons and search for existence of these different roles.

REFERENCES

- Bolici, R. & Mora, L. (2015). *Urban regeneration in the digital era: how to develop smart city strategies in large European cities.* TECHNE: Journal of Technology for Architecture and Environment, 5(2) [in press]
- Dawes, S. S., Creswell, A. M., & Pardo, T. A. (2009). *From "need to know" to "need to share": Tangled problems, information boundaries, and the building of public sector knowledge networks.* Public Administration Review, 69(3), 392-402.
- Etzkowitz, H. (1993). *The Triple Helix: A North American Innovation Environment.* Available: http://taisurpjoe.tripod.com/NIS-PDF/America3.html
- Etzkowitz, H., & Leydesdorff, L. (1995). *The Triple Helix--University-industry-government relations: A laboratory for knowledge based economic development*. Easst Review, 14(1), 14-19.
- Etzkowitz, H., Dzisah, J., Ranga, M., & Zhou, C. (2007). The triple helix model of innovation: University-industry-government interaction. Asia Pacific Tech Monitor, 24(1), 14-23.
- Jetzek, T. (2015). *Managing Complexity across Multiple Dimensions of Liquid Open Data: The Case of the Danish Basic Data Program*. Government Information Quarterly, in Press: doi:10.1016/j.giq.2015.11.003
- Kitchin, R. (2014). *The real-time city? Big data and smart urbanism.* GeoJournal, 79(1), 1-14.
- Marceau, J. (2008). *Introduction: Innovation in the city and innovative cities*. Innovation: Management, Policy & Practice, 10(2-3), 136-145.
- Nam, T., & Pardo, T. A. (2011). *Conceptualizing smart city with dimensions of technology, people, and institutions.* In Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times (pp. 282-291). ACM.
- Murray, R., Caulier-Grice, J. & Mulgan, G. (2010). *The Open Book of Social Innovation.* Available: http://www.nesta.org.uk/publications/open-book-socialinnovation
- OECD: *Fostering Innovation to Address Social Challenges.* Workshop Proceedings (2011), http://www.oecd.org/sti/inno/47861327.pdf
- Ranga, M., & Etzkowitz, H. (2013). *Triple Helix systems: an analytical framework for innovation policy and practice in the Knowledge Society.* Industry and Higher Education, 27(4), 237-262.
- Weber, E. P., & Khademian, A. M. (2008). *Wicked problems, knowledge challenges, and collaborative capacity builders in network settings.* Public Administration Review, 68(2), 334-349.

Conducted interviews

- Interview 1: CEO, Picodat
- Interview 2: Project Manager, Lyngby-Taarbæk City of
- Knowledge and Urban Development
- Interview 3: Entrepreneurship Lead, Climate KIC
- Interview 4: CIO, Lyngby-Taarbæk Municipality
- Interview 5: Associate Professor, DTU (Follow-up also included a professor at DTU)