TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Software Science

Keegan David Braun McBride IVGM156414

GOVERNMENT AS A PLATFORM: EXPLOITING OPEN GOVERNMENT DATA TO DRIVE PUBLIC SERVICE CO-CREATION

Master's Thesis

Supervisor: Robert Krimmer
Prof. Dr.
Co-Supervisor Maarja Toots
MA

Tallinn 2017

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond
Tarkvarateaduse instituut

Keegan David Braun McBride IVGM156414

VALITSUS KUI PLATVORM: AVALIKU SEKTORI AVAANDMETE KASUTAMINE AVALIKE TEENUSTE KOOSLOOMEKS

Magistritöö

Juhendaja: Robert Krimmer
Prof. Dr.
Kaasjuhendaja: Maarja Toots
MA

Tallinn 2017

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Keegan David Braun McBride

04.05.2017

Abstract

New technologies are changing our current understanding of public services. One example of this is the emerging concept, and exploitation of, Open Government Data (OGD). Governments are able to release OGD and, through this action, act as a platform. This Open Government Data Platform allows anyone to use, exploit, and analyze government datasets to co-create new and innovative services which provide public value and empower communities in combination with multiple stakeholders. The aim of this thesis is to explore this new phenomenon and attempt to gain a better understanding of the process in which stakeholders are able to use OGD to co-create these new public services. An exploratory case study is conducted on an ongoing pilot project within Estonia which is co-creating a new public service based on OGD. The case seems to show that in order for OGD driven public service co-creation to occur effectively a new understanding of the role of stakeholders is needed, and that when governments release OGD and act as a platform they inherently become involved in the co-creation of new public services, even if this is not the goal. As a result of the research a general architecture for a co-creation OGD driven public service web application is also derived and presented.

This thesis is written in English and is 47 pages long, including 6 chapters, 3 figures and 3 tables.

Annotatsioon

Valitsus kui platvorm: avaliku sektori avaandmete kasutamine avalike teenuste koosloomeks

Uued tehnoloogiad muudavad meie arusaama avalikest teenustest. Üks niisugune tehnoloogiline innovatsioon on avaandmed, mis avaliku sektori kontekstis tähendab valitsuse andmete avaldamist masinloetaval ja taaskasutataval kujul. Avaldades avaliku sektori andmeid avaandmetena, on valitsustel võimalik toimida platvormina. Selline avaandmete platvorm võimaldab igaühel kasutada ja analüüsida avaliku sektori andmeid, et luua eri huvipoolte koostöös uusi ja innovaatilisi teenuseid, mis edendavad ühiskondlikku hüve ja võimestavad kogukondi. Magistritöö eesmärk on uurida seda uut nähtust ja mõista protsessi, mille kaudu huvipooled saavad rakendada avaliku sektori avaandmeid uute avalike teenuste koosloomeks. Selleks viiakse läbi kirjeldav juhtumianalüüs, mille keskmes on avaandmete-põhise avaliku teenuse koosloomega tegelev pilootprojekt. Uuritud juhtum paistab viitab, et tulemuslik koosloome nõuab uut arusaama huvipoolte rollist avalike teenuste loomes ning et avaandmeid avaldades ja platvormina toimides muutub valitsus loomupäraselt avalike teenuste koosloome osapooleks, isegi kui see ei ole olnud valitsuse eesmärk. Uurimuse tulemusena pakub magistritöö välja avaandmete-põhiste avalike teenuste veebirakenduste koosloomet toetava üldise arhitektuuri.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 47 leheküljel, 6 peatükki, 3 joonist, 3 tabelit.

List of abbreviations and terms

API	Application Programming Interface
CSS	Cascading Style Sheets
CSV	Comma Separated Values
EC	European Commission
EU	European Union
EREPP	The Estonian Real Estate Pilot Program
GaaP	Government as a Platform
H2020	Horizon2020
HTML	Hyper Text Markup Language
ICT	Information and Communication Technologies
IT	Information Technology
JS	JavaScript
MKM	Ministry of Economic Affairs and Communications
NPG	New Public Governance
NPM	New Public Management
OECD	Organization for Economic Cooperation and Development
OGD	Open Government Data
OGI	OpenGovIntelligence
PAM	Public Administration and Management
PIA	Public Information Act
PS	Public Service
PSD	Public Service Delivery
PSO	Public Sector Organizations
TTÜ	Tallinn University of Technology
UK	United Kingdom

Table of contents

List of figures

List of tables

1 Introduction

In a recent report, titled “A vision for public services”, the European Union outlined their understanding of the future of public services, how they are designed, the importance of technology, and the importance of inclusivity of citizens in the design and implementation process [1]. The report states that technology has empowered citizens, increased their connectivity and interactions with government, and led to innovation within the public sector. The Organization for Economic Cooperation and Development (OECD) has also released similar writings in the realm of new public services, the role of technology, and citizen involvement [2]. In a 2011 report titled “Together for Better Public Services” the OECD finds that technology allows for new innovations in public services due to the increased ability for citizens to play an active part in the design and implementation of the service. These writings seem to be in line with current scholarly literature, which also proclaims that public services are changing, Information Communication Technologies (ICTs) are helping to drive this change, and that ICTs are empowering citizens, which may lead to higher rates of innovation in the public sector [1]–[5].

If ICTs are, in fact, changing our understanding of how public services are created and delivered the current theory for understanding this process must be evaluated. In the past, one of the main theories for examining this process was New Public Management (NPM) [3], [6]. However, NPM is also heavily criticized [3], [7]–[9], and many new theories have risen to try to create a new theoretical paradigm for understanding public administration systems, public services, and public service delivery in today’s ICT driven world [10]. One theory which seems to address the shortcomings of NPM, and acknowledges the importance of ICTs in society, is New Public Governance (NPG). At the foundational level, NPG can be said to be based on the following: network theory, organizational sociology, service management, and coproduction [11]. NPG claims that public services should be viewed as public service delivery systems [6], [12], this understanding differs from previous transaction or top-down based understandings [12]. Furthermore, NPG believes that there needs to be a focus on the relationships among members or actors within the public service delivery system, this focus leads to a stronger importance being assigned to the relationships within the organization and among stakeholders [5], [10], [13].

If it is true that public service delivery is changing, and that there will be increased citizen involvement in the creation of public services as outlined previously, then NPG does seem to be well suited for the task of explaining and analyzing this change and the role in which ICTs may play. At this stage, a literature review was conducted, and what became clear was that there are two major areas where research gaps exist and future research is needed:

1. What is the role of ICT in coproduction? (See: [9], [12], [14], [15]).
2. How does ICT influence our understanding of public service delivery? (See: [3], [5], [12], [16]).

With a research gap identified, the next step in the thesis was to draft a research question which would help to address this gap. However, it was also important to narrow down the scope of the research. Thus, instead of looking at ICTs as a whole, the decision was made to focus on one type of upcoming technological innovation: open government data (OGD). The main reason for this decision is due to the fact that OGD is something that is believed to increase interactions between government and outside stakeholders, lead to new innovative public services, and change our understanding of public services [17]. Taking the aforementioned information into account, a research question was drafted, along with two sub research questions which should help to drive the research for this thesis.

- **RQ**: How can OGD contribute to the co-creation of new public services?
    - **SQ1**: How can an OGD driven co-created public service be designed?
    - **SQ2**: How does OGD influence our understanding of stakeholder roles in the public service delivery process?

Currently a Horizon2020 (H2020) project is being conducted within Estonia, and deals with similar issues raised by the research questions. This project is titled OpenGovIntelligence (OGI), and aims to show how OGD may be used to help drive co-creation and innovation within the public sector [18]. To demonstrate how this happens, six pilots are being conducted, one of which is taking place in Estonia. The Estonian Real Estate Pilot Program (EREPP), is a pilot program which is carried out in cooperation between The Estonian Ministry of Economic Affairs and Communications (MKM) and Tallinn University of Technology (TTÜ). The author of this thesis was directly involved with OGI project and in charge of the implementation of the pilot program in Estonia. This entailed planning and implementing the pilot throughout all stages of development, stakeholder meetings, organizing workshops, and writing the

initial code for the pilot project. The design and implementation of the pilot program focuses on the co-creation of a new public service and the exploitation of OGD[1]. This project gives the researcher direct access to a project which aims to co-create a new public service that is based on OGD; for this this reason, in combination with the similar aims and goals of the research, the project was selected as the case for this thesis.

This thesis is interdisciplinary in nature and aims to explore the relationship between public administration and management (PAM) theory and literature with some aspects from information technology (IT) studies. Thus, the aim of this thesis is to explore, discuss, and provide some initial answers to the research questions presented. The thesis begins this process by outlining the theoretical framework in Chapter 2. The theoretical framework starts with literature and theory from Public Administration and Management, and then continues on to a discussion of open government data (OGD). Chapter 2 also presents a conceptual model, which emerges from the interdisciplinary theoretical framework, which is then able to be used as a lens for examining the case. Chapter 3 will present the research methodology for the thesis and the research design of the case. Chapter 4 will present the case background, description, and a summary of the case. Chapter 5 will provide discussion on the case, the results of the research, and answer the research questions. Finally, Chapter 6 provides a conclusion and summary of the thesis research, and will also provide suggestions for future research.

1 OpenGovIntelligence has received funding under Horizon2020 grant agreement 693849, the author is thankful for the support and funding received from this project.

2 Coproduction, New Public Governance, and Open Data- Theoretical Overview

2.1 Coproduction and Co-creation

The term coproduction was first mentioned by Elinor Ostrom 1972[19]. What Ostrom noticed and articulated was that police agencies had different levels of effectiveness depending on the attitude of the citizens involved. Ostrom found that, in areas where citizens were more forthcoming and welcoming to law enforcement, there was a higher level of public service, or a higher production of public value, compared to areas where citizens were not as cooperative with the police [20], [21]. She thus concluded that the value of a public service was very much determined by not just the provider of the service, but by the interaction between the consumer of the service and the provider [20], [22], [23]. Since its conception, coproduction has been a constantly evolving subject within PAM literature, and the definition of coproduction is still debated up to this day.

When looking at the current state of understanding of coproduction, it is hard to keep the ideas straight as there are currently many different approaches. Table 1 represents some of the current interpretations and understanding currently visible in the academic literature.

Table 1. Types of Coproduction. Source: Author.

Authors	**Types of "coproduction"**
Whitaker 1980:	▪ Citizens requesting assistance from public agents ▪ Citizens providing assistance to public agents ▪ Citizens and agents interacting to adjust each other's service expectations and actions
Pollitt, Bouckaert, Loeffler 2006:	▪ Co-Planning ▪ Co-Design ▪ Co-Delivery ▪ Co-Evaluation
Bracci et. al 2016, Pestoff 2012:	▪ Individual acts of coproduction ▪ Collective acts of coproduction ▪ A mix of individual and collective acts of coproduction
Osborne, Strokosch 2013:	▪ Consumer coproduction ▪ Participative coproduction ▪ Enhanced coproduction
Voorberg et. al 2015:	▪ Citizen as a co-implementer ▪ citizen as a co-designer ▪ citizen as an initiator
Osborne, Radnor, Strokosch 2016:	▪ Coproduction ▪ Co-Design ▪ Co-Construction ▪ Co-Innovation

Paletti 2016:	▪ Coproduction through ICT:
	  ▪ Applications to report problems
	  ▪ Applications to crowdsource data
	  ▪ Applications to involve citizens in the public service delivery

There does seem to be many different ways in which coproduction may be understood. However, it is also important to outline how the term coproduction was initially understood and defined. This importance is assigned to the original definition of

coproduction as it forms the basic building block for future theories, and in order to assess the validity of advancements, the theory it is based off of must be understood.

When Ostrom talks about coproduction, she defines it as "the process through which inputs used to produce a good or service are contributed by individuals who are not "in" the same organization" [24]. Ostrom also notes that using the term "client" when defining a service is not necessarily the best term as client is "a passive term", and in her understanding of coproduction citizens can "play an active role in producing public goods and services of consequences to them" [24]. This mirrors a common complaint against New Public Management, as it is believed that by defining citizens as clients or customers they are viewed as unable to function without access to government services [24], [25]. In essence, what Ostrom is stating, is that coproduction must be a requirement for public services as some inputs are always present by individuals who are not service providers. This can be taken a bit further to state that public service cannot succeed without input from the service consumers [21], [23], [24]. Though it is obvious that participation of service consumers is paramount for the success of a public service, one should not count on service consumers to be automatically engaged and active citizens once a new service is provided [26]. A new public service needs to motivate active coproduction, however, if a public service requires higher levels of motivation for participation it is also likely that there will be an increased effort "required of service consumers to overcome hurdles to participation" [27].

One way to lower the barriers to coproduction is to involve citizens at every stage of the public service creation process, this is known as co-creation and the concept is also clearly present in Table 1. It is important to differentiate between the terms coproduction and co-creation as the concepts are quite similar, but the following statement should demonstrate the difference: all public services are coproduced, but not all public services are co-created.

Though there are many different ways to understand co-creation, the Government as a Platform (GaaP) typology[1] proposed by Linders in the paper "From eGovernment to weGovernment" seems to be a reasonable approach when looking at OGD and co-creation of new public services. Government as a platform emerges from the current situation where dissemination of government data and computer services is becoming

1 Government as a Platform is one of three classifications presented by [4], the other two are "Citizen Sourcing" and "Do it Yourself Government".

increasingly cheap, and that by making data and digital information available to citizens the state is able to "help citizens improve their day-to-day productivity, decision-making, and well-being. The government is not responsible for the resulting activity, but can leverage its platform and influence to foster greater public value" [4]. What we have here is a typology which posits low barriers to participation are, where citizens are able to take advantage of government data as needed, and this data may be used to create their own services and produce greater levels of public value. The government is, in effect, providing the basis for any actor to use this data to create a public service. Not only does this remove previous "top-down" approaches of public service creation and delivery, it also allows for higher levels of customization, citizen input, and citizen empowerment.

Some interesting conclusions may be drawn from the literature on coproduction and co-creation.

- Coproduction is an inherent part of any public service, it is necessary for effective public services, and that it is not an "add-on" to a public service, but a core attribute of a public service [9], [12], [22], [28].
- Public service delivery systems should focus on reducing the cost a consumer must 'pay' for playing an active role in the creation or involvement with a public service.
- New public services that are co-created have the potential to allow for higher levels of user empowerment, as it allows for consumers to "control their experience of a public service and contribute to their own desired outcomes" [12].
- The GaaP co-creation typology seems to be a strong contender for understanding how OGD can help drive the co-creation of new public services.

2.2 New Public Governance

One theory which heavily relies upon coproduction and co-creation theory, is the theory of New Public Governance. Stephen Osborne is credited with bringing the idea of NPG to light with his 2006 paper "The New Public Governance". When looking at the initial conception of NPG, there are two important things to note. The first: the NPG views the state as both pluralist and plural – there are multiple processes and actor's which influence the policy-making system and public service delivery [11]. The second is that

NPG does have strong theoretical background, and is defined as being "rooted firmly within organizational sociology and network theory" [11], recently service theory has also been utilized to explain public service delivery systems [6], [12]. NPG also believes that previous theories, such as NPM, fall short of providing an accurate presentation and understanding of governance [8]–[10], [28]. As NPM was focused on efficiency, borrowing heavily from the private sector ideas, costs are cut where possible and citizens are treated as clients or customers [3], [7], [28]. This emphasis on efficiency within NPM often causes public sector organizations (PSOs) to forget about the goals of the service which leads to a decrease in effectiveness of the PSO and the services they are providing [29].

NPG attempts to address these shortcomings of previous theories and places an emphasis on the improvement of public services, improving policy outcomes, and bases itself around collaboration and trust [10]. One of the most important takeaways from NPG is that any service which produces public value may be able to be considered a public service [1], [30]. This leads to a new radical idea: a public service does not need to be initiated on a government level, rather any stakeholder can be a producer or a consumer of a public service. Public services are defined by the interactions between the provider and the consumer. As such, it is possible to say that public services require "attention to the processes of their delivery and relationships with users, and not simply to service design" [28]. Since it is the interaction between parties which defines a service, it is important to pay attention to these interactions and the relationships.

NPG provides a new and radical way of looking at how public services are created, designed, and delivered. NPG believes that public services should be looked at as public service delivery systems, public services may be created by anyone, coproduction is an inseparable part of public services, that interactions and relationships may well be more important than the design of a service, the value of a public service comes from the interactions which take place, and that through collaboration and cooperation more effective public services can be produced and delivered [5], [11].

2.3 Open Government Data

In the previous two sections of the theoretical framework some important concepts emerge: public services are changing, collaboration and cooperation are incredibly important, outside stakeholders should have the ability to play a role in the co-creation

of new services, and that in the current digital age governments now have the ability to act as a 'platform', thus playing a role in the change of how public services are understood. These points emerge from PAM theory and literature, but many of these changes are enabled by new technological ideas and innovations. One such new technological idea that appears to be able to aid in the understanding of NPG, coproduction, and co-creation is the concept of OGD. OGD is said to increase government accountability, increase cooperation and communication with outside stakeholders, and is one of the core ways government is able to start acting as a 'platform' [31]. With the following in mind, it seems that OGD has the potential to play a direct role in enabling co-creation of new public services to take place. The benefits and motivations behind OGD directly relate to the ideas proposed in sections 2.1 and 2.2, and this allows for a bridge to be built between IT and PAM. With the bridge and relation between coproduction, co-creation, NPG, and OGD in place it is then possible to further explore the concept of OGD.

The purpose of this section will be to provide an overview of OGD theory: What is OGD? What are the proposed benefits of OGD? What are the drivers/barriers of OGD? Finally, how can OGD be used for the co-creation of new public services.

2.3.1 Defining OGD

When discussing OGD it is important to differentiate between the terms "open data" and "open government data"; the former refers to data which is open and provided by any source, whereas the latter directly refers to open data which is created and released by government agencies. There are currently many different definitions of OGD [31]–[33], but most definitions share some core components: data must be machine readable, it should be licensed in a way to allow easy sharing and reuse of data, and it should be usable and understandable by humans. With this in mind, for the purpose of this paper the following definition of OGD will be used: OGD is non-confidential data which is gathered, and subsequently released by a government organization in a machine readable format which is discoverable, usable, and freely available [31], [33]–[35].

2.3.2 Benefits of Open Government Data

There have been recent academic works that present some of benefits that may be provided by OGD such as, [31], [35]. In this section, the commonly cited benefits will be presented; Table 2 provides an overview of societal and governmental benefits that

may emerge from the use of OGD. The goal of 2.3.1 is not to provide a comprehensive overview of how these benefits may come to be realized, but it aims to provide an overview of what benefits may be likely to emerge for society and government agencies if OGD becomes available and it is used.

Table 2. Societal and Governmental Benefits Emerging From the Use of Open Government Data. Source: Author.

Benefit	Source
▪ Increase transparency	[31], [36], [37]
▪ Encourages social participation	[31]
▪ Drives innovation	[31], [35], [37]
▪ Drives creation of new services	[31], [35]–[37]
▪ Increased government accountability	[31], [36]
▪ Improved policy making process	[31], [36]
▪ Enhanced citizen services	[31], [36]
▪ Creation of new business models	[31], [35], [37]
▪ Optimized administrative processes	[31], [36]
▪ Improved data quality	[31], [36]
▪ Improved decision making	[31], [36]

2.3.3 Barriers Relating to the Use and Release of OGD

If OGD is released, and it is truly open (it meets the requirements set out in the definition provided in the introduction to this section), then it has the potential to create major benefits for society [31], [35]. However, as pointed out by [31], just making OGD available is not enough, as "OGD on its own has little intrinsic value; the value is created by its use". OGD usage generally refers to any interaction an actor (a user of OGD can come from any sector be it private or governmental [35]) has with the data, such as downloading, analyzing, or exploiting the data [38], [39]. As the use of OGD is paramount for benefits to be realized, there must be a clear understanding of the barriers related to the usage, and release of, OGD. The aim of this section will be to provide an overview of the current understanding of the barriers which prevent OGD from being used or realized.

OGD is becoming an increasingly popular topic to academics and practitioners, but it is also a fairly new one. That being said, there have been many attempts to provide a better understanding of OGD barriers [31], [35], [40]–[42]. When discussing the barriers for the use and release of OGD, it becomes increasingly clear that a majority of the barriers lie within government agencies and only a minority directly relate back to the user of OGD.

On the user level, commonly cited barriers are:

- Lack of technological understanding/ability; lack of domain-knowledge; language barriers to the data; lack of time to use data [35], [40], [41], [43];

On the government level, commonly cited barriers are:

- Missing political motivation; no understanding of the potential benefits of open data; missing technical infrastructure or technical know-how; poor data quality; confidentiality or personal data issues related to the release of data [31], [35], [40], [41], [43].

On the government level, a majority of the barriers are directly related to the release and publishing of open data whereas, on the user level, a majority of the barriers relate to their ability to use or understand OGD. However, in the case where these user barriers do not apply, namely when the user of the data has a strong understanding of data analytics and a personal interest in open data, they may often struggle to use the OGD which is provided. In this situation, the most relevant barriers are related to the poor quality (encoding issues, missing values, lack of metadata, etc) of the OGD, lack of interesting information, outdated data, and lack of an application programming interface (API) functionality [43].

2.4 OGD and Co-Created Public Services

In this section, the main goal is to discuss how OGD can be used to create new public services. In order to achieve this goal different forms of OGD based services will be presented, as well as how these public services create public value. When looking at different services that create public value by exploiting OGD, the three more common ways in which it is done: creation of new web/mobile applications; creation of OGD portals; and through the use of data analytics (this includes creation of machine learning (ML) models for outlier analysis or predictive modelling, as well as creation of dashboards for easy usability and understandability of data). For each of these three ICT mediums the following will be presented: how OGD is used in the co-creation of the public service, how users can interact with the public service, and how the public service creates public value.

2.4.1 Web/Mobile Applications

Example Service: Hackcessible[1]

Hackcessible is an application which was developed to "enable safe, accessible trip planning on pedestrian ways for people with limited mobility" [44]. It is able to do this by utilizing different OGD sets which relate to the condition of sidewalks, availability of ramps, location of street crossings, and elevation changes [44]. The service provides a map based visualization which color codes sidewalks in the city of Seattle based on how accessible they would be for those with limited mobility. A user simply needs to navigate to the service, type in a current address, type in a destination, and is then able to select the best route based on different factors: maximum incline/decline, avoiding construction sites, or requiring the presence of curb ramps (it is also possible to choose default settings based on different types of mobility handicaps) [44]. This application is completely open source and utilizes only open data sets, it provides a service which was not previously provided by the government, and it provides great value for those who suffer from mobility impairments.

2.4.2 Open Government Data Portals

Example Service: United Kingdom (UK) OGD Portal[2]

Open Government Data portals are where governments, normally, host all of their OGD sets in one place. This allows for easy access to all information the government provides as OGD. In the case of the UK OGD Portal, data.gov.uk, users are able to search for different open data sets by keyword or theme, access information via an API, hold discussions on forums, request information which is not currently available, and access over 400 applications which have been developed using OGD sets from data.gov.uk [45]. OGD portals play a critical role in the co-creation of new OGD based public services, as they allow stakeholders to easily access and request OGD sets. So, OGD portals create public value in multiple ways: they provide access to OGD sets, they allow for easy construction of applications based on OGD, and they provide a way for application developers to easily advertise newly created services. It is also important to note that it is not just developers who benefit from OGD portals, anyone is able to go

1 To see this service, please visit: www.accessmap.io/ (accessed 04 May, 2017)

2 To see this service, please visit: data.gov.uk/ (accessed 04 May, 2017)

and look at different datasets; often data is able to be visualized with just the click of a few buttons – this, then, allows for those without the technological knowledge to create applications or conduct data analysis to see and understand data in an easy manner.

2.4.3 Data Analytics

Example Service: Chicago Food Inspections[1]

The Chicago Food Inspections application uses OGD sets to create a model that allows for food inspectors within the city to target higher risk food service providers in a more efficient manner. This model allowed for critical food violations to be found, "on average, 7 and a half days earlier" [46]. The development was done in a completely open manner, and the datasets which were used, as well as the different models, are available on GitHub. So, open data was used in the training of new ML models, these models were then put into production and allowed for serious/critical food violations to be caught in an earlier manner, and the code is open so that it can be copied or implemented by others. Though users do not directly interact with this, there is a large amount of public value created as these models do have the potential to stop many individuals from getting seriously sick due to unsafe food handling by food establishments.

2.5 Co-Creation in an Open Government Data Driven Public Service

In section 2.1 the ideas of coproduction and co-creation were presented, in section 2.2 NPG's approach towards public services were laid out, in section 2.3 the concept of OGD was discussed, finally, in section 2.4, examples of OGD driven public services were provided. Following the discussion and presentation of these different concepts, the next step is to demonstrate how these concepts work together and put them together as a framework. One such framework which is built upon the aforementioned ideas was first put forth by [17] and deals specifically with OGD driven public service co-creation, this framework is demonstrated in Figure 1.

1 To see this service, please visit: github.com/Chicago/food-inspections-evaluation (accessed 04 May, 2017)

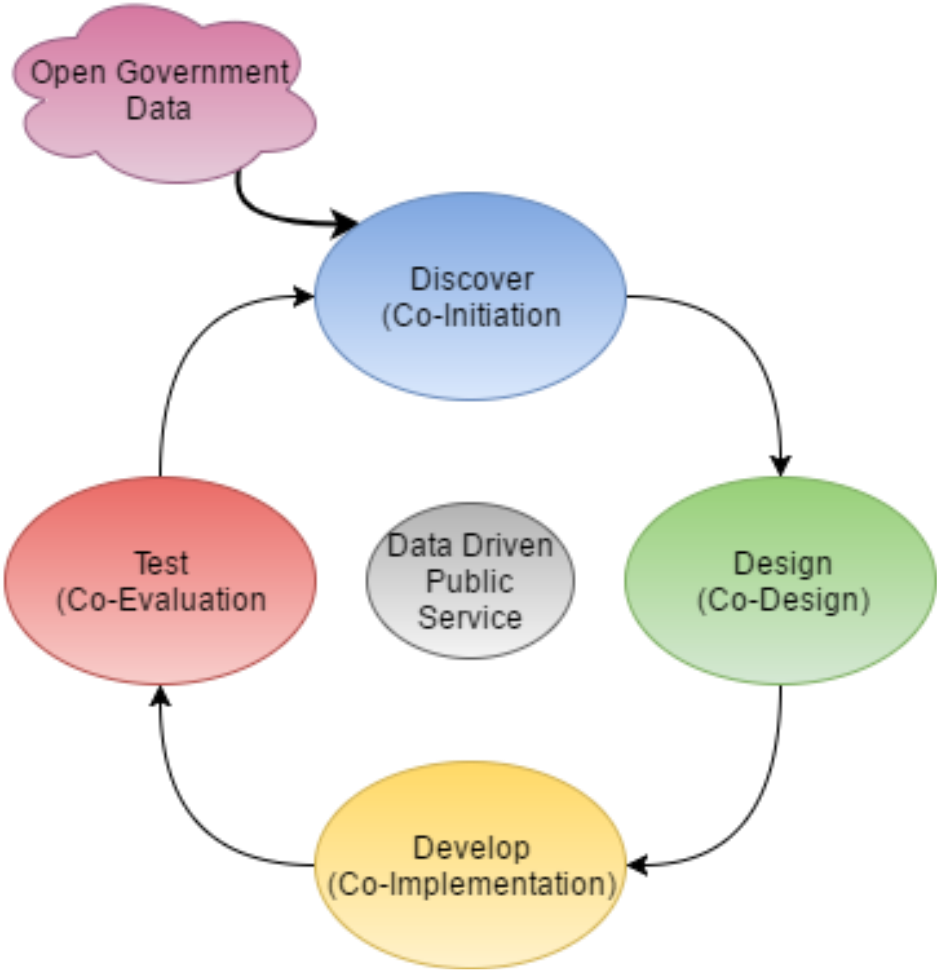

Figure 1. OpenGovIntelligence Co-Creation Framework. Source: [17], [47].

The framework draws on the ideas of agile development, lean methodology, co-creation, and aims to provide a way to support open government "data-driven service creation" [17]. In the rest of this section each aspect of the framework will be discussed and then an example of a co-created OGD driven public service will be presented.

Agile development is based around the idea of 'sprints'. Sprints are, in essence, short time periods were development is done following a four stage cycle: plan, build, test, release [17], [48]–[50]. Agile development allows for services to be developed more efficiently and makes them responsive to user suggestions and changes in initial development plans. This flexibility that comes with agile development is necessary for any service which is being co-created. The reason for this is that, as users are constantly providing input to the service, it must be able to adapt quickly.

In Figure 1 the different stages and ways in which stakeholders may contribute to the co-creation of the service is also presented. At the beginning stages, stakeholders are able to act as a co-initiator and play a role in the initiation of a new public service. In the

second stage they participate in co-design, this is where the service is being designed in cooperation between the service user and service producer (this may take place, for example, as user workshops where initial functionalities are described and defined). Thirdly, there is a co-implementation stage, this is where service producers and consumers work on implementing the service together, in practice this normally appears on GitHub with open source code where anyone has the ability to fork the code and make a contribution. The final stage is co-evaluation, at this stage all stakeholders are able to play a role in the evaluation of the service and provide feedback so that future iterations are improved.

The final aspect of the model is the role of OGD. OGD is the catalyst which allows for this process to take place. In essence, government agencies release their data and make it open. Once this has happened, anyone is able to use and exploit this data to drive public service co-creation. As OGD is freely accessible and exploitable by any stakeholder, having OGD available allows for any actor to start the process of co-creating a new public service.

Thus, what emerges is an initial conceptualization of the process in which OGD enables any actor to exploit OGD and thus become an active participant in the co-creation of a new public service. This process is able to start when access to OGD becomes available and is exploited. Drawing on the concepts put forth in sections 2.1 through 2.4, some key ideas also emerge from the framework:

1. Government and citizens should be partners throughout the creation and implementation of a new service [17].
2. There should be an initial release of the public service at an early stage, which allows for the development cycle to start as quickly as possible [17].
3. The new public service should be able to quickly respond to user feedback [17].
4. User input should be sought at all stages of public service development [17].

If this framework is followed, it allows for "citizens and service users to be put in the driver's seat and allow for a genuinely user-centric, adaptive and collaborative innovation process" [17].

2.6 Summary and Propositions

The theoretical framework constructs the following picture: anyone is able to create a new public service, public service provision should be viewed as a system or network,

the value of new public services emerges from the interactions which take place between service provider and service consumer, and the availability and exploitation of OGD has the potential to enable and drive public service co-creation. However, it is also clear that in order for OGD driven public service co-creation to take place, it must be enabled. In order to understand this process, and how it emerges, four propositions have been put forth. These propositions will also aid in the answering of the research questions presented in Chapter 1.

An investigation of these propositions will allow further insight to be gained about the effectiveness and accuracy of the model proposed in the previous section. Currently no work has been done to evaluate this framework, and thus empirical data is needed. This thesis aims to fill this gap by providing a detailed description and analysis through the use of an empirical case study.

Proposition 1 (P1): In order for OGD to be able to contribute effectively to the co-creation of new public services, a change in the current understanding of stakeholders in the public service creation process must take place.

Proposition 2 (P2): If OGD is released by government agencies, and this data is used to create a new public service, then the government is a participant in the co-creation of new public services.

Proposition 3 (P3): OGD may be used as a base, or platform, from which new and innovative public services may be co-created between government agencies, who maintain and release the data, and with outside stakeholders.

Proposition 4 (P4): OGD has the potential to enable, encourage, and drive public service co-creation.

3 Research Methodology

At the beginning stages of the development of the thesis a thorough overview of the literature was conducted. Out of this literature the following became clear: public service delivery is changing, coproduction is an inherent and important part of public services, ICTs and innovative technologies are changing how public services are created/delivered, and OGD has the potential to create major benefits for society (such as the creation of new services). A research gap also began to manifest itself, there was an obvious gap in the academic literature exploring how OGD may be used to create new public services and how OGD may contribute to the co-creation of new public services. The research question that drives this thesis was born out of the discovered research gap. With the research question in place, it was important that the research methodology chosen to produce this thesis made sense and had the potential to allow for this thesis to contribute to the current relevant scholarly work. As the research is qualitative in nature, two main research methodologies were initially compared, case study and design science (DS).

A **case study**, in general, may be defined as "an empirical method aimed at investigating contemporary phenomena in their context" [51]. According to Yin 2013, a case study is appropriate when "a 'how' or 'why' question is being asked about a contemporary set of events, over which the investigator has little or no control" [52]. Furthermore, a case study may often be used to help "contribute to our knowledge of individual, group, organizational, social, political, and related phenomena" [52]. If the case study meets the aforementioned criteria, the next step would be to decide how to design the case study research. There are multiple approaches to this. Case study research can be either holistic or embedded, it may also be designed to include a single case or multiple cases [51], [52]. A case study may also be further defined depending on what the goal of the case study research is. Yin 2013 outlines five main rationales for single-case-study research: critical cases, extreme or unique cases, representative or typical cases, revelatory cases, and longitudinal cases [52].

A case study may be appropriate when the research question being asked is a 'how' or 'why' question; the research question presented in this thesis is a 'how' research question and thus the first criteria is met. The current research is dealing with a contemporary set of events so the second criteria is met. The third criteria posits that the

researcher has little or no control over the events, this is also true in this case and therefore the third criteria is met. The research question is attempting to understand a new and contemporary process, which deals with organizational, political, and social phenomena so it does appear that a case study will be able to help contribute.

Design science is one methodology which is increasingly gaining support within information systems research [53]. The core principle behind design science is that "knowledge and understanding of a design problem and its solution are acquired in the building and applications of an artifact" [53]. Design science research methodology is often appropriate when there is some inefficiency and the research scientist wants to design some artifact to address this inefficiency [54]. Recent papers have tried to draft a conceptual model for DS research, the nominal process sequence is as follows: Problem identification and motivation; objectives of a solution; design and development; demonstration; evaluation; communication [54]. Though there have been attempts to create a conceptual or mental model for DS research, one of the main weaknesses of DS research is that there is still not a *widely agreed upon* model for carrying out DS research and evaluating the outputs [53], [55]. It is true that one part of the research being undertaken for this thesis is the creation of a new artifact, however, when looking back at the research question, the artifact itself is *not* the main-focus of the question. It is also important to note that at the time of writing this thesis, the artifact itself is not completed and is therefore out of scope (it is not possible to evaluate its success as it yet to have been used or evaluated). Given all of this information, it appears that a case study may be a suitable research methodology for the current thesis.

3.1 Case Study Design

The unit of analysis is a critical part of any case study design, this is what drives and guides the case study research and it should emerge naturally out of the main research question [51], [52]. Thus, the unit of analysis for this case study is the *process* of how an open data driven public service is created. The focus of this case study is not the artifact that is being created; rather, it is the *process* of how the artifact was created. It is also important to point out that this process is not yet complete, so the case will only focus on the process from project inception to the beginning of the artifact creation. Within this case there is only one unit of analysis, and it is for this reason that the case is holistic rather than embedded.

In certain cases, an exploratory case study may be conducted. Ultimately, the end goal of an exploratory study is to begin to construct new theory, propositions, or gain an understanding of some current happenings [52]. In the case of this thesis, an exploratory case study will allow for a further exploration into how OGD can be used to contribute to the co-creation of new public services.

The case that has been selected for research is the Estonian Real Estate Pilot Program, which is being carried out as part of the H2020 OpenGovIntelligence project. The reason for selecting this is due to the researcher's direct access to the project, the relevance to the research question, and due to the diverse nature of stakeholders involved in the project. The aim of this project is to use OGD to create a new service within Estonia. The project focuses on open source development, prioritizing the use of OGD, and citizen co-creation. As the author of the thesis is heavily involved with this project, it may represent a revelatory case. A revelatory case is one rationale that may be used for choosing a single-case over a multiple case study approach. Revelatory cases, in general, occur when the investigator or researcher has the ability to play an active role within an ongoing project [52].

3.2 Data Sources and Collection

When conducting a case study, one way to improve the validity of the study is to utilize multiple sources of evidence [51], [52]. Multiple sources of evidence will allow the investigator to triangulate the data and lead to the emergence of stronger conclusions from the case study research. For this thesis, four different sources of evidence have been used: semi-structured interviews, workshops, survey, and documents.

Interviews were conducted with governmental and nongovernmental actors in an attempt to gain a better understanding of the context surrounding the use of OGD to co-create new public services within Estonia. The conducted interviews were semi-structured interviews. Semi-structured interviews work well for exploratory case studies as they allow for improvisation and exploration of different ideas or issues that emerge naturally from the conversation [51], [52]. For the semi-structured interviews, a list of five basic questions was prepared (Appendix 4), and these were then used as the starting point for the interviews; however, the interviews were not limited to only these questions as may be the case in more formally structured interviews [51]. The interviewees that were selected represented both Estonian public administrators who

deal with service development and OGD, as well as Estonian civil society members who had an interest in the topic. A total of six interviews were carried out, five interviews were done face to face and recorded, whereas one was done through email. A list of interviewees may be found in Appendix 1. After the interviews the responses were transcribed and then coded. Ultimately, semi-structured interviews are one of the best ways to get an understanding of how stakeholders interpret or view the issue; they also often lead to new or unknown ideas and facts to be brought forth due to the informal nature of the interview [51], [52].

Two different workshops will provide evidence for the case. One was organized by the author; the other was attended by the author. The first workshop was structured in a way to follow Nominal Group Technique [56]. The workshop was attended by seven different public sector agencies and two private sector Estonian companies. The goal of this workshop was to involve many different stakeholders and get an understanding of how the new Estonian service should be built/designed. This understanding comes in the form of user stories and personas which were generated by workshop participants. The second workshop was an unofficial Estonian working group that deals with the use of data analytics for the improvement of public services. This workshop was attended by both public sector officials and private sector individuals. This workshop represented another way for feedback to be received from stakeholders representing different sectors.

As part of a larger study, a survey which aimed to gain insight into how OGD could be used in the co-creation on new public services was conducted. This survey received 63 responses, however the responses received represented six different countries so only the responses from the Estonian stakeholders may be included in this thesis. Nine responses were received from Estonian stakeholders, three from the non-governmental sector and six from the public administration sector, survey respondents are shown in Appendix 3. For a further write up and discussion of this survey, please see [35].

The documents that have been analyzed are Estonian legal text relating to OGD, the "Public Information Act", and an Estonian government Green Paper, "Green Paper on The Organization of Public Services". These documents provide background on the legal and social situation of OGD and public service design in Estonia.

In this section the data sources have been presented as a hierarchy, in order of importance. The semi-structured interviews provide direct insight into how stakeholders who are involved with public service creation and OGD currently perceive the situation.

The information that emerged from these semi-structured interviews provided new information, which was not currently available elsewhere. The workshops also provided access to relevant stakeholders and practitioners, but the information gathered here largely related to how OGD could be used in the co-creation of a new public service rather than their personal experience with the topic. The survey provided insight into the specific topic of OGD driven public service co-creation. Finally, the documents which were analyzed provided an overview of the current legal and organizational understanding of the theoretical concepts at play in this thesis.

The document analysis and survey provided an initial overview and understanding of the situation in Estonia in regards to OGD and public service co-creation. The workshops allowed for further exploration into the drivers and barriers related to OGD public service co-creation, as well as how OGD driven co-created public services should be designed. Finally, the interviews allowed for further exploration on topics which emerged, as well as those that did not, from the documents, survey, and workshops. Ultimately, the use of different sources of evidence allows for results to be compared and contrasted, thus strengthening the initial results.

Though the reasons and benefits of each source of evidence have been presented here, the limitations and risks have not been. In the next section of this methodology chapter, the main limitations and risks will be presented and discussed.

3.3 Risks and Limitations

When looking at the potential limitations of case study research, two of the most commonly mentioned are the lack of rigor of case studies and the case study's lack of generalizability [52]. In order to address the first limitation case study research must ensure that multiple sources of evidence are utilized, the research design is valid, and that the case study is carried out in a way consistent with widely agreed upon case study methodology. In regards to the second limitation, lack of generalizability, there is some truth to this. However, as pointed out by [52], case study research may well be generalizable to "theoretical propositions and not to populations or universes"; that is to say, single case study research may be used to make theoretical generalizations, but single case studies are not able to make generalizations about entire populations or universal laws. With these limitations in mind, the case study research has still been

deemed the appropriate strategy and steps have been taken to counteract these limitations and ensure the rigor and validity of the presented study.

There are, of course, also risks involved with case study research which have the potential to ruin the validity of the study; one of the main risks is bias. This is especially true in the case of the current thesis, as the author is an active participant throughout the entire project. However, it is also possible to mitigate this risk. One of the best ways to do this is to present all evidence that has been collected in a neutral manner, thus allowing the reader to agree or disagree with the authors' conclusions [52]. Though the direct involvement of the researcher in a case study may be viewed as a needlessly risky way to bring bias into the research, it is also, often, a critical form of evidence providing new and interesting insight into the issue which is being studied [52]. Being directly involved with the research throughout all stages raises the risk of the research being done in a subjective manner, though steps have been taken to mitigate this risk, it is possible that subjective ideas emerge in the thesis. In order to mitigate the risk of bias and subjectivity the research design must be strong and valid and the evidence must be presented in a fair and neutral way.

It is also important to discuss the weaknesses or limitations of the different sources of evidence that have been selected. The four most commonly cited weaknesses of the chosen evidence sources are bias, reflexivity, selectivity, and availability [52]. Though it is true that these are valid limitations of the selected sources of evidence, utilizing multiple sources of evidence, having a strong theoretical foundation, having a strong and valid research design, and presentation of evidence in a neutral and fair manner all help to mitigate these limitations and weaknesses [51], [52].

4 The Case

4.1 Case Background

The Estonian Real Estate Pilot Program aims to fight information asymmetry in the Estonian real estate market by exploiting OGD to create a new public service; this development is done in cooperation with private sector stakeholders, universities, and government agencies. This pilot project is the case that is the focus of this thesis and is one of six pilots that is being conducted within the H2020 OpenGovIntelligence project. Though the Estonian Pilot is the focus of the case, it is also important to understand the context behind the pilot as it does play a strong role in influencing the overall pilot direction. As such, the aim of this background section will be to provide a brief introductory overview of the OGI project, the pilot framework, and the current OGD and co-creation situation in Estonia. Once the context is in place it is possible to discuss the service that is being developed, the goal of the service, and the aspects of the service that are in-scope for this thesis. This section should provide enough of an informational foundation for a more in-depth discussion and analysis of the case to be conducted.

The OGI project is a 2.8million EUR project which aims to demonstrate how open data and statistical analysis can be used to create new services and drive innovation within the public sector [57]. To this end, the OGI project selected six different countries for pilot projects to be carried out. Each pilot partner must use OGD to co-create a new service in cooperation with stakeholders from different spheres, such as citizens, public sector, or private sector. These pilot projects should addresses a relevant and current societal issue and, if done successfully, the pilot should demonstrate how OGD is able to drive innovation, improve service quality, and overcome barriers relates to the use and exploitation of OGD in the public service creation context [47]. In essence, OGI posits that the definition of a public service is changing, that co-creation and cooperation with stakeholders is important part of this change, and that OGD is likely to play an important role in changing our understanding of what a public service is and how it is created.

While the OGI project was being drafted, and partners were being chosen, every potential pilot partner was asked to describe a current problem that may be addressed with OGD, the OGD datasets that were available, and how these datasets could be used

to address the aforementioned problem. Estonia was able to provide a problem, potential datasets, and articulate a reasonable solution for the problem.

- Problem: There is information asymmetry in the Estonian real estate market. Information is out there, but there is no one place where all information is available. Currently, the information is only available by navigating to many different ministry websites and querying the X-Road to gather information. In addition to this, some OGD sets are available, but they were generally hard to find and unusable in their current form [18].
- Datasets: When the Estonian pilot program was initially proposed, approximately 11 different sources of data were mentioned as being potentially relevant for addressing the problem. These datasets included information coming from the building register, environment information, crime statistics, and real estate valuation data [18].
- Solution: The initial solution proposed for the Estonian pilot was a search engine based application. The general idea was that any individual could input an address and all datasets that contained information on this address would be returned; this application would also have the ability to display relevant geospatial data on a map.

In addition to this information, Estonia also came with the image of being a world leader and expert when it comes to ICT innovation, e-services, and e-society [58]. Thus, it follows that Estonia may well be an ideal location for using OGD, in combination with newly created ICT tools, to create a new public service which has a high potential to create new benefit for society.

Once Estonia had been selected as a pilot country, the process to understand the current situation of OGD and co-creation began. The initial research included an overview of the potentially relevant legal texts, an overview of relevant government ministry documents, a survey, and further on in the project unstructured interviews were undertaken to get further knowledge of stakeholder perception of OGD and co-creation of publics services. The results of this research provide the necessary contextual foundation for understanding a more in depth discussion of the case.

Inside Estonia OGD is regulated by the Public Information Act (PIA). In the act it is stated that all data that may be used for public purposes, that is to say it is not restricted by law, shall be opened to the public [59]. The PIA continues on to say that data should be released in machine readable format and come without any restriction on reuse of the

data; however, it also states that this is only required if it would involve "disproportionately great effort" [59]. Agencies should be following an "open by default" policy, but often what ends up happening is agencies hire a person to deal with all incoming data requests (Interview D). Estonia has made steps in the right direction for OGD, but as it is not a concrete requirement for all data to be made available in an open and machine-readable format, the availability of OGD in Estonia is lacking (Interviews A&C, Survey EENG1, EENG3, EEPA5).

As many agencies are not taking or playing a leading role in the release of OGD, true OGD is hard to come by. In Estonia, the main OGD portal is opendata.riik.ee/en. This portal provides access to 58 datasets, however many of the listed datasets simply point to HTML pages and thus do not meet the criteria to be considered OGD. Due to the lack of quality of the data on the OGD portal, it has largely gone unused. Most of the usable OGD sets are present on government agency web pages or repositories and are available for download as either CSV or XML files; there is currently no OGD repository in Estonia with API access (Interview F).

The second important part of the background information relates to the current situation and understanding of co-creation of public services within Estonia. The idea of co-creation of public services within Estonia is one that does not have much government support (Interview A,B,C,D,E,F, Survey EENG1, EENG2, EENG3, EEPA3, EEPA5, EEPA6). Some of the reasons for this include not enough funding, not enough citizen demand, low levels of collaboration between citizens and government, and lack of understanding of the concept (Interviews A,B,C,D,E,F). With this in mind, steps are being taken to try to move towards creation of new public services that have been co-created with citizens (Interview E&F). Many agencies are going out to end users to ask for their input on what services they need and then trying to involve them in the design and creation process (Interview D). The Estonian Ministry of Economic Affairs and Communications (MKM), has also attempted to increase government-citizen interaction by requiring agencies to conduct an analysis of end-user needs before receiving European Structural Funds for new public services (Interview E&F).

Though Estonian public agencies are beginning to consult potential service end users at the beginning stages of development, this is often where the cooperation ends (Interview B). When looking at the usage of OGD for the co-creation of new public services inside Estonia there is not, yet, an example as it is not currently occurring. Some of the main reasons for this are the infancy stage of the Estonian OGD

infrastructure and agencies not being willing to participate in co-creation of services with citizens and other stakeholders (Interviews B&F). Other examples of barriers for OGD co-creation were brought up in the conducted survey and included: not requiring compliance with Estonian OGD recommendations, lack of technical understanding, lack of resources, unwilling leadership, X-Road, and small demand for OGD [35], [18].

4.2 Case Description

The Estonian Real Estate Pilot Program (EREPP) aims to address the current information asymmetry in the Estonian real estate market by exploiting OGD to create a new service that provides easy access to all relevant real estate information for any given address [18]. As there is currently a lack of access to quality OGD in Estonia, one of the aims of the pilot program is also to open up datasets that are currently not available. The pilot also aims to include users in the design, development, implementation, and evaluation as often as possible. Ultimately, the goal of EREPP is the co-creation of a new OGD driven public service.

The Estonian Real Estate Pilot Program will follow the OGI model for pilot program development. That is to say, there will be three releases of the pilot and each iteration will follow the co-creation cycle shown in Figure 1. The timeframe of events, which are being presented within this case study, is from June 2016 until April 2017. As the pilot project is still in development, the stages, which are in scope for this thesis, are the co-initiation phase, the co-design stage, and the beginning of the co-implementation stage. In addition, as EREPP is still in the first stage of development, the application is currently only focusing on providing data for Tallinn; future releases of the service aim to include more of Estonia. In the following subsections of this chapter, each stage will be presented and discussed in depth. Evidence that was gathered at each stage will be used to provide greater insight into the events that happened or the actions that were taken.

4.2.1 Outlining Available Open Government Data Within Estonia

As one of the main goals of EREPP is to demonstrate how OGD can be used to create new public services, the first task that had to be completed was to get an overview of the currently available OGD sets. In order to identify these datasets, three different approaches were used. The first and initial starting point was a simple google search

using the Estonian term for open data, “Avaandmed”. The second approach was to use Google Dorks[1] to try to target sites within Estonia that may provide access to OGD. The third approach was to make requests to government agencies for datasets which could be useful for pilot development. In the following paragraphs each approach will be discussed and what was returned from each approach will be provided; it is also important to note that OGD sets which were not related to real estate will not be presented as they are out of scope for the pilot and were thus automatically excluded.

A simple search on Estonian google using the Estonian term “Avaandmed” mostly returned links to news articles discussing the term, Estonian government documents that discussed OGD, and links to government agencies were also present. When navigating to the government ministry webpages, often the pages redirected to either www.rik.ee/et/avaandmed or opendata.riik.ee/; the first link is a simple repository hosted by the Centre of Registers and Information Systems (RIK) that provided links to different OGD sets and the latter is the official OGD portal of Estonia. However, some government organizations such as Tallinn City, Tartu City, Maa-amet, and Riigi Kinnisvara provided OGD sets directly on their own webpages. Though useful datasets were found in this manner, other search methods were tried to make the process easier or more efficient.

As a follow up to the initial search for OGD within Estonia new search techniques (Google Dorks) were utilized. Google Dorking originates from the “google hacking” community and they are used to find vulnerable systems or passwords/login information in plaintext [60]. However, Google Dorks are quite useful for finding information, as one is able to target your search to specific parts of a webpage. Some of the Google Dorks that were used for obtaining OGD sets are outlined below:

- opendata and estonia site:github.com (This searches for the term open data and Estonia on the site github.com)
- opendata and estonia site:.ee (This searches for the term open data and Estonia on all sites with a .ee domain)
- avaandmed and csv OR xml OR JSON site:.ee (This searches for the term avaandmed and either csv, xml, or JSON on .ee sites)

1 For a discussion of Google Dorks, please see: whatis.techtarget.com/definition/Google-dork-query (accessed 03 May, 2017).

- inurl:opendata site:.ee (This searches for any URLs which have open data and .ee as part of the URL).
- inurl:avaandmed site:.ee (This searches for any URLs which have avaandmed and .ee as part of the URL)

As many individuals who work with OGD are active users of GitHub, the first Google Dork aimed to find and identify repositories that existed, and dealt primarily with, Estonian Open Government Datasets. This did turn out to be the case, for example see: github.com/jaakk/EstoniaOpenData. The other Google Dorks also successfully helped identify new OGD sets such as opendata.mkm.ee/EHR/. After this second attempt at gathering data, some datasets still appeared to be missing, mostly those regarding road safety and criminal activity.

The last approach used to gather datasets was to contact directly the data owners to attempt to gather the datasets. One dataset that had not yet been discovered, but was mentioned as an important dataset to have when the pilot was initially drafted, was OGD relating to criminal activity within Tallinn. A formal request for data was sent to the Ministry of Justice Statistics division, they are in charge of criminal statistics, the data asked for needed to include a breakdown of crime by type, the linnaosa (city district) where it occurred, the count of each type of crime, and this data should be provided for at least 5 years. The initial request for data was sent on July 18, 2016 and the data was received on September 6, 2016. The datasets that were received were individual CSV files for every linnaosa in Tallinn outlining the type of crime, the count, and for the years 2006 - 2015. The second dataset that was obtained with this method was data on all car accidents that occurred in Tallinn, the time they occurred, what the situation was, and the total damage. However, this dataset did not come with an open license and it is therefore unable to be considered OGD.

At this point all known datasets that pertained to real estate had been mapped out and put into a list, a copy of this list may be found in Appendix 1. At this stage, the task that concerned the gathering of OGD for Tallinn had been completed. However, if the EREPP was to be developed and launched in a quick and agile manner, the datasets that were used for the initial MVP release had to be selected and narrowed down.

4.2.2 Estonian Real Estate Pilot Program Workshop

The Estonian Real Estate Pilot Program Workshop was conducted on 16 September 2016[1]. The workshop had nine attendees (excluding the organization team) who represented seven different government agencies and two private sector companies; a full list of attendees can be found in Appendix 2. The workshop followed the structure as described by Nominal Group Technique: introduction, silent ideation, group discussion, and clustering of ideas [56]. The four main goals of the workshop will be presented below, as well as how the workshop aimed to achieve these goals.

1. To get feedback from multiple stakeholders.
    a. By bringing in stakeholders from multiple government agencies and from the private sector it was possible to receive feedback on the idea behind the new service from multiple perspectives. The workshop also went beyond just receiving feedback and allowed every participants ideas and thoughts to be taken into account, thus allowing them to drive the direction of the service.
2. To improve the initial MVP of the service
    a. The MVP of the service should focus on doing one thing really well, in the EREPP case an understanding was needed in regards to what the most important service the pilot could provide, as well as what datasets would be needed to do this. The participants in the workshop would play a direct and important role in outlining the initial functionality and datasets that should be offered.
3. To involve stakeholders in the design of the service
    a. Here the participants were able to help with the design by developing a set of initial functionalities for the service, creating simple UI wireframes, and creasing user stories and personas.
4. To raise awareness of EREPP.
    a. In Estonia, there is currently not much perceived demand for OGD (Interviews A,E,C,F,D, Survey EEPA1, EEPA2, EEPA3, EEPA6). One of the main goals of EREPP is to demonstrate that if OGD is available

[1] For a blog post writeup of the workshop please see: medium.com/opengovintelligence/opengovintelligence-pilot-showcase-the-estonian-pilot-511fb4647e8 (accessed 03 May, 2017).

then new innovative applications and services can be built. If this demonstration is done successfully, it is likely government agencies and OGD developers may become more involved in providing or using OGD (Interviews A,C,E,F).

These four goals were divided between two different session within the EREPP workshop. The first session was titled "Developing the Estonian Real Estate Pilot Program" and was focused on the following questions:

- What problems do you see with this pilot program and its goals?
  - What are the main areas in the real estate sector where you see issues?
- What are the potential solutions for these problems?
  - How could this pilot program better address these issues?
- In what ways would this new public service be beneficial for Estonia?

All participants spent some time writing down answers to these questions silently, then each participant shared their answers and the answers were discussed by all parties. A list of the problems and potential solutions is provided in Table 3 and they are ordered by number of occurrences/mentions in the workshop starting from the most common to the least common.

Table 3. Problems and Solutions Generated at Estonian Real Estate Pilot Program Workshop. Source: Author.

Problems	Solutions
1. Data Quality	Open API solutions, Automatic Dataset updates
2. Confidentiality Issues	Anonymizing the data
3. Competition from Existing Real Estate Portals	Involve users in design process, talk with current users of existing portals
4. Data is not Open	Open API solutions, common license template
5. Data Integrity	Update national registries, involve government agencies
6. Needed data not collected	Invite users to help collect and provide needed data

When examining the problems, it seems that a majority require government intervention or policy changes. The problems that may be addressed from the end of the development team are problems three and six, there is also a possibility to address problem two by potentially utilizing new technical solutions that create new methods to analyze confidential data.

After the problems and solutions for the problems had been presented and discussed the potential benefits of this new service were presented by all participants. The top five benefits were listed and voted on by all participants and are listed below.

1. Increased transparency in the Estonian real estate sector
2. Fairer pricing
3. Happier citizens
4. One stop shop for real estate data
5. Increased available and usage of real estate information

The second session was titled "Constructing the Functionality" and dealt primarily with constructing user stories and personas. In any agile development project, personas and user stories play a critical role as they allow the development to reflect better the actual needs of the users [61]. These user stories and personas that were developed help to define the pilot program's initial user target group. Though many personas were presented, there were two core target groups which were consistent throughout the participants' work; foreign students and foreign employees who are moving to Tallinn. For this reason, the following personas have been kept in mind for the initial MVP development of EREPP:

- "As a student moving to Tallinn, I need information on location of universities, public transportation, safety of the neighborhood, and the potential cost of rent" [62].
- "As a foreign IT specialist coming to Tallinn, I need a safe, environmentally friendly place to live; I need information on safety and public transportation so that I can live in a safe place and move easily around the city" [62].

After the personas had been presented, and the two most likely initial target personas were selected as the target group, the participants were asked to create ten to fifteen user stories for each persona. The participants were asked to stick with agile methodology and develop user stories in the following style: As a <type of user>, I want <some goal>

so that <some reason> [62]. Once these user stories had been drafted and discussed five were selected for the initial release of the pilot program.

1. As a user, I want to be able to search by address so that I can see the address location on a map.
2. As a user, I want to be able to see how safe a new home is so that I can feel safe when moving to a new location.
3. As a user, I want to be able to measure the distance between my house and different points of interest so that I can travel more efficiently.
4. As a user, I want to be able to find addresses where I could live based off of my own criteria (for example, find housing which is super safe without paying attention to price).
5. As a user, I want to be able to find out how often public transport comes by a given address so that I can gain a better understanding of how much time will be spent commuting via public transport.

Taking into account the personas and the user stories, the initial datasets that should be included in the pilot program were discussed, voted on, and selected. These five datasets were public transport data, safety data, price data, point of interest data (schools or doctor's offices), and property information (such as age of the building, amenities within the building, or the accessibility of the building).

This workshop represents a major milestone in the Estonian Real Estate Project Pilot and the OGD movement in Estonia. The workshop brought stakeholders from private and public sector together to discuss the benefits of open data, and there was genuine interest from both sides. On the private sector side, the company Teleport willingly granted the development team access to their datasets that provided information into what individuals who were moving to Estonia valued most highly when looking for a new home. On the public sector side, one attendee noted that, as their goal ultimately was to make citizens happy new services, such as EREPP, should be supported as it has the potential to create new public value.

Reflecting back on the four goals of this workshop it does appear all four goals were achieved. At the end of the workshop feedback had been received, problems were identified, initial target groups and functionalities were created, and stakeholders from the private sector, public sector, and education sector sat at the same table and discussed how they could work together to create a new service through the use of OGD.

4.2.3 Working with Open Government Data

After the workshop, the initial datasets had been chosen, and the initial functionalities of the new service had been outlined. In order to advance with the pilot development, each OGD set that had been chosen needed to be played with to see how usable it was, what needed to be done to make it usable, and what role it may play in the provision of the initial EREPP functionalities. In this subsection, each selected dataset will be presented along with what work was done to it and how it could be used in the initial service.

- Public Transport Data (Tallinn)
    - soiduplaan.tallinn.ee/data/stops.txt
        - Structure: ID; SiriID; Lat; Lng; Stops; Name; Info; Street; Area; City; Pikas
    - What was done?
        - This file is only available as a text file and thus needed to be converted into a CSV. In addition, many stops in this list were outside of Tallinn so a value of 1 or 0 was appended to each stop. If it had a value of 1 then it was within the city of Tallinn, if it had a value of 0 then it was outside of Tallinn. This would make it easier to use the data for the initial pilot program, as the focus was the city of Tallinn.
    - How it may be used in the initial service offering:
        - Bus stops may be plotted on a map using the latitude and longitude, individuals could then see how many bus stops are close to any address or the distance to the closest bus stop.
- Crime Data
    - Data not yet available for download was received via research request to Ministry of Justice.
        - Structure: CriminalCode; Year(Count); Total
    - What was done?
        - An individual XLSX file was received for each linnaosa inside of Tallinn. The first step to make this data usable was to combine all excel files into one file. As the dataset only contained the criminal code that was violated, a human readable/understandable name needed to be provided. Thus, six categorizations were

created: Offences against the person, offences against public health, offences against property, offences against public peace, offences dangerous to public, and traffic offences. After this has been done empty values were also removed which brought down the number of rows for each linnaosa from over 400 to just 13.

- How it may be used in the initial service offering:
  - Crime data is available only at the linnaosa level; however, this may still be useful. Individuals would be able to type in an address and see what crime has happened in that address' city district, look at change in crime over time, or common types of crime thus giving a general idea about the safety of the area.

- Price Data
  - www.maaamet.ee/kinnisvara/htraru/Start.aspx
    - Structure: Location; Area(m2); Number; Price(Eur)Min; Price(EUR)Max; PricePerUnitArea(EUR/m2)Min; PricePerUnitArea(EUR/m2)Max; PricePerUnitArea(EUR/m2)Median.
  - What was done?
    - This data did not need additional work done to it in order to be useful.
  - How it may be used in the initial service offering:
    - Due to confidentiality issues this dataset does not display transaction values in areas where there have not been at least five transactions [63]. However, this dataset may be used in combination with the property dataset to get an idea about potential sale value. This information may be useful for those who are moving to Tallinn and wish to buy a new apartment and want to make sure they are getting a fair price. It is also possible to use this data in combination with other datasets such as school location/performance and crime data to look at how they may or may not affect price.
- School Data
  - info.haridus.ee/Asutused, www.haridussilm.ee/?leht=alus_yld_0
    - Structure: Name; Type; Teaching_Language; Address

- What was done?
    - In order to map data a latitude and longitude coordinate would be needed, to accomplish this the dataset was imported into R and then the package “GGMAP2” [64] was used to geocode each address thus allowing a latitude and longitude coordinate to be attached to each school.
- How it may be used in the initial service offering:
    - This dataset will initially be used to map schools and kindergartens on a map, if individuals have young children they would then be able to see how close the nearest school is to a new potential address.

- Property Data:
    - www.maaamet.ee/kinnisvara/htraru/Start.aspx
        - Structure: EHIT_ID; EHR_KOOD; NIMETUS; SEISUND_KOOD; SEISUND_NIMI; KAOS_ID_PEAMINE; KAOS_KOOD_PEAMINE; KAOS_PEAMINE; OMANDI_LIIK_KOOD; OMANDI_LIIK_NIMI; AADRESSTEKSTINAKEHTIV; RAJATIS_HOONE; MAAKOND_KOOD; OMAVALITSUS_KOOD; MAX_KORRUSTE_ARV; EH_ALUST_KP; KAV_KASUTUS_KP; AJEH_KASUTALG_KP; AJEH_KASUTLOPP_KP; MARKUSED; SOLTUMATU_VALLASASI; KINNISTAMINEAVALDUSEKUUPÄEV; DATE_CREATED; DATE_UPDATED; ADSOID; ABS_0_KORGUS; EHITISALUNE_MAAPEALNE_PIND; EHITISALUNE_PIND; ESMANE_KASUTUS; KOETAV_PIND; KORGUS; LAIUS; LIFT; MAHT_NETO; MAHT_MAAALUNE; MAX_0_KORGUS; MAX_0_SYGAVUS; PIKKUS; SULETUD_NETOPIND; PIND (üldkasutatav pind); MAAALUSTE_KORRUSTE_ARV; ID; LIIK; NIMETUS; TAHIS; RODUDE_PIND; TERASSIDE_PIND; HOONE_KUJU_ID; HOONE_AADR_ID; DATE_CREATED; DATE_UPDATED; GAASIPAIGAL_ARV; PIND;

TUBADE_ARV; LODZADE_PIND; ADOB_ID; EHR_KOOD; OSA_KOOD; KOOGINISH; KOOK; SISSEPAASU_KORRUS; TUALETTRUUM

- What was done?
  - This OGD set is a 300+mb CSV file with many blank/missing values and encoding issues. When downloading the file the Estonian characters üõöäžš were not present due to encoding issues so a series of gsub replacements was done in order to fix this issue. The next issue that needed to be fixed was issues with data quality. For example, the Estonian word 'Garaaž' could was spelled in multiple ways such as garz, garaaz, gar a az, gara ž, Garage. In order to fix this issue the tool "OpenRefine" was used[1]. This tool allows for large datasets to be searched for values that are similar and then replace these values into one consistent value throughout the dataset. Another issue with this dataset is that the coordinate system is the LEST97 system [65], in order to be compatible with a majority of mapping software programs all coordinates needed to be converted into the international format[2]. Another issue was the large size of the CSV file and large amounts of missing values. To address this, many columns were cutout; the columns that remained dealt specifically with the size, purpose, and location of the property.
- How it may be used in the initial service offering:
  - This dataset provides more in-depth information about specific addresses. For example, it would be possible to type in an address and find out when it was built, how many rooms it has, what the size is, what the state of the property is, etc.

1 For more information on Open Refine, please see: openrefine.org (accessed 03 May, 2017).

2 The R function which was used for this conversion may be found by following the link: github.com/OpenGovIntelligence/EstonianRealEstatePilot/blob/master/Functions/eestiGPSConvert.R

4.2.4 Hackathon – Testing the Waters

While the cleaning of the datasets was progressing, the author of this thesis was approached to act as a mentor at a Garage48 Hackathon[1], sponsored by MKM, which dealt with big and open data. This mentorship role involved discussing and disseminating information about the currently available OGD sets originating from Estonian government organizations. As this hackathon dealt with open data, it also seemed to be the perfect opportunity to see what applications could be developed with the data sets we had thus far obtained and cleaned. The hackathon took place from 21 to 23 October, 2016 and a brief overview of this event will be discussed below.

At the beginning of the hackathon, everyone with an idea was able to pitch his or her topic to the crowd; it was at this point the EREPP service was pitched. Not all ideas successfully made it through this round; however, the EREPP pitch was received with good reviews. After the pitching was over a team was formed with members from the OGI team at Tallinn University of Technology and big data developers from the company of Nortal. As a compromise between these two groups, the datasets, which had been cleaned and obtained by the EREPP team, would be used, but it was to be a commercially oriented service rather than a free service; however, the creation of public value would still be the main goal. Over the next 48 hours, a new service MVP was built which used OGD to rate different addresses based on an individual's preferences[2].

After 48 hours, this idea was presented to the audience where it received an honorary mention for providing valuable location based information. Though no prizes were won, many important goals were achieved by participating in this hackathon.

1. It demonstrated that OGD data could be used to create a new service which provided public value
2. It demonstrated that the creation of an MVP for a new service could be done in a short amount of time
3. It helped bring awareness to the OGD situation within Estonia, this is especially important and was highlighted by Interviewees (Interviews A,C,D,E,F), where

---

1 For a description of this hackathon, please see: garage48.org/events/openbigdata (accessed 03 May, 2017).

2 The website which was created during the hackathon may be found here: www.locata.eu (accessed 03 May, 2017).

hackathons were listed as one of the major ways to drive the OGD movement in Estonia

4. It demonstrated that individuals from different sectors could come together and volunteer their free time to use OGD to build a new service

The hackathon marked a major stage for the Estonian pilot program. This hackathon had provided validation for the datasets which had been collected and cleaned, it provided validation for the idea (a major real estate portal in Estonia expressed interest in the service during the hackathon), and it allowed EREPP to gain more attention among the public and private sectors.

4.2.5 Data Analytics Working Group (Andmeanalüütika Töörühm)

After the hackathon, the focus of the pilot development team remained on gathering feedback and promoting awareness[1]. In essence, the task group, "Andmeanalüütika töörühm", is made up of members from multiple public and private sector organizations. The presentation given by the research team to the working group focused on what the goal of EREPP was, and what barriers had so far been countered. At the end of the presentation, three questions were asked to the audience:

- Has OGD had an effect on your organizational operations? If yes, how?
- How could your organization's data be used to create a new innovative public service?
- How could OGD change how you interact with other stakeholders? How does your organization currently involve outside stakeholders?

What was discovered, and matched the literature and the results of the previous workshop, was that OGD was generally looked upon favourably and as a needed innovation, but that there was no political will or user demand for better OGD. The private sector representatives took a critical approach towards the presented service, the general criticism was that this work seemed to be better suited for private sector agencies and that the data that had thus far been cleaned would be very valuable for some private sector companies' business models; related to this, Interviewee C stated that if more OGD became available, some individuals had discussed creating new

1 For more information on the working group, please see their homepage here: itpraktikud.eesti.ee/dokuwiki/doku.php?id=itari:toogrupid:erasektor:start

businesses to use and exploit said OGD. Many public agencies also echoed previous thoughts in that if this pilot program was developed and delivered successfully, it had the potential to help drive the release and improvement of OGD sets in Estonia (Interviews C,E,F).

4.2.6 Initial Pilot Development

Initially, MKM was intended to develop the new service in cooperation with TTÜ and outside stakeholders. However, there was much organizational push back which eventually led to TTÜ taking the lead in pilot development. The initial development sprint took place between 15 and 18 March, 2017; the initial goal of the sprint was to develop a fairly simple and easy to understand user interface for the pilot project. As EREPP aims to encourage others to participate in the design and development of the service, the code is completely open source and hosted on GitHub[1]. The application is being developed using the open source language "R" and relies heavily on the "R" package "Shiny".

"R" is a language that is best known for its ability to carry out advances statistical computations and graphing techniques [66]. The "Shiny" package for R is a package that allows responsive web applications to be built without an understanding of HTML, CSS, or JS. This, then, lowers the technical understanding needed in order to contribute to the project. Using Shiny and R also allows for prototypes to be developed quite fast, thus assisting to get a functioning MVP service online with less than 200 lines of code (the current code for the new public web service is included in Appendix 5). The initial application allows for individuals to type in an address, find the location on a map, plot schools on a map in Tallinn, and analyze car accident data in Tallinn. The application uses the R package leaflet for presenting data on a map; leaflet is a package that allows for OpenStreetMaps to be accessed via R. A picture of the current UI is shown in Figure 2.

The use of GitHub and developing the service in an agile and open source manner is critical for the success of the pilot program. By using GitHub the developer is able to raise issues and then allow outside contributors to fork the code, address the issue, and then merge the fix back to the master branch. GitHub would also allow for anyone interested in said service to create and add new functionalities to the service, the

[1] To see the current project, please navigate to: github.com/OpenGovIntelligence/EstonianRealEstatePilot

changes simply need to be approved and merged by the project administrator. A third benefit of developing the service in an open source manner and using free/open source tools is that this service is totally exportable. This service should be able to be packaged and used by any other city, country, or town in the world; in order to do this, the datasets simply need to be changed. Overall, using GitHub for service development has the potential to allow for quick collaboration between multiple stakeholders on a new service.

The goal of the pilot is to have a fully functioning service with all of the initial functionalities described in section 4.2.1 present by the end of April 2017; however, this is out of scope for this thesis and thus only the initial UI development can be presented.

,

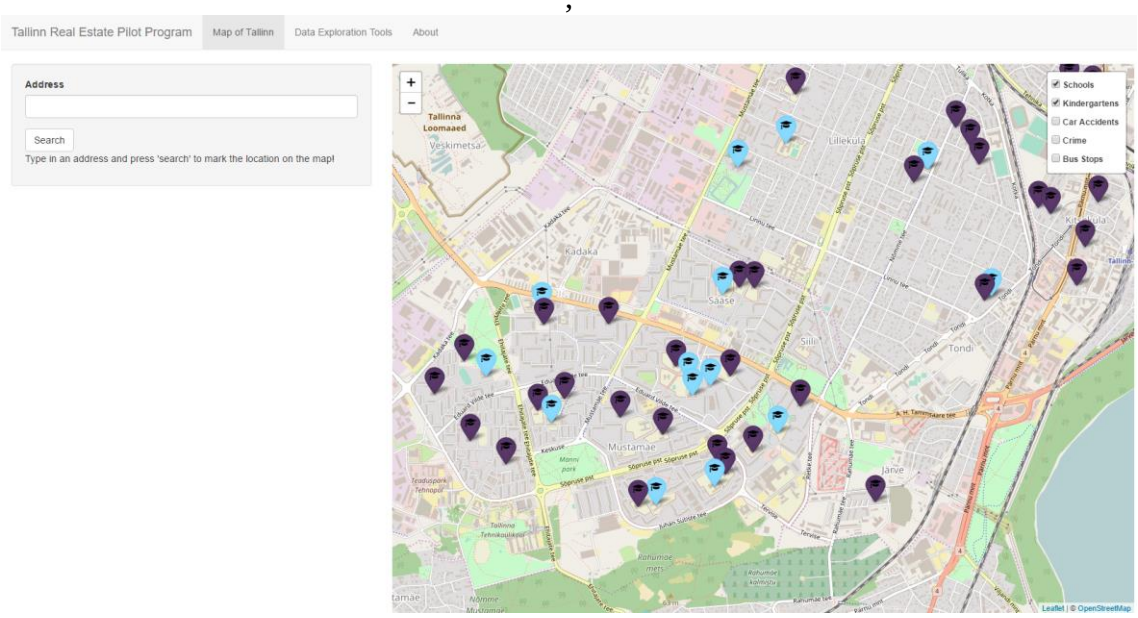

Figure 2. Tallinn Real Estate Pilot Program User Interface

4.3 Summary of the Case

The aim of the case description was to present it in a way consistent with the framework discussed in 2.5. For example, the subsections 4.2.1-4.2.6 may all be divided into one of three stages (co-initiation, co-design, and co-implementation). Sections 4.2.1 deals with the co-initiation stage, at this point OGD sets were being found and other stakeholders were consulted in order to gain access to these datasets so that the service could begin to be designed. Section 4.2.2 through 4.2.5 deal with the co-design aspect of the framework. Throughout this process the initial functionalities of the pilot program were designed in cooperation with other stakeholders. Sections 4.2.5 and 4.2.6 deal with the

co-implementation of the new service as this is when initial functionalities began to be implemented. It is also important to note that there is often overlap between the stakeholders involved, and the events, which are included in each part of the OGD driven public service co-creation cycle. This is demonstrated in section 4.2.5 where both co-design and co-implementation elements were present. Overall, the case aimed to present the overall process which occurred from the inception of the project to the initial implementation.

5 Discussion

The aim of the discussion section is to look back on Chapter 2 and reflect back on the theories and frameworks presented. This chapter will provide a discussion of the propositions which were raised in section 2.6, and also offer improvement and critiques to the framework presented in 2.5.

P1 stated that a change in the current understanding of stakeholders in the public service creation process must take place, if OGD was to be able to contribute effectively to the co-creation of new public services. Along with a change in understanding of the role of stakeholders, there must also be an organizational change in how public services are understood.

Though there are many different understandings of the definition of public service in Estonia, one which is often referred to is as follows: "A public service is a service that the state, local government, or a person in private law performing public duties provides at the will (including the presumable will) of a person for the performance of their legal obligations or exercise of their rights" [67]. In the interviews it was claimed that a public service was, in essence, something that was paid for by public money and carried out by a public agency (Interviews B,C,D,E,F). Thus, according to these definitions, citizens could not (should not) be able to create new public services. This understanding of public services also carries out into the understanding of the roles of the stakeholders in the public service creation process. When asked, interviewees often said that citizens should be consulted at the beginning stages of a new service and asked for feedback throughout (Interviews B,D,E,F). However, when asked if citizens should be able to play a role in the creation and design of a new public service, the answer appears to be no (Interviews B,F). The interviewees did state that though outside stakeholders are currently not able to play a role in public service co-creation, this may change in the future (Interviews B,F).

There seem to be two primary reasons for why stakeholders are not currently viewed as being able to play a role in the co-creation of new public services in Estonia. The first relates back to the definition of a public service in Estonia, and the second is that citizens are referred to as clients or customers rather than as partners or collaborators (Interviews D,E,F). Throughout the case, resistance to the notion that government agencies could work with citizens as partners was clear. The clearest example of this is

through the actions of MKM. During the case TTÜ took over the pilot implementation from MKM. This was an interesting development, as TTÜ and MKM still worked together, but instead of the government agency developing the new public service, a university had taken the lead. This further demonstrates the changing role of stakeholders and the ability of non-governmental actors to play a major role in the development of new OGD driven co-created public services. Though a new public service is in the process of being created, and it does exploit some OGD sources, the effectiveness of the pilot program has been hindered by the lack of access to OGD and the organizational belief that outside stakeholders should not be able to play a part in the co-creation of new public services.

In regards to P2, there seems to be an interesting paradox currently in place in Estonia. On the one hand, some interviewees stated that citizens should not or are not able to play a role in the co-creation of a new public service (Interview B). On the other hand, government agencies do make some of their data open, and this data can be exploited to create public value (as demonstrated by this case). Thus, by releasing open data, government agencies are willingly becoming a participant in the co-creation of new services (whether they mean to or not). OGD relates directly back to the idea that the current understanding of stakeholders in public service design and delivery is changing. When government agencies release open data, citizens have the possibility to use, analyze, and exploit this data. The interactions which take place at this stage between government agencies providing data and service providers who are using it are quite important. In the Estonian case, government agencies were constantly communicated with to discuss issues in relation to data availability, data quality, data structure, etc. This communication accomplished a few different things: it increased communication between service developers and government agencies, it increased awareness of data issues, and, as noted by Interviewee (Interview A), these conversations help government agencies become aware of what data they have, what they do not, and what the current issues are. So, by opening up datasets government agencies allow other stakeholders to create public value through the exploitation of their data, while at the same time gaining valuable information in regards to their own data. Though this interaction takes place and does seem to provide tangible benefits for both government sector and other stakeholders, it is also one of the largest barriers present. The reason for this is that releasing data requires government agencies to acknowledge that there is an alternative

way to create public services and that other stakeholders may come to be seen as partners or as collaborators rather than as customers or dependents.

To further demonstrate the ideas presented in P3, and to relate back to the concept of Government as a Platform, Figure 3 has been created to outline what an OGD driven co-created public service would look like, as well as where and how co-creation happens. Figure 3 also represents an improvement and addition to the model which was presented in section 2.5. Whereas the previous model looked at the process of how a new public service may be co-created, the new figure uses elements of enterprise architecture to create a general overview of a co-created OGD driven public service web application.

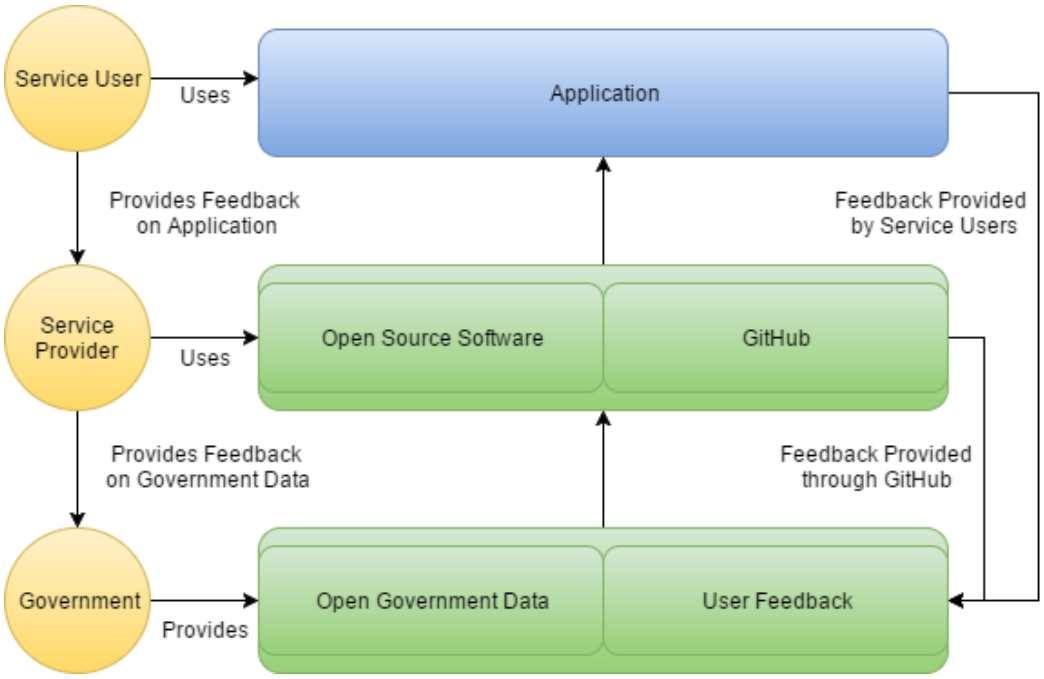

Figure 3. OGD Co-Created Web/Mobile Application General Architecture. Source: Author.

In regards to the stakeholders, the government will always have some role as they are the provider of the OGD. However, service provider and service user are left open. The reason for this is that anyone has the potential to use and exploit OGD to create a new public service, and any stakeholder has the ability to consume this application. It is also interesting to note that the service provider will always, at the same time, be a service user as in this process they are utilizing data which has been opened by government actors. Figure 3, also demonstrates the different ways in which a stakeholder is able to influence the design of the service as well as the flow of feedback. The service provider is able to receive feedback directly from the service user as well as through application usage statistics. Any actor is able to fork the code from GitHub and make a pull request

to influence the design and implementation of the service. The government is also able to receive feedback in regards to their data quality and availability from the service provider. Received feedback and input goes directly back into the cycle and helps drive future releases and changes to the service.

Though the framework presented in 2.5 has been largely applicable to the current case, there does seem to be some room for improvement. For example, though OGD does have the potential to drive the co-creation of new public services, it must be pointed out that not all public services may be co-created using OGD; services which require access to confidential information seem to be the best example here. Furthermore, the contextual requirements for OGD driven public service co-creation need to be integrated into the framework design. In Estonia, some of the biggest barriers emerged from organizational beliefs which are directly opposite the ideas put forth by the framework in 2.5. For example, in order for OGD driven co-creation to take place data must be available, the data must be of usable quality, government organizations must be willing to co-create with outside stakeholders, and new definitions of public services must be in place. The current case demonstrated that all of the contextual factors in the previous sentence played a role in hindering the effectiveness of the EREPP, though it is likely that there are also other contextual barriers which did not manifest in this case and this should be further explored.

In line with P4, the case does seem to confirm that there is a relationship between OGD, coproduction, co-creation, and NPG. If OGD is available, any actor is able to exploit or analyze this data to create new public value. If the definition of a public service put forth by the EC is to be followed, then this creation of new public value may be also viewed as the creation of a new public service. Throughout the process in which OGD becomes exploited, co-creation is occurring, at a minimum, between the government agency and the actor that is exploiting the data. Furthermore, the service provider is also acting a service user at the same time as they are reliant upon the government's open data. Finally, if an application is developed on top of exploited OGD, a complex public service delivery system begins to emerge. These public service delivery systems are discussed and put forth by NPG, and based on this one case it does seem to be true that networks, relationships, and feedback have a strong role in influencing the design, development, and implementation of the new public service.

6 Conclusion and Summary

The idea of opening up government datasets is beginning to take hold with many governments. As a result of this, it appears that new ways of creating public services may be here. However, there is currently a scholarly gap in regards to how OGD can be exploited to create new public services, and how OGD can involve outside stakeholders to co-create these services. The case of the Estonian Real Estate Pilot Program provided the means to gain knowledge which may help address the current gap in the research. The case at hand has the potential to contribute to the current scholarly understanding of the relationship between OGD and co-creation of new public services, which in turn leads to increased levels of coproduced public value.

The case study was rooted in the theory of coproduction, NPG, and OGD. The combination of theories proposes that a new understanding of public services is needed, that OGD has the potential to play a major role in influencing this new understanding, and that OGD can be used to co-create new services, which leads to increased levels of coproduced public value. One way to understand this new relationship between the theories is the concept of GaaP, in essence, OGD is released by government agencies and then exploited by outside stakeholders. This OGD platform plays a key role in enabling OGD driven public service co-creation. The case used these theories to springboard further exploration into the process in which OGD could be used in the co-creation of new public services. This initial exploration does seem to confirm that OGD does lead to a change in our understanding of public service delivery (SQ2): governments begin to act as partners with other stakeholders, anyone can become a service provider, and that these new public services become more effective if there is increased levels of feedback and communication.

In addition to providing increased understanding of the process in which OGD contributes to the co-creation of new public services, the thesis has provided new understandings of the architecture OGD driven co-created public services. In the theoretical framework section, it is stated that there are three overarching ways in which OGD may be exploited for the co-creation of new public services: web/mobile applications, OGD portals, and data analytics. As a result of the case, a general architecture for OGD driven co-created web/mobile applications was created and presented (SQ1). This architecture allows for an easy way to understand how co-created

OGD driven web/mobile applications come into being, how co-creation and coproduction take place, and demonstrates the changing roles of stakeholders in public service delivery.

The initial case of the Estonian Real Estate Pilot seems to confirm that OGD may be able to contribute to the co-creation of new public services, but it also demonstrates that there is currently many factors which impede on this process from occurring, such as: data quality, organizational push back, inadequate legal framework, and lack of government interest (RQ). As the case is limited to the initial process of creating a new public service, further research should be conducted on the later stages of this process (from initial service implementation to completion and usage). Further research could also be conducted on government motivations for resisting or embracing OGD, exploring the general architectures of co-created OGD driven services, and exploring the networks which emerge when OGD is used in the co-creation of new public services.

References

[1] European Commission, “A vision for public services,” p. 16, 2013.

[2] OECD, *Together for Better Public Services: Partnering with Citizens and Civil Society*. 2011.

[3] P. Dunleavy, H. Margetts, S. Bastow, and J. Tinkler, “New public management is dead - Long live digital-era governance,” *J. Public Adm. Res. Theory*, vol. 16, no. 3, pp. 467–494, 2006.

[4] D. Linders, “From e-government to we-government: Defining a typology for citizen coproduction in the age of social media,” *Gov. Inf. Q.*, vol. 29, no. 4, pp. 446–454, 2012.

[5] S. P. Osborne, Z. Radnor, and G. Nasi, “A New Theory for Public Service Management? Toward a (Public) Service-Dominant Approach,” *Am. Rev. Public Adm.*, vol. 43, no. 2, pp. 135–158, 2013.

[6] S. P. Osborne, “Delivering Public Services: Time for a new theory?,” *Public Manag. Rev.*, vol. 12, no. 1, pp. 1–10, 2010.

[7] W. Drechsler, “The rise and demise of the new public management,” *Post-Autistic Econ. Rev.*, vol. 33, no. 14, pp. 17–28, 2005.

[8] S. Osborne, “The new public governance: Emerging perspectives on the theory and practice of public governance,” *From Local Gov. To Local Gov.*, 2010.

[9] V. Pestoff, “Collective action and the sustainability of co-production,” *Public Manag. Rev.*, vol. 16, no. 3, pp. 383–401, 2014.

[10] J. Torfing and P. Triantafillou, “What's in a Name ?: Grasping New Public Governance as a Political-Administrative System,” *Pap. Present. ECPR Gen. Conf. Bordeaux, 4-7. Sept. 2013*, pp. 4–7, 2013.

[11] S. P. Osborne, “The new public governance?” Taylor & Francis, 2006.

[12] S. P. Osborne and K. Strokosch, “It takes two to tango? Understanding the co-production of public services by integrating the services management and public administration perspectives,” *Br. J. Manag.*, vol. 24, no. S3, 2013.

[13] E. Sorensen and J. Torfing, “Enhancing Public Innovation through Collaboration, Leadership and New Public Governance,” *New Front. Soc. Innov. Res.*, pp. 145–169, 2016.

[14] A. Meijer, “Co-production in an Information Age: Individual and Community Engagement Supported by New Media,” *Voluntas*, vol. 23, no. 4, pp. 1156–1172, 2012.

[15] W. H. Voorberg, V. J. J. M. Bekkers, and L. G. Tummers, “A Systematic Review of Co-Creation and Co-Production: Embarking on the social innovation journey,” *Public Manag. Rev.*, vol. 17, no. July 2014, pp. 1–25, 2014.

[16] S. P. Osborne, “Debate: Delivering public services: Are we asking the right questions?,” *Public Money Manag.*, vol. 29, no. 1, pp. 5–7, 2009.

[17] M. Toots, K. McBride, T. Kalvet, R. Krimmer, E. Tambouris, E. Panopoulou, E. Kalampokis, and K. Tarabanis, “A Framework for Data-Driven Public Service Co-Production,” *EGOV-EPART St. Petersburg.*, pp. 1–13, 2017.

[18] M. Janssen, T. Kalvet, R. Krimmer, R. Matheus, B. Roberts, E. Tambouris, K. Tarabanis, M. Toots, and D. Zeginis, “Deliverable 1.1 OpenGovIntelligence Challenges And Needs,” pp. 1–44, 2016.

[19] E. Ostrom, “Metropolitan reform: Propositions derived from two traditions,” *Soc. Sci. Q.*, pp. 474–493, 1972.

[20] E. Ostrom, R. B. Parks, G. P. Whitaker, and S. L. Percy, “Formation of Police and law enforcement policy,” *Policy Stud. J.*, pp. 381–389, 1978.

[21] S. P. Osborne, Z. Radnor, and K. Strokosch, “Co-Production and the Co-Creation of Value in Public Services: A suitable case for treatment?,” *Public Manag. Rev.*, vol. 18, no. 5, pp. 639–653, 2016.

[22] E. B. Sharp, “Toward a New Understanding of Urban Services and Citizen Participation: The Coproduction Concept,” *Am. Rev. Public Adm.*, vol. 14, pp. 105–118, 1980.

[23] G. P. Whitaker, “Coproduction : Delivery in Service Participation Citizen,” *Public Adm. Rev.*, vol. 40, no. 3, pp. 240–246, 1980.

[24] E. Ostrom, “Crossing the Great Divide : Synergy , and Development,” *World Dev.*, vol. 24, no. 6, pp. 1073–1087, 1996.

[25] E. Vigoda, “From Responsiveness to Collaboration: Governance, Citizens, and the Next Generation of Public Administration,” *Public Adm. Rev.*, vol. 62, no. 5, pp. 527–540, 2002.

[26] V. Pestoff, “Co-production as a social and governance innovation in public services,” *Polityka Spoleczna*, vol. 11, no. 1, pp. 2–8, 2015.

[27] B. Verschuere, T. Brandsen, and V. Pestoff, “Co-production: The State of the Art in Research and the Future Agenda,” *Voluntas*, vol. 23, no. 4, pp. 1083–1101, 2012.

[28] S. P. Osborne, Z. Radnor, T. Kinder, and I. Vidal, “The SERVICE Framework: A Public-service-dominant Approach to Sustainable Public Services,” *Br. J. Manag.*, vol. 26, no. 3, pp. 424–438, 2015.

[29] W. Drechsler, “Good Governance and New Public Management : The Case for Estonia.”*Trames* 8(58/53). 2004.

[30] R. B. Denhardt and J. V. Denhardt, “The New Public Service: Serving Rather than Steering,” *Public Adm. Rev.*, vol. 60, no. 6, pp. 549–559, 2000.

[31] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, “Benefits, Adoption Barriers and Myths of Open Data and Open Government,” *Inf. Syst. Manag.*, vol. 29, no. 4, pp. 258–268, 2012.

[32] F. Gonzalez-Zapata and R. Heeks, “The multiple meanings of open government data: Understanding different stakeholders and their perspectives,” *Gov. Inf. Q.*, vol. 32, no. 4, pp. 441–452, 2015.

[33] “What is Open Data?” [Online]. Available: http://opendatahandbook.org/guide/en/what-is-open-data/. [Accessed: 21-Apr-2017].

[34] S. M. Burwell, S. Vanroekel, T. Park, and U. S. Chieftechnolo, “Open Data Policy - Managing Information as an Asset,” *Exec. Off. Pres.*, vol. M-13-13, pp. 1–12, 2013.

[35] M. Toots, K. Mcbride, T. Kalvet, and R. Krimmer, “Open Data as Enabler of Public Service Co-creation : Exploring the Drivers and Barriers,” pp. 1–11.

[36] S. S. Dawes, L. Vidiasova, and O. Parkhimovich, “Planning and designing open government data programs: An ecosystem approach,” *Gov. Inf. Q.*, vol. 33, no. 1, pp. 15–27, 2016.

[37] M. Janssen, “Open Data , Information Processing and Datification of Government,” pp. 2668–2670, 2017.

[38] M. Foulonneau, S. Martin, and S. Turki, “How Open Data are turned into services?,” *Lect. Notes Bus. Inf. Process.*, vol. 169 LNBIP, no. June, pp. 31–39, 2014.

[39] M. Foulonneau, S. Turki, G. Vidou, and S. Martin, “Open data in service design,” *Electron. J. e-Government*, vol. 12, no. 2, pp. 99–107, 2014.

[40] P. Conradie and S. Choenni, “On the barriers for local government releasing open data,” *Gov. Inf. Q.*, vol. 31, no. S1, pp. S10–S17, 2014.

[41] A. Zuiderwijk, M. Janssen, S. Choenni, R. Meijer, and R. S. Alibaks, “Socio-technical Impediments of Open Data,” *Electron. J. Electron. Gov.*, vol. 10, no. 2, pp. 156–172, 2012.

[42] A. Zuiderwijk, M. Janssen, and S. Choenni, “Open Data Policies: Impediments and Challenges,” in *12th European Conference on eGovernment (ECEG 2012)*, 2012, pp. 794–802.

[43] M. Young and A. Yan, “Civic Hackers’ User Experiences and Expectations of Seattle’s Open Municipal Data Program,” *Proc. 50th Hawaii Int. Conf. Syst. Sci.*, pp. 2681–2690, 2017.

[44] “AccessMap.” [Online]. Available: https://www.accessmap.io/. [Accessed: 21-Apr-2017].

[45] “data.gov.uk.” [Online]. Available: https://data.gov.uk/. [Accessed: 21-Apr-2017].

[46] T. Schenk, “Food Inspection Forecasting - City of Chicago.” [Online]. Available: https://chicago.github.io/food-inspections-evaluation/. [Accessed: 21-Apr-2017].

[47] R. Krimmer, T. Kalvet, M. Toots, and K. McBride, “Deliverable 2.1 OpenGovIntelligence framework--first release.”

[48] K. Taveter, “On Agile Design in Software Engineering and beyond,” 2015.

[49] K. Beck, M. Beedle, A. Van Bennekum, A. Cockburn, W. Cunningham, M. Fowler, J. Grenning, J. Highsmith, A. Hunt, R. Jeffries, and others, “Manifesto for agile software development,” 2001.

[50] J. Highsmith and A. Cockburn, “Agile software development: The business of innovation,” *Computer (Long. Beach. Calif).*, vol. 34, no. 9, pp. 120–127, 2001.

[51] P. Runeson, M. Host, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering*. 2012.

[52] R. K. Yin, *Case study research: Design and methods*. Sage publications, 2013.

[53] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design Science in Information Systems Research,” *MIS Q.*, vol. 28, no. 1, pp. 75–105, 2004.

[54] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A Design Science Research Methodology for Information Systems Research,” *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–78, 2007.

[55] S. T. March and V. C. Storey, “Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research,” vol. 32, no. 4, pp. 725–730, 2016.

[56] A. L. Delbecq, A. H. de Ven, and D. H. Gustafson, *Group techniques for program planning: A guide to nominal group and Delphi processes*. Scott, Foresman Glenview, IL, 1975.

[57] F. Innovation, “OpenGovIntelligence Fostering Innovation and Creativity in Europe through Public Administration Modernization towards Supplying and Exploiting OpenGovIntelligence framework – first release,” pp. 1–44.

[58] T. Kalvet, “Innovation: a factor explaining e-government success in Estonia,” *Electron. Gov. an Int. J.*, vol. 9, no. 2, pp. 142–157, 2012.

[59] *Public Information Act – Riigi Teataja*. 2016.

[60] NullByte, “How to Use Google to Hack(Googledorks),” 2015. [Online]. Available: https://null-byte.wonderhowto.com/how-to/use-google-hack-googledorks-0163566/. [Accessed: 21-Apr-2017].

[61] S. Chamberlain, H. Sharp, and N. Maiden, “Towards a framework for integrating agile development and user-centred design,” in *International Conference on Extreme Programming and Agile Processes in Software Engineering*, 2006, pp. 143–153.

[62] K. D. Mcbride, “Tallinn User Workshop Writeup,” 2016.

[63] MaaAmet, “Real property price statistics.” [Online]. Available: http://www.maaamet.ee/kinnisvara/htraru/Start.aspx. [Accessed: 21-Apr-2017].

[64] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009.

[65] MaaAmet, “Maa-amet: Geoportaal: L-EST97 ja teised koordinaatsüsteemid.” [Online]. Available: http://geoportaal.maaamet.ee/est/Andmed-ja-kaardid/Koordinaatsusteemid-ja-kaardilehtede-jaotused/L-EST97-ja-teised-koordinaatsusteemid-p173.html. [Accessed: 21-Apr-2017].

[66] R Development Core Team, “R: A Language and Environment for Statistical Computing.” Vienna, Austria, 2008.

[67] Estonian Ministry of Economic Affairs and Communications, “Green Paper on the Organisation of Public Services,” 2013.

Appendix 1 – List of Interviewees

1. Interview A – Representative of a NGO – Audio Recording, 10.03.2017
2. Interview B – Employee of Ministry of Environment – Interview via Email, 13.03.2017
3. Interview C – Representative of a NGO – Audio Recording, 14.03.2017
4. Interview D – Employee of Ministry of Education – Audio Recording, 20.03.2017
5. Interview E – Employee of Ministry of Economic Affairs and Communications – Audio Recording, 22.03.2017
6. Interview F – Employee of Ministry of Economic Affairs and Communications (involved with EREPP) – Audio Recording, 22.03.2017

Appendix 2 – List of Workshop Participants' Organizations

1. European Commission
2. Mooncascade
3. Ministry of Economic Affairs and Communications
4. Finance Ministry
5. Teleport
6. Land Board
7. Estonian Board of Statistics
8. Tallinn City Government
9. National Registers and Information Systems Center

Appendix 3 – Survey Respondents

- EENG1 – Non-governmental/civil society organization, Senior Expert
- EENG2 – Private company, Product Owner and Board Member
- EENG3 – University/Research Institution, Senior Researcher
- EEPA1 – Central/Federal Government, Head
- EEPA2 – Central/Federal Government, Chief Architect
- EEPA3 – Central/Federal Government, Expert
- EEPA4 – Central/Federal Government, Specialist
- EEPA5 – Central/Federal Government, Advisor
- EEPA6 – Central/Federal Government, Head

Appendix 4 – Interview Questions

1. How would you define a public service?
2. How do you believe citizens should be involved in the service creation process?
3. Can you provide any examples of government – citizen interaction in the service creation process?
4. Do you think that open data could be used in this service design process?
5. What are the biggest reasons that government agencies do not interact with citizens in the service creation process?

Appendix 5 – Web Service Code

UI.R

```
library(shiny)
library(leaflet)
library(ggmap)
navbarPage(
  "Tallinn Real Estate Pilot Program",
  tabPanel("Map of Tallinn",
           sidebarLayout(
             sidebarPanel(
               textInput("address", "Address", value = ""),
               verbatimTextOutput("value"),
               actionButton("addressButton", "Search"),
               p(
                 "Type in an address and press 'search' to mark the location on
the map!"
               )
             ),
             mainPanel(
               tags$style(type = "text/css", "#map {height: calc(100vh - 150px)
!important;}"),
               leafletOutput("outputmap", width = "100%", height = "800px"),
               hr(
                 "This project has received funding from the European Union's
Horizon 2020 research and innovation programme under grant agreement No
693849."
               )

             )
           )),
  tabPanel(
    "Data Exploration Tools",
    a("Cube Explorer", href =
"http://wapps.islab.uom.gr/CubeVisualizer/crashes/", target =
        "_blank"),
    br(),
    a("QB OLAP browser", href = "http://wapps.islab.uom.gr/qbOLAPbrowser",
target =
        "_blank")
  ),
  tabPanel(
    "About",
    "The Estonian pilot program is one of 6 pilot programs being carried out
    by the OpenGovIntelligence project. This pilot program is being carried out
by The Estonian Ministry
    of Economic Affairs and Communications and Tallinn University of
Technology.
    The purpose of this pilot program is to fight information asymmetry in the
```

```
real estate market and
    provide an easy way to access real estate data. The pilot is intended to
give real estate agents, property developers,
    investors, and those involved in the real estate market (buyers, sellers,
renters, students, new arrivals etc.)
    a deeper knowledge of the marketplace.",
    br(),
    br(),
    "For questions on current stage of pilot please contact: ",
    a("Keegan.mcbride@ttu.ee", target = "_blank")
  )
)
```

Server.R

```
library(shiny)
library(leaflet)
library(data.table)
library(ggmap)


##Load Data Here
schoolData <- readRDS("datasets/schoolData.rds")
lasteaedData <- readRDS("datasets/lasteaedData.RDS")
crashData <- readRDS("datasets/crashDataCleanedFixed.rds")


##Subset and Select data here
names(schoolData)[6] <- paste("Type")
lasteaedData <- subset(schoolData, schoolData$Type == "lasteaed", select =
V1:ads_oid)
schoolData <- subset(schoolData, schoolData$Type != "lasteaed", select =
V1:ads_oid)


##Create Custom Icons Here
schoolIcons <- awesomeIcons(
  icon = 'graduation-cap',
  markerColor = 'lightblue',
  library = 'fa',
  iconColor = 'black'
)


lasteaedIcons <- awesomeIcons(
  icon = 'graduation-cap',
  markerColor = 'darkpurple',
  library = 'fa',
```

```
  iconColor = 'black'
)


addressIcons <- awesomeIcons(
  icon = 'home',
  markerColor = "red",
  library = 'fa',
  iconColor = 'black'
)


function(input, output, session) {

  ##Initializes the leaflet map for the page.
  output$outputmap <- renderLeaflet({
    map <-
      leaflet() %>% addProviderTiles(providers$OpenStreetMap.Mapnik)  %>%
      setView(lng =  24.753574,
              lat = 59.436962,
              zoom = 12) %>%
      addAwesomeMarkers (
        lng = schoolData$Lon,
        lat = schoolData$Lat,
        icon = schoolIcons,
        popup = paste(
          "Name:",
          schoolData$Nimi,
          "<br>",
          "Type:",
          schoolData$Type,
          "<br>",
          "Address:",
          schoolData$Aadress
        ),
        group = "Schools"
      ) %>%
      addAwesomeMarkers (
        lng = lasteaedData$Lon,
        lat = lasteaedData$Lat,
        icon = lasteaedIcons,
        popup = paste(
          "Name:",
          lasteaedData$Nimi,
          "<br>",
```

```
            "Type:",
            lasteaedData$Type,
            "<br>",
            "Address:",
            lasteaedData$Aadress
          ),
          group = "Kindergartens"
        ) %>%

        addAwesomeMarkers(
          lng = crashData$Lon,
          lat = crashData$Lat,
          clusterOptions = markerClusterOptions(),
          popup = paste(
            "Date:",
            crashData$Kuupäev,
            "<br>",
            "Time:",
            crashData$Kellaaeg,
            "<br>",
            "Damage (EUR):",
            crashData$Kahju.suurus..euro.
          ),
          group = "Car Accidents"
        ) %>%

        addLayersControl(
          overlayGroups = c("Schools","Kindergartens", "Car Accidents",
"Crime", "Bus Stops"),
          options = layersControlOptions(collapsed = FALSE)
        ) %>%
        hideGroup("Schools")    %>%
        hideGroup("Kindergartens")    %>%
        hideGroup("Car Accidents")     %>%
        hideGroup("Crime")     %>%
        hideGroup("Bus Stops")
    })


    ##Geocoding of custom address input
    geocoding <-
      eventReactive(input$addressButton, {
        geocode(input$address)
      })
```

```
  #Takes an address input and places a marker on the map at given location
  observeEvent(input$addressButton, {
    newMarker <- geocoding()
    leafletProxy('outputmap') %>% addAwesomeMarkers(
      lng = newMarker$lon,
      lat = newMarker$lat,
      icon = addressIcons,
      label = input$address
    )
  })

}
```

Acknowledgements

I would like to say thank you to my supervisors Robert Krimmer and Maarja Toots. Robert, thank you for helping me throughout this whole process, your thoughts and comments were incredibly beneficial; you have also helped me grow in my current academic career and an individual and for this I am incredibly thankful.

Maarja, thank you for always being willing to listen to my ideas, provide feedback on them, and being involved throughout the whole process.

Secondly, I would like to thank the e-Governance department for supporting my studies. The e-Governance study program has allowed me to conduct interdisciplinary research and I am incredibly thankful for this opportunity. The support and knowledge passed on by the department has played an instrumental role in giving me the opportunity to study, learn, and research the field of e-Government.

Next , I would also like to thank the staff and faculty at both Ragnar Nurkse Department of Innovation and Governance as well the e-Governance department. Having the ability to conduct research with the department has been an honor.

I would like to thank all of those who helped critique and read my thesis, especially Veiko Lember, Margit Kirs, and Karin Oolu.

Finally, thank you to my family at home, my fiancée Kelian Kaio, and my friends in the e-Government program for playing a role in supporting me throughout the entire duration of the course.