**DOCTOR OF ECONOMICS AND BUSINESS MANAGEMENT**

# Open Data in Cities

## A Role Perspective on Misalignments in an Open Data Value Chain

Gebka, Elisabeth

*Award date:*
2022

*Awarding institution:*
University of Namur

UNIVERSITÉ DE NAMUR

FACULTÉ
DES SCIENCES ÉCONOMIQUES,
SOCIALES ET DE GESTION

# Open Data in Cities

## A Role Perspective on Misalignments in an Open Data Value Chain

A Doctoral Thesis Submitted in August 2022 by

**Elisabeth Gebka**

In fulfilment of the requirements for the degree of Doctor in Economics and Business Management at the University of Namur

FACULTY OF ECONOMICS, SOCIAL SCIENCES AND BUSINESS ADMINISTRATION
UNIVERSITY OF NAMUR | RUE DE BRUXELLES 61, NAMUR

## Doctoral Committee

**Prof. Jean-Yves Gnabo (President)**
Faculty of Economics, Social Sciences, and Business Administration
University of Namur, Belgium

**Prof. Annick Castiaux (Supervisor)**
Faculty of Economics, Social Sciences, and Business Administration
University of Namur, Belgium

**Prof. Mélanie Latiers**
Faculty of Economics, Social Sciences, and Business Administration
University of Namur, Belgium

**Prof. Catherine Elsen**
Faculty of Applied Sciences
University of Liège, Belgium

**Prof. Corentin Burnay**
Faculty of Economics, Social Sciences, and Business Administration
University of Namur, Belgium

**Prof. Ulf Melin**
Department of Management and Engineering, Information Systems, and Digitalization
Linköping University, Sweden

# Foreword

At the beginning of my research, I had never heard about open data. I think it is important for the readers to know from where I started, based on what previous knowledge and encounters my curiosity grew, to eventually, make me see open data as an interesting research topic.

I have a background in user-centred innovation and citizen participation, which I used to scrutinize open data initiatives addressed to the citizens. It helped me suspect possible misalignments, especially between the actors and roles involved in the open data value chain. The possible misalignments became the research problem of my thesis, and the roles a core perspective to investigate them. Hence, my research purpose was to explore, understand, and describe how the role concept can bring new insights into the possible misalignments in the open data value chain, envision new solutions, and increase the chance to realise the expected outcomes. Specifically, my research investigates the roles of the citizens, the municipalities, and the end-users.

While conducting my research, I found myself standing at the crossroads of several disciplines, building my reasoning with concepts coming from, among others, design thinking, co-creation, open innovation, and open data. At the time of writing the thesis, I realised that my audience could be unfamiliar with some of the concepts and be experts in others.

Therefore, I made editorial choices that aim at providing as many insights as possible for the readers to understand my perspective on open data while giving an overview of the previous knowledge I used. I introduce my research topic the way I came across open data: I develop my journey in the world of open data, anchored in empirical cases, to, then, posit my research problem, purpose, and questions. In the background, I choose to introduce the schools of thought influencing my understanding of open data, with the risk that some sections give an "*I-already-know*" feeling to the experts in the field. To help the readers navigate between the concepts, I also list the most used ones with their definitions on p. 7.

I hope that this thesis will arouse the curiosity of the readers and encourage them to explore original ways, outside of their discipline, to tackle open data's challenges.

# Executive summary

Nowadays, public authorities, civil society associations, scholars and enthusiastic practitioners encourage municipalities to publish their data as "open data". Open data is defined as data that can be freely used, re-used, and redistributed by anyone for any purpose - subject only, at most, to the requirement to attribute and share-alike. The people or organisations that publish data take the role of the publisher, and those reusing it take the role of the re-user. The higher purpose of open data is to create value, and a way to visualise the value creation is in a value chain. It shows how actors, in certain roles, contribute to a process to reach expected outcomes. Those of open data are to foster innovation (economic value), participation in governance, and transparency (public and social value). However, open data is not reaching full fruition. In cities, the roles of the citizens, end-users and municipalities in open data initiatives are ambiguous in previous research, and the participation of laypersons is scarce. I suspect that it can be a source of misalignments in the value chain. A misalignment consists of dysfunctions and unfitted arrangements between the components of a process and actors in their roles, tasks, and purposes that can cause problems and impede the realisation of the expected outcomes. Hence, my thesis purpose is to explore the concept of role to bring new insights into possible misalignments in the open data value chain.

I survey the different participatory paradigms nurturing the rise of open data in the last decades, such as open innovation and open government. In this literature, the concept of participation can have different meanings and forms among the actors involved. I use this knowledge to synthesise an open data value chain and pinpoint possible misalignments. In open data practices and previous literature, I suspect them at three levels: between the actors (e.g., municipalities, developers, and citizens) and roles (open data publishers, re-users, and end-users); between the actors and their respective expected outcomes; and between the roles and process' main components (i.e., input, output, needs, expected outcomes). I use these possible misalignments to highlight three research gaps: the fuzzy citizens' roles in open data initiatives, the diverse municipalities' roles and their expectations towards their users, and the lack of end-user involvement and integration of their information needs.

I follow a pragmatic constructivist paradigm, using two methods: a multiple case study, with qualitative abductive research, and Design Science Research. The produced knowledge and recommendations should help the actors reflect on their own roles and alignment with others' roles. This thesis contributes to both research and practice, showing the relevance of the concept of role in bringing new insights into possible misalignments and start solving them. The specific contributions of the three studies are a conceptual framework representing the citizens' roles in an open data value chain, a typology of municipalities' roles and expectations towards their users, and design principles to develop methods and capture end-users´ information needs. For research, it implies using the role combinations, issues, and expectations as a new perspective to study and solve the misalignments in the open data value chain. For practice, the findings prompt the practitioners' critical thinking and reflections on their assumptions about other roles and actors involved, and the implementation of their expected outcomes. It implies making one's role in interactions with other actors and envisioning singular value chains for singular expected outcomes, which goes against the current "open arena" thinking of open data's rhetoric.

Keywords: open (government) data, innovation, citizens, roles, value chain

# Acknowledgements

This thesis would not have been completed without the support and collaboration I found among my colleagues, friends, and family. I would like to thank, especially:

Annick, for her trust all along this journey and our regular meetings ending up with discussions about everything else than the thesis. I enjoyed sharing ideas and brainstorming together. Thank you for allowing me to reconcile research with my other professional projects and private life. It helped me to get the energy to fight and overcome the challenges inherent to a thesis.

My sub-supervisors, Catherine and Mélanie, for their support and interest in my work, even if it went beyond their initial scope of research.

My jury members, Ulf and Corentin who accepted to share their expertise and give a bit of their time to improve my thesis.

My colleagues from the Trakk. Véronique, who warmly welcomed me into the department with a very appreciated coffee on my first day. Claire-Anaïs, with who I shared my initial steps in research and a memorable conference in Manchester. I will never forget that hotel and room next door, a crime scene that spiced up our stay. And Vanessa that I occasionally helped willingly in our common engagement for the scout movement. You three opened your office when I felt a bit isolated in my private one, which I would have loved to use more if the covid pandemic had not sent us all home.

My colleagues from the Computer Science faculty, Antoine and Anthony. You invited me to join a collaborative research project with a Swedish colleague, Jonathan. This project gave a real start to my thesis and an unexpected turn to my private life.

My colleagues from Linköping University. Ulf, again, who included me in the life of the department INDIG from my first visit (a Christmas dinner, a writing seminar). Johanna, who welcomed and hosted me two times. My stay at Linköping broadened my research network, among others in SWEG and the e-Government community, which helped concretise three papers.

Karin, co-author, friendly reviewer, and friend who helped me several times with her relevant feedback. I remember also our weekend of work in Fjällbacka, an amazing place to take resourceful breaks in the beautiful nature of Sweden's West Coast.

My partners from Futurocité, Nicolas, Julie, and Valérie. You gave me access to a valuable field that contributed to my understanding of open data and the completion of this thesis. It was a pleasure to work with you on the program "Ouvrir ma Ville" and the hackathon "Citizens of Wallonia". Your kindness made me feel part of your team.

Last but not least, Jonathan. Partner in research and now in life. You gave me the strength to keep on with your unconditional support, both intellectual and emotional. Thank you for your patience, your love, and the most beautiful output of our "collaboration": Lucia. Her first kicks, first smile, and first laughs cheered me up in the last mile of this journey.

## Table of content

## List of figures



## List of tables

## List of concepts and definitions

| | |
|---|---|
| **Actor** | Entity (a person or group of people, organisation, company) who does something or participates in something. |
| **Data** | Unprocessed information, representation of facts that can be captured, manipulated and stored (in, for example, databases or repositories). |
| **End-user** | Person or organisation that a service or product is designed for. The term is based on the idea that the "end goal" of a product or service is to be useful to the consumer. |
| **Information** | Told fact, data that people have made sense of in a context. Information related to other information provides knowledge. |
| **Information need** | Secondary need, which appears when people realise that they lack knowledge on a subject. They may wish to satisfy this secondary need, motivated by a higher need, namely, to accomplish a task or work effectively, to solve a problem, or to pursue a hobby or personal interest in a satisfactory manner. It is specific to each individual and varies according to the prior knowledge accumulated, the status occupied in society, and the context in which this need appears. |
| **Misalignment** | Dysfunction and unfitted arrangement between the components of the process, or the actors in their respective roles, tasks, and purposes, that can cause problems and hinder the realisation of the expected outcomes. |
| **Open data** | Data that can be freely used, re-used, and redistributed by anyone for any purpose, subject only to the requirement to attribute and share-alike. |
| **Open government data** | Publicly produced and funded open data, usually published by public administrations, governments, public institutions or agencies. |
| **Outcome** | Result or effect of an output generated when an output is consumed and completes its higher purpose. |
| **Paradigm** | Set of theories that explain the way a particular subject is understood at a particular time. In a non-academic context, it refers to a model of something or a very clear and typical example of something. |
| **Publisher** | Organisation or person who releases data to be reused by anyone for any of its purposes. |
| **Rationale** | Reasons or intentions that cause a particular set of beliefs or actions. |
| **Re-user** | Organisation or person who reuses data, for example, to find facts, create visualisations, or applications. |
| **Role** | Position, function, or purpose that an actor has in a situation, process, organisation, society, or relationship. |
| **User** | Organisation or person who uses a product or service. |
| **Value chain** | Description of the full range of activities, which are required to bring a product or service from conception, through the different phases of production, and delivery to final consumers. |

# Chapter 1. Introduction

## 1.1. The promise of open data: value and expected outcomes

I started my thesis by exploring the dynamics of participation and innovation in cities, especially within the fields of open innovation and smart cities. The research project was part of a large research portfolio called "Wal-e-Cities", funded by the European Regional Development Fund (ERDF H2020) and the Walloon Region, in Belgium. It aimed at supporting the development of an ecosystem of smart cities in Wallonia by investigating the opportunities offered by the new technologies and data. In the literature, Smart cities are presented as a vision for the future of cities where the main idea is to improve the well-being of the citizens in urban areas and encourage their participation in governance and innovation, with the support of data, and Information and Communication Technologies (ICT) (Caragliu, Bo, & Nijkamp, 2009; Nam & Pardo, 2011). Hence, the initial objectives of my project were to explore new ways to create value and new services in cities, with enhanced citizen participation in the innovation process and an enhanced release and reuse of data, both considered assets to tackle the current challenges faced by cities.

In my initial literature review, I noticed that open innovation and smart cities have in common the idea of involving outsider innovators or citizens in solutions development. Open innovation relies on knowledge transfers, inside or outside of the firms, to increase innovation success and market shares (Chesbrough, 2003). Similar ideas broke through the public sector, within the New Governance Paradigm (S. P. Osborne, 2006) and Public Service Logic (S. P. Osborne, 2018) where it is believed that the creation of public value is no longer the prerequisite of the administration: *public services can be created by any actor, as long as there is a focus on public value*. It is acknowledged that stakeholders, such as citizens and consumers, can innovate (von Hippel, 1994), share user expertise and ideas (Visser, Stappers, van der Lugt, & Sanders, 2005), or resources (Nyström, Leminen, Westerlund, & Kortelainen, 2014), and contribute to the development of solutions and decision-making processes. They can result in new digital services, products, open governance in public management, or participatory urban planning.

By exploring studies and projects using data in smart cities, I came across the concept of open data, which seemed to combine the ideas of citizen participation and open innovation. I became more curious after a visit to the website of Open Knowledge Belgium, the local chapter of a foundation that promotes open knowledge and open data. It stated the following definition: "***open data*** *is data that can be freely used, re-used, and redistributed by anyone for any purpose - subject only, at most, to the requirement to attribute and share-alike"* (Open Knowledge Foundation, 2015). With such a definition, I wanted to find out what value open data could bring to society, especially in cities to citizens.

I learned that the Open Knowledge Foundation and its members are driven by values such as transparency, democracy, access, and participation. They claim that openness should enhance universal participation: anybody should have access to data and information to create and share knowledge without restriction, for self-empowerment, improved efficiency of products, services or governments, and democratic control (Open Knowledge Foundation, 2015). I soon realised that the public authorities and municipalities (local governments) also surf on this wave and promote open data for innovation, participation, and governance. Open data should help the citizens to be more informed, which, in democratic processes, should improve citizens' ability to contribute, understand and accept decisions affecting them, and shape the situations in which they live (Birkinshaw, 2006; Meijer et al., 2012 as cited in Ruijer et al., 2017).

In Europe, open data became more visible after the European Public Sector Information directive (European Parliament & European Council, 2003) encouraged governments to release publicly produced and funded data, called **open government data** (OGD) (Zhang, Hua, & Yuan, 2017). Since then, the municipalities have become one of the most important sources of reused open data (Berends, Carrara, Engbers, & Vollers, 2017), and open data initiatives are said to merge with smart city initiatives (Ojo, Curry, & Zeleti, 2015). The reuse of data is expected to foster innovation, create economic value (new jobs, start-ups, services), and improve public governance and transparency. In sum, open data is expected to create economic, public, and social value (Attard, Orlandi, Scerri, & Auer, 2015).

I was intrigued by the expected outcomes of open data, which seemed promising, and I was tantalized by the underlying values of democracy and participation. I saw in open data an interesting research topic that connects the dots between my two phenomena of interest: citizen participation and open innovation in the context of cities.

At the same time, as a lay citizen, I had a hard time understanding and seeing the impact of open data in my life: like most citizens (Weerakkody, Irani, Kapoor, Sivarajah, & Dwivedi, 2017), I am not capable of reusing data by myself and was not aware of collaborative open data initiatives. Where was the value created by open data and how could I participate or benefit from it? This realisation, as a growing cloud in the tale of open data promises, was the spark of my thesis. Starting from my role as a citizen of a municipality publishing data, I dug deeper into the roles involved in the process of value creation with open data. I wanted to understand how the open data outcomes can be realised with and for the citizens, and if they are not, the possible sources of problems in the process. Hence, I looked at the value creation process through a *value chain* and conceptualised its problems as *possible misalignments*.

## 1.2. In the shadow of open data: ambiguous roles and possible misalignments in the value chain

To understand how open data can lead to value creation and who is involved, I use a process perspective that I portray as a **value chain**. "*A value chain describes the full range of activities which are required to bring a product or service from conception, through the different phases of production, and delivery to final consumers"* (Kaplinsky & Morris, 2001, p. 4). The value chain is a concept of business management that became popular after the seminal work of Porter (1985). A value chain has several benefits as an analytical framework. It discloses all the activities required for a product or service to gain value, which implies understanding the way different actors contribute and capture value. It focuses on the dynamics and linkages between actors, which goes beyond firm-specific analyses. At the same time, it sets clear boundaries: those of a product or service development (opposite to sectorial or ecosystem analyses, where boundaries might be blurry) (Kaplinsky & Morris, 2001). I believe a value chain has also a high explanatory potential despite being a simplification of reality and it puts value creation in the foreground, which was one of the Wal-e-cities research project's priorities.

An important preliminary consideration to conceptualise an open data value chain and later, conceptualise its possible misalignments, is to distinguish the concept of **role** from the concept of **actor.** Following the Collins Dictionary, I define an **actor** as an entity (a person or group of people, organisation, company) who does something or participates in something (Collins Dictionary, n.d.). In other words, an actor is a participant in an action or process (Oxford Dictionary, n.d.). A **role,** however, is the position, function, or purpose that an actor has in a situation, process, organisation, society, or relationship (Cambridge Dictionary, n.d.-c). The

roles can be defined by the tasks they encompass, which means that the more we describe a process, the more we can create roles or extend their scope. Two important characteristics of roles, as presented in Figure 1, are that they can be shared between actors (e.g., a manager can have two employees occupying the function of the assistant) and an actor can take several roles at the same time (e.g., one of the assistants can also be a worker representative in the trade union).

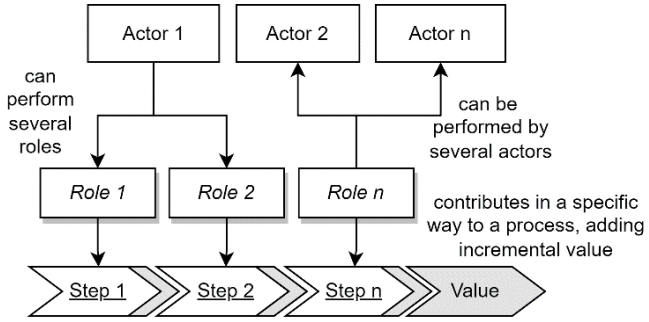


*Figure 1. Actors and roles in a value chain*

The process to produce value with open data requires at least two roles (Zuiderwijk, Janssen, & Davis, 2014): the **publisher**, an organisation or a person who releases data (***an input***) to be reused by the **re-users** for any of their purposes. Then, the re-users can clean and release new data for their peers, investigate the data for facts, and create visualisations, or applications, for their own use (Davies, 2010). These are the ***outputs*** of the process. The re-users can also make their reuses available to others (e.g., consumers, citizens) and become intermediaries (van Schalkwyk, Willmers, & McNaughton, 2016) for those I will call **end-users**. In my definition, the end-users are not in direct contact with the data, because they are not capable or interested in reusing it. They are thus dependent on re-users to capture the benefits of open data. In the end, ***the output*** can generate ***an outcome***. An outcome is the result or effect of an output, the fulfilment of a higher purpose. For example, as defined by Benington and Moore (2010), public value is the outcome of a production process of different public services pursued by public agencies, to fulfil the collective goals approved by the citizens in democratic processes. An outcome is generated when an output (in our example, public services) is consumed and completes its higher purpose (Panagiotopoulos, Klievink, & Cordella, 2019). By extension, outcomes can relate to other types of value. As mentioned earlier (Attard et al., 2015), the expected outcomes of open data can be economic value (e.g., paid applications can create new jobs and revenues), social value (e.g., the citizens' well-being is improved because they can better plan their mobility with an open data mobile application), or public value (e.g., public services and processes are more efficient with the knowledge gained from reusing data) (Berends et al., 2017).

In Figure 2, I depict a simplistic open data value creation process (a value chain) articulated around the generic process components: input, output and examples of expected outcomes, using a swim lane diagram. It allows us to see how the roles are involved and their main activities, and how along the process, the value created benefits different types of actors. The raw open data is valuable for the actors that can reuse it, taking the role of re-user. The cleansed and re-shared open data creates value for the re-user and its peers because they are the ones capable to use and process it. The open data transformed into information or

solutions, when shared, can create value for end-users of any sector, because they can make better decisions or improve their work by using the solutions, for example (e.g., citizens, civil servants, employees from private companies; Lassinantti, 2019).

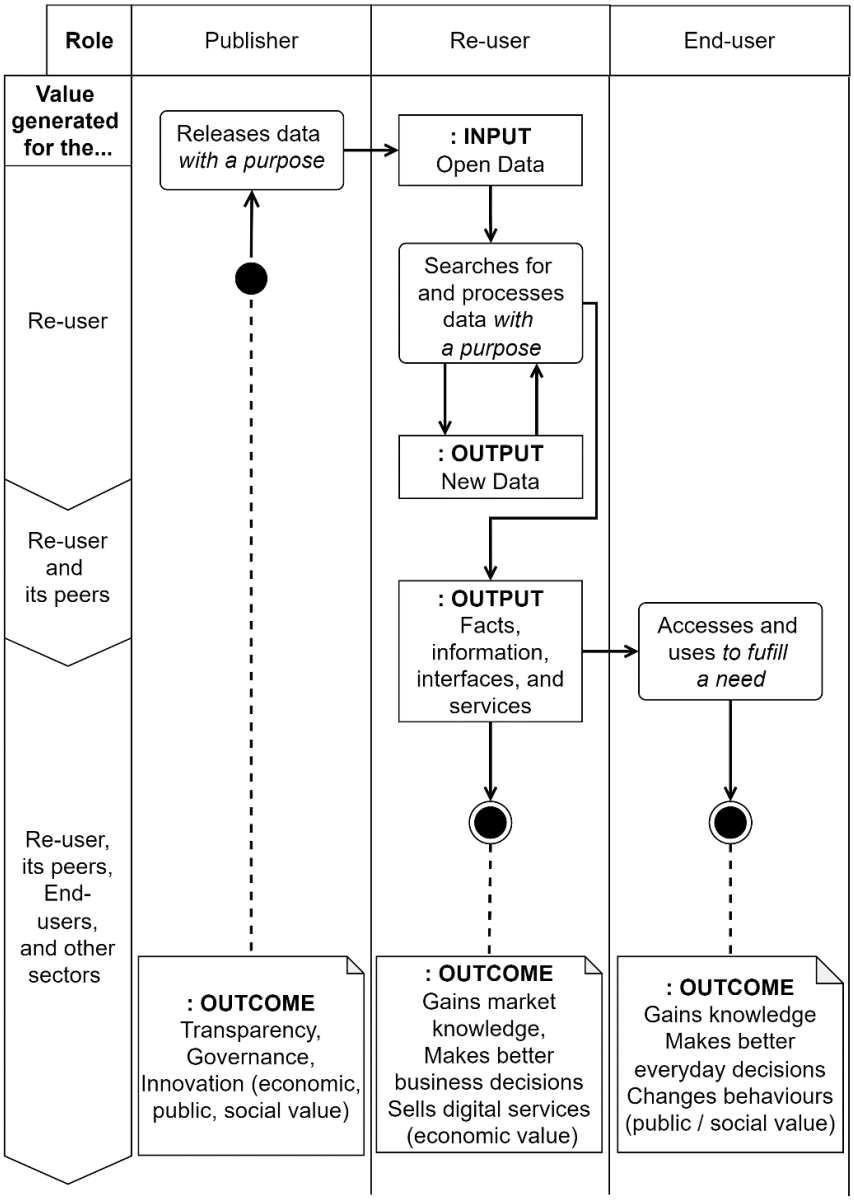


*Figure 2. Roles and activities in a simplified open data value chain*

I dug into the scientific literature, attended open data grassroots conferences, and interviewed a couple of practitioners. After a while, I saw more than one issue in the process and articulation of actors' contributions towards the realisation of the expected outcomes. In fact, given the intrinsic characteristics of the roles (see Figure 1) and the ideals of open data (for anyone, any

purpose), the roles and tasks of the actors are not as neatly distributed as suggested in Figure 2. Previous research has already developed several elaborated categories of roles and tasks (e.g., Attard, Orlandi, & Auer, 2016; Berends et al., 2017; Crusoe & Ahlin, 2019). However, they have a strong focus on data-related activities. They neglect the existence and possible influence of actors' other roles and rationales, which is relevant when open data is studied and used in contexts that imply other roles and multiple stakeholders, like in cities. It is known that publishers or re-users with a public mission define their goals and open data's expected outcome through the lens of the common good (Lupi, Antonini, De Liddo, & Motta, 2020) and that re-users from the private sector use the lens of profit (Zuiderwijk, Janssen, Poulis, & Van De Kaa, 2015). In previous literature, citizens, for their part, are categorised under the roles of developers and providers of solutions, or end-users of open data applications and services, possibly with a civic purpose (e.g., voicing their opinion regarding an urban problem through an application) (Desouza & Bhagwatwar, 2012). Under a label, a role can relate to different tasks for the researchers and practitioners using it.

I argue that the lack of understanding of open data roles can create confusion and problems in the implementation of the value chain and the realisation of the expected outcomes. I call these confusions and problems "misalignments". A **misalignment** is "*a situation in which parts of a system are different, so the system does not work well"*(Cambridge Dictionary, n.d.-b). Misalignments are studied in several research fields. For example, research in management and strategy uses misalignments to conceptualise flaws between actors' logics of action within an organisation (Bacharach, Bamberger, & Sonnenstuhl, 2018), between a firm strategy and internal or external variables (Drori, Sheaffer, Weber, & Landau, 2009), or between suppliers and customers' perception of needs and solutions (Corsaro & Snehota, 2011). Research in information systems investigates misalignments between software or information systems and organisations' structures and strategies (Heath, Singh, & Shephard, 2013; Wei, Wang, & Ju, 2005). Misalignments reflect an understanding of a gap between an effective (and ideal) model or process and its realisation in a complex reality, which does not result in an orderly and proper implementation (Drori et al., 2009). Misalignments are perceived as pitfalls to be addressed to complete successfully a process (Pereira de Castro Casa Nova, Costa Lourenço, & Ferreira LeitãoAzevedo, 2018) or the implementation of a project (Wei et al., 2005). Sources of misalignments have to be identified to solve them (Baker & Singh, 2019; Soh & Sia, 2004).

In the context of an open data value chain, I define **misalignments** as **dysfunctions and unfitted arrangements between the components of the process, or the actors in their respective roles, tasks, and purposes, that can cause problems and hinder the realisation of their expected outcomes**.

Previous research in open data occasionally refers to misalignments at different levels. I have found six studies mentioning misalignments in open data: misalignments between the priorities of stakeholders (Almirall, Lee, & Majchrzak, 2014), between the ambitions of the expected outcomes and the input provided (open data sets and portals quality) (Elbadawi, 2012), between the publisher's goal and re-users' expectations (Park & Gil-Garcia, 2021), and between the input provided and re-users' information needs (Lupi et al., 2020; Neves, de Castro Neto, & Aparicio, 2020). However, misalignments have not been identified in a systematic way. The five studies mentioned above come up with misalignments in their findings and discussion. To my knowledge, only one study introduced a misalignment as an initial problem to solve (Toro, Carrion, Albertella, & Brovelli, 2019). Toro et al. (2019) argue that different publishing practices

led to misalignments, lack of reuse, and repetition of information, which they suggest solving with standardisation.

**The research problem** of my thesis is **the possible misalignments** related to open data roles and their contribution to the value chain, in the release and reuse of open data in cities.

My reasoning, further developed in the next subsections, is the following. I see and show that these roles are not clear neither for research nor for practitioners. It made me suspect possible misalignments in value creation because if the actors involved do not agree on who does what and how, the realisation of the expected outcome is likely compromised. Previous research mentions, indeed, the current lack of value and timid outcomes of open data. Hence, I conceptualise possible misalignments and pose research questions, centred on the roles, assuming that understanding an issue is a first step towards the development of solutions.

## 1.3. Possible misalignments in the open data value chain in cities

To show the relevance of the research problem and introduce the possible misalignments related to roles at the core of this thesis, I use previous research and a few thought-provoking quotes from my empirical material as illustrations. The latter come from practitioners' websites used as cases in my research and a seminar about open data governance in the local public administrations in Wallonia I attended in June 2021[1] (organised by Futurocité, an organisation active in stimulating the open data community in Wallonia).

I focus on the roles in the shadow of the "publisher – re-user" duo, in cities: the ambiguous role of the citizens, the blurry scope of municipalities' roles and their expected users, and the distant end-user.

#### *The ambiguous meaning of citizen participation in open data*

Citizen participation can have different meanings, according to the rationales and background of the actors encouraging it. In cities, the citizen can be understood either as a political actor who has a voice in a democratic process or as a consumer who can be involved in a facilitated innovation process (Anttiroiko, 2016), inspired by user innovation (von Hippel, 1986) and open innovation (Chesbrough, 2003). Part of the open data literature presents citizen participation and engagement as an objective of open data, leading to outcomes such as transparency, accountability, and empowerment (Charalabidis et al., 2018a). However, Safarov, Meijer, & Grimmelikhuijsen (2017) show that this mere objective is lacking empirical evidence and creates vagueness about the citizen´s role, which is comprised of a broad and undefined group of either indirect beneficiaries of open data or direct re-users.

In practice, increasing citizen participation is an incentive for municipalities to release data. One speaker at the seminar, an employee of the municipality of Mons developing a first open data initiative, declared:

*"[Open data] would also allow us to ease the development of projects aiming at citizen participation which is very trendy at the moment."*

[1] The quotes are extracted from the recordings of the sessions, available online (Futurocité, 2021), and translated into English by the author.

A second speaker, an employee of the municipality of La Louvière, working on the municipality's open data strategy explained:

*"Another repercussion of open data is citizen participation (…). We want to give them the opportunity to collect and improve this data for the common good, (…) participate in other projects such as the measurement of air quality in the schools' surroundings. We also think that open data can encourage a change in behaviour from the citizens' side, (…) for example regarding soft mobility."*

He understands citizen participation as participation in projects beneficial to society through data collection and changes in citizens' behaviour.

One last example is from a private actor, the start-up *Wallonie en Poche* which provides an open data application for the citizens. This company sees the citizens as users with needs. In 2019, they launched a marketing campaign called "*let's make room for your ideas"* where the citizens, users of their solution, were invited to participate in the development of the application by sharing ideas and voting on development propositions.

The diversity of stakeholders, who can have diverging perspectives on open data (Gonzalez-Zapata & Heeks, 2015), generates a diversity of citizen participation forms. It made me suspect **a possible misalignment between the different actors' perceptions of other roles**, specifically, the citizens' roles (in terms of definition, scope, and purpose), **and between the municipalities' expected and realised outcomes,** the outcome from reusing data and involving citizens.

In open data initiatives addressed to or involving the citizens, little is known about the variety of citizens' roles comprised under the popular concept of "citizen participation". It can increase the risk of misalignment between the involved actors in terms of citizens' contributions to the value chain and realisation of the expected outcomes, and could explain the limited impact on the end-users and society (Zuiderwijk & Janssen, 2014).

*The blurry scope of municipalities' roles*

The lack of clarity in the role definition concerns also the municipalities, which can take several roles: public actor, publisher of open data, and in some cases, re-user of their data (Mergel, Kleibrink, & Sörvik, 2018). In open data, all roles are interchangeable: a publisher can be a citizen or a municipality; a re-user can be a municipality or an entrepreneur (Attard et al., 2016b).

In practice, we see municipalities combining the roles of the publisher and re-user. For example, the employee of the municipality of La Louvière introduced the concept of "internal open data".

*"The trigger of all our reflections is the will to create an internal open data with the aim to improve the efficiency of the municipal services."*

He meant that their open data strategy would start by sharing data within the organisation and its departments to fulfil the municipality's objectives: improvement of the municipal processes, better services to the citizens, and quality of life in their territory. In La Louvière, the municipal employees are the first intended re-users. The open data manager of Namur, a city having a long experience with open data, explained that as well. At the time of the seminar, they had a portal with 317 datasets "*allowing [anyone] to create innovative connected applications*" (Ville de Namur, n.d.-b). In fact, 150 datasets are closed and only used by civil servants. The municipality is a re-user for its processes and projects, and a re-user (intermediary) for the

citizens. They have developed ready-to-use applications for the citizens, such as one to inform citizens about ongoing and planned works in the public space.

The diversity in their approaches to open data can mean a diversity of ways of taking their role as publisher and/or re-user and can imply a diversity of expected users. I suspect again **a possible misalignment between the different actors' perception of other roles**, specifically, the roles of the municipalities and the role of their expected users (in terms of definition, scope, and purpose). Consequently, there is **a possible misalignment between the roles in the process**: what should be the activities of a publisher and a re-user in a value chain. Different types of re-users have different types of concerns and purposes for open data (Lassinantti, Ståhlbröst, & Runardotter, 2018). If the publishers (municipalities) have also different purposes, or do not understand those of their re-users, it could explain why we still do not know precisely what needs to be done to guarantee the achievement of transparency, participation, efficiency and innovation (Jetzek, Avital, & Bjørn-Andersen, 2014).

The municipalities' combination of roles is currently under-researched (Mesquita, Luciano, Rafael, & Wiedenhöft, 2020). Little is known about the scope and limits of the roles of municipalities and the one of the actors whose municipalities' input is addressed to.

***The distance between us: unknown re-users, imperceptible end-users***

The common definition of open data, "open for any re-use, without any restriction", discourages the publishers to implement registration and usage reporting systems perceived as a barrier to reuse (Janssen, Charalabidis, & Zuiderwijk, 2012) and, as a result, prevents them to know the actual re-users and reuses of their data, and by extension, the end-users (Susha, Grönlund, & Janssen, 2015). Without direct contact with the re-users, one possible publishing strategy is the "supply-driven" one, consisting in publishing what is available and believed to be important (Ham, Koo, & Lee, 2019; Susha et al., 2015). For example, the open data manager of Namur explained:

> *"Every time we have new datasets available for the citizens, we upload them on the platform, which allows us to accumulate. The aim is not to accumulate [open data], but to prepare the data and make it ready for usage by people. We don't know yet what the users will do with it, but at least there are available."*

A way to estimate the reuse of open data is then counting the number of downloads, which can generate an inflated perception of usage (Neves et al., 2020). The open data manager of Namur continued:

> *"We notice the interest for open data from the number of users per month. We have got already more than 4000 users per month. Well, 4000 uses, they might not be different users, but 4000 persons per month, either civil servants, colleagues in our department, or citizens"*.

Neves et al. (2020), based on previous research (Liu, Jiang, & Li, 2017; Montano et al., 2014; Ruijer et al., 2017; Walker, Frank, & Thompson, 2015), argue that this indicator does not show how the data is being used and by whom, and if the data quality is meeting re-users' needs. Following that argument, Lupi et al. (2020) state that the design of local open data portals is centred on what is known, i.e. the needs and purposes of the publishers, instead of being centred on the *information needs* of the re-users. Let alone the information needs of end-users.

The end-users, not in direct contact with the data, have their needs understood through intermediaries: local actors who use data for local actions and the common good (Lupi et al., 2020), or organisations that develop applications, platforms, and services to ease the use of data by others (Janssen & Zuiderwijk, 2014). Gonzalez-Zapata & Heeks (2015, p. 449) argue that the citizens who are not direct re-users are "*silent stakeholders*" because of "*lacking either direct "voice" that inputs to OGD or direct "hands" that make use of OGD outputs*". This absence of end-user involvement is surprising, given that participation is viewed as a primary objective in open data (Charalabidis et al., 2018a) and understanding end-users' needs by giving them an active role in the elicitation of these needs is widely recognised as an essential activity in software engineering. It has a long history in the field of Information Systems (Wing, Andrew, & Petkov, 2017), Human-Computer Interaction (Karat & Karat, 2003), and service design (Sanders, 2002), among others.

As noted by Lupi et al. (2020) and Neves et al. (2020), there is a **possible misalignment between the input provided by the publishers, i.e. the data, and re-users' requirements** to develop defined solutions. Knowing this, and the rare involvement of citizens without data literacy, I suspect that it can lead to a **possible misalignment between the re-users' output,** the solutions resulting from the data available, **and the end-user's needs**. It could explain why open data is criticized for its low impact on society (Zuiderwijk & Janssen, 2014).

Because, data is information to be (Hey, 2004), I argue that open data reuses must satisfy **information needs**. Little is known about the information needs of end-users from an open data perspective, and the capture of this type of need is tricky because they are personal, contextual, temporary (T. D. Wilson, 1981) and hard to express (Westbrook, 1993).

*Open data's timid outcomes*

Finally, previous research points out that the expected outcomes of open data are not fully realised (Zuiderwijk, Shinde, & Janssen, 2019), open data is not reused as much as expected, and open data initiatives have a low and unpredictable impact on society (Worthy, 2015).

Cautious municipalities can be discouraged to start allocating resources to open data initiatives (Zuiderwijk, Volten, Kroesen, & Gill, 2018) because of the lack of open data reuse and real impact on society (Worthy, 2015). On the other hand, they are cradled by rhetoric based on expected outcomes that should come to reality *when and only if* open data is published. Advocating municipalities spread the word (Zuiderwijk et al., 2018). For example, the open data manager of Namur finished his presentation at the open data seminar by stating:

> *"You have to start [releasing open data] before you need it. (…) We have a lot of datasets in Namur (…) we didn't know we could use them before we published them, but we need them (…). You shouldn't wait to need open data to publish open data. It will be, very often, useful to many people."*

I understand that, for some practitioners, open data is almost a question of faith: people need to trust and believe that it *will* bring outcomes. At the same time, the same practitioner (see earlier quotations from the open data manager of Namur) and previous research (Janssen et al., 2012; Ahlin & Crusoe, 2022) admit that the outcomes are unpredictable and uncertain.

As a result, convinced municipalities tend to mimic each other's initiatives (Zuiderwijk et al., 2019). The employee of the municipality of Mons, for example, started and concluded his presentation with:

*"Following the path of the municipalities of Namur and Liège, our municipality wishes to go into open data". (…) "Our objective is really to do what other municipalities have already done, that's why we have got in touch with them to try to reach the same situation [i.e. outcomes] they have achieved."*

**Encouraging critical thinking** against the rhetoric and practices of open data is, in my opinion, **a prerequisite to exploring new paths to improve the realisation of the expected outcomes**. I believe the usage of roles in combination with a value chain perspective will allow me this opportunity.

## 1.4. Theoretical and practical problems

In sum, I suspect four possible misalignments in the value creation process that could partly explain open data's timid outcomes. I visualise them in an open data value chain (further developed in Section 2.3.), at three levels:

- At the level of actors in their roles: a possible misalignment between the actor's perception of other roles (in terms of definition, scope, and purpose).
- At the level of actors towards an outcome: a possible misalignment between the municipalities' expected and realised outcomes.
- At the level of roles in the value chain: a possible misalignment between the publisher's input and the re-user's requirements and purposes, and between the publisher's input, the re-user's output and the end-user's needs.

The ***theoretical problem*** I focus on in this thesis is the lack of understanding of roles in open data initiatives in the context of cities, specifically the roles in the shadow of the publishers and re-users: the citizens, the municipalities and their expected users, and the end-users.

I argue that roles are worth being explored to understand possible misalignments of the value chain and find new ways to solve them, to potentially unlock more value with open data for the citizens and society.

The ***practical problem*** at the core of my thesis is the lack of understanding of municipalities regarding the needs and the scope of the roles, tasks, and contributions of the different stakeholders (e.g., citizens, developers), including themselves, to reach their expected outcomes. This lack of understanding can lead municipalities to avoid open data, publish data based on their priorities instead of the re-users' and end-users' needs, or mimic each other's initiatives with the belief that outcomes will come (Zuiderwijk et al., 2019).

I argue that critical thinking can be enriched by the elicitation of roles. It should help to bring light on discrepancies between what is done and what should be done to succeed in a given outcome.

## 1.5. Research purpose and questions

My ***thesis purpose*** is to explore **how the concept of role can bring new insights to capture the possible misalignments in the open data value chain and envision new solutions**. I assume that a value chain whose actors, roles, and tasks are aligned can achieve its outcomes, and that understanding misalignment is the first step to solving them.

Thereby, the produced knowledge and recommendations should help actors reflect on their own roles and alignment with others' roles and actions, to improve the realisation of their expected outcomes. The research output intends to provide actionable knowledge, relevant to practice.

To reach this purpose, I have divided my research project into three studies, each one focusing on one or two actors, roles, and possible misalignments.

The first study explores the role of citizens in open data-driven initiatives according to their innovation approach, assuming that the rationales behind the innovation approaches shape the way the initiative´s leaders define participation. It was guided by the following research questions:

- *What are the innovation approaches taken by participatory data-driven initiatives in urban settings?*
- *What can be the citizens' roles and patterns per innovation approach?*

The second study classifies the roles of municipalities, based on their goals for open data, tasks, and output delivered to external users. In addition, the study explores the roles expected from external users. The research questions are:

- *What are the possible municipalities' roles within OGD release and reuse?*
- *What expected users' roles are implied by the municipalities' roles?*

The third study focuses on the role of users of finished solutions, i.e., the end-users, who are not in direct contact with the data. The end-users tend to have a passive role and to be left aside in the ideation stages of the solution development. I argue that the end-users do not need data, but information. I explore ways to involve them in the expression of their information needs to help publishers and re-users align their work to the needs of their final consumers. I posed the following research question:

- *What design principles could data providers and intermediaries follow to design methods to identify the information needs of groups of users?*

Altogether, these parts should provide insights into how misalignments occur and how the concept of role contributes to our understanding of them. The comprehensive (yet not exhaustive) set of insights into the misalignments should open new perspectives of solutions and research to improve the value gained from publishing and reusing open data.

## 1.6. Knowledge contribution and audience

The three studies and the thesis are addressed to two types of audiences: researchers and practitioners. Their focus and complementary expected contributions are summarised in Table 1.

*Table 1. Studies and Thesis' focus, expected contributions, and audience*

| | Focus: roles and actors | Expected contribution (research output) | Audiences |
|---|---|---|---|
| **Study 1** | Roles of the "citizens" as actors/third parties in the open data value chain.<br><br>The citizens as *participants* in open data initiatives. | Clarification of the citizens' roles, in relation to the initiative leaders' rationales and open data value chain (process).<br>*Misalignments: actors in roles – actors' expected outcome* | **Research:** management of innovation community, use of new technologies in cities.<br>**Practice:** publishers and re-users of open (government) data interested in involving citizens. |
| **Study 2** | Roles of the municipalities in open data, and related expected roles of their output's "users".<br><br>The users' or end-users' roles *as expected* by the municipalities. | Clarification of the municipalities' roles in open data and expectations they project on the intended users.<br>*Misalignments: actors in roles – roles in the process* | **Research:** Open/e-Government and open data communities.<br>**Practice:** municipalities publishing or eager to publish open government data. |
| **Study 3** | Implication of the end-user, in a defined role, to inform publishers and/or re-users about their needs.<br><br>Ways to *engage with* the end-users and make them express their needs. | Guidelines (design principles) to develop flexible methods to capture end-users' information needs and inform the work of publishers and re-users.<br>*Misalignment: roles in the process* | **Research:** data science community, data use in a most effective, efficient, ethical way in promoting knowledge.<br>**Practice:** publishers and re-users of open government data, end-users or their representatives. |
| **(Compilation) Thesis** | Roles and value chain behind the dominant duo publisher – re-user, in cities: municipalities, citizens and re-users. | Roles as a catalyser to capture possible misalignments in open data value chains and envision new solutions | **Research:** participatory innovation management, Open/e-Government<br>**Practice:** managers of open data initiatives, actors involved in the stimulation of open data release and reuse (public and private actors) |

## 1.7. Delimitations

The present thesis focuses on the **open data roles in the context of cities**. I understand a city as a place where people live, work, and travel regardless of its size (in terms of population or surface area) or type (rural or urban), and which is governed by a local government called a municipality (an actor). The context of cities is present across the three studies. In the first study, it allows me to scope and select the cases: initiatives (solutions or services) for and with the involvement of citizens, who are the inhabitants and users of cities. In the second study, I focus on municipalities releasing and reusing open data. In the third study, the city is used to create scenarios that circumscribe the reflection of the participants in a given role. The choice of the city as a context is justified by the objectives of the research portfolio Wal-e-cities, funding this research. It is also justified by the observation that in practice, open data is

promoted by the public actors, governments, and civil society associations, which especially encourage the municipalities to engage in open data. The municipalities are sitting on plenty of data and the national laws related to public information access and open data release are also enforced at the local level. At the same time, new governance paradigms and leading visions for the future of cities, such as the open government and the smart city concepts, are encompassing the use of new technologies, information, and open data for innovation and transparency. In cities, open data is in the air, while local public actors are either mesmerized and advocate for open data or are cautious to reluctant to open data (Zuiderwijk et al., 2018). However, municipalities, as publishers, play a decisive role in the value chain, as one of the first links, and cities are an ideal testbed for participation and collaborative innovation.

Value creation with a value chain perspective is the second important delimitation of my thesis. The focus on value is determined by the objectives of the ERDF project (Wal-e-cities) and a research gap: open data does not fulfil all its promises of value creation. It also delimits my work around research gaps regarding the problem of value creation for a broader public and society, focusing on municipalities, citizens and end-users' roles in relation to others (publishers, re-users). We know that open data can generate value for its direct re-users (Lassinantti, 2019) and publishers/re-users (Mergel et al., 2018). The attention given to public value and impact on society (e.g., solving societal problems with open data) is however more subject to discussion and criticism (Zuiderwijk & Janssen, 2014). Open data value creation can be studied, among others, in terms of life-cycle (Attard et al., 2016b), ecosystems (Abella, Ortiz-de-Urbina-Criado, & De-Pablos-Heredero, 2017), or value-generating mechanisms (Jetzek, Avital, & Bjørn-Andersen, 2013). I chose a process perspective presented in a value chain because it allows putting the actors, their roles, and activities in the foreground, and it has good explanatory power. I want to find a balance between the obvious importance of the technical features of open data to create value, without side-lining the human factor and a holistic understanding of open data. I use a general process view depicted in a value chain that includes the notions of input (what is put into the system), output (what is produced) and outcome (the result or effect of the output). I do not enter the debate regarding the nuances between outcome and impact, as I use the outcome as part of the context and motivations of the actors, not as a research focus. I identified one possible misalignment in relation to the expected outcomes, but I do not write a precise research question about outcomes. The studies do not evaluate nor measure the outcome but understand the expected outcome as an element of the process, a higher purpose for the actors involved, and an element for discussion.

Finally, my objective is to produce knowledge that is relevant to practice and can become actionable, which has framed my research questions, methods, and research outputs. I further explain my paradigm and its influence on my research in Section 3.1.3.

## 1.8. Thesis outline

This thesis is structured as follows.

Chapter 2 starts with the development of the context of the emergence of open data, at the crossroads between the influences of innovation and open government paradigms (Section 2.1.). It provides a bigger picture of the paradigms underlying the rise of open data and explains the multiple perspectives and rationales of the involved actors. Section 2.2. continues with a presentation of open data, structured around the basic components of a value chain: the input, the actors and roles involved in the process, the output, and the outcome. Finally, I present a framework that combines all the relevant concepts and the research questions. To close the

chapter in Section 2.3., I state four possible misalignments, some of their sources, and potential consequences on the realisation of the expected outcomes of open data. They are addressed by the research questions posed in Section 2.4.

Chapter 3 focuses on the research design: the research paradigm, methods, and empirical contexts used throughout the thesis. It is structured logically and sequentially. It starts in Section 3.1. with the axiology, ontology and epistemology. I present my background and assumptions, and the research paradigm wherein I identify myself: the constructivist pragmatism. I map the influence of its main assumptions on the research design. In Section 3.2., I give an overview of the research. I expose the two research approaches followed in the thesis: qualitative abduction and design science research. I explain my research process and present the empirical context. I give in Section 3.3. a summary of the methodological alternatives, the chosen research strategies, data collection and data analysis techniques, and contribution to the main research problem. I conclude with the methodological limitations in Section 3.4.

Chapter 4 answers the first research question, centred on the role of the citizens in the value creation process. The study explains the innovation approaches influencing the citizen participation activities in open data initiatives in cities, identifies the citizens' roles, and highlights the patterns per innovation approaches.

Chapter 5 answers the second research question. It explores and classifies in a typology the multiple roles of the municipalities, as publishers and re-users of open data. The typology also allows us to set in parallel the expected roles of the "users", depending on the nature of the output provided by the municipality.

Chapter 6 answers the third and last research question by developing and testing methods to capture the information needs of the end-users with their involvement. The focus is thus the understanding of end-users' needs, who are at the bottom of the open data value chain, as a possible way to align the publishers' and re-users' work on them.

Chapter 7 discusses what we have learned about the roles and what insights they bring about the misalignments, opening new perspectives to solve them. I take a reflective outlook on the pragmatic paradigm and the methods used.

Finally, in Chapter 8, I conclude by summarising the key learnings and answers to the research purposes and questions. I finish with the limitations and possible development of the present contributions in future research.

Figure 3 shows the articulation of the chapters in the thesis.

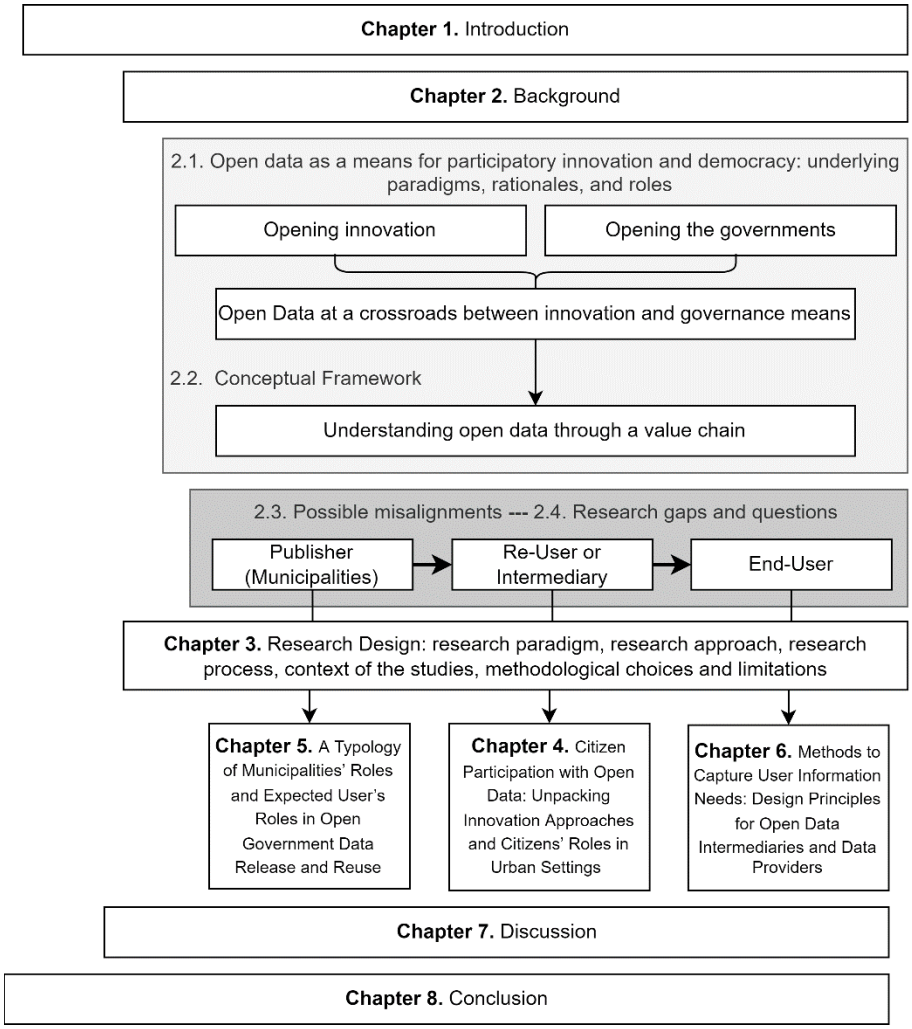


*Figure 3. Outline of the thesis and articulation of the chapters*

# Chapter 2. Background

## 2.1. Open data as a means for participatory innovation and democracy: underlying paradigms, rationales, and roles

Let us take a step back from open data to look at the larger context of two of its purposes: (1) innovation and (2) governance with transparency and participation (Charalabidis et al., 2018a). It helps us to understand the ambiguities of the roles and the sources of possible misalignments.

The next sections elaborate on the concepts of innovation and governance, with a focus on the trends and paradigms where the increasing interest for openness and involvement of third parties originates.

In Section 2.1., I briefly develop the roles of those third parties, being called users, re-users, end-users or citizens depending on the context, and the types of methods or interactions that can exist between them and the actors involving them (private companies, designers, public actors who might encourage the third parties' involvement). This background information leads us to some of the reasons for the rise of open data, pushed by top-down and bottom-up actors, as well as scholars, and its importance for municipalities.

Subsequently, in Section 2.2., I briefly present open data as an object (what open data is, the input), a process to create value with data, the main roles involved, and the possible outputs. The purpose of this section is to introduce open data to readers not familiar with it. It scrapes on the surface of a complex subject and explains the basics to understand the possible misalignments I conceptualise.

Finally, the reader has all the fundamental concepts I used to understand my perspective on the issues at the core of this thesis. In Section 2.3., I develop possible misalignments in the value chain. Thereafter, the research gaps and research questions are detailed in Section 2.4.

Figure 4 illustrates the structure of this chapter.

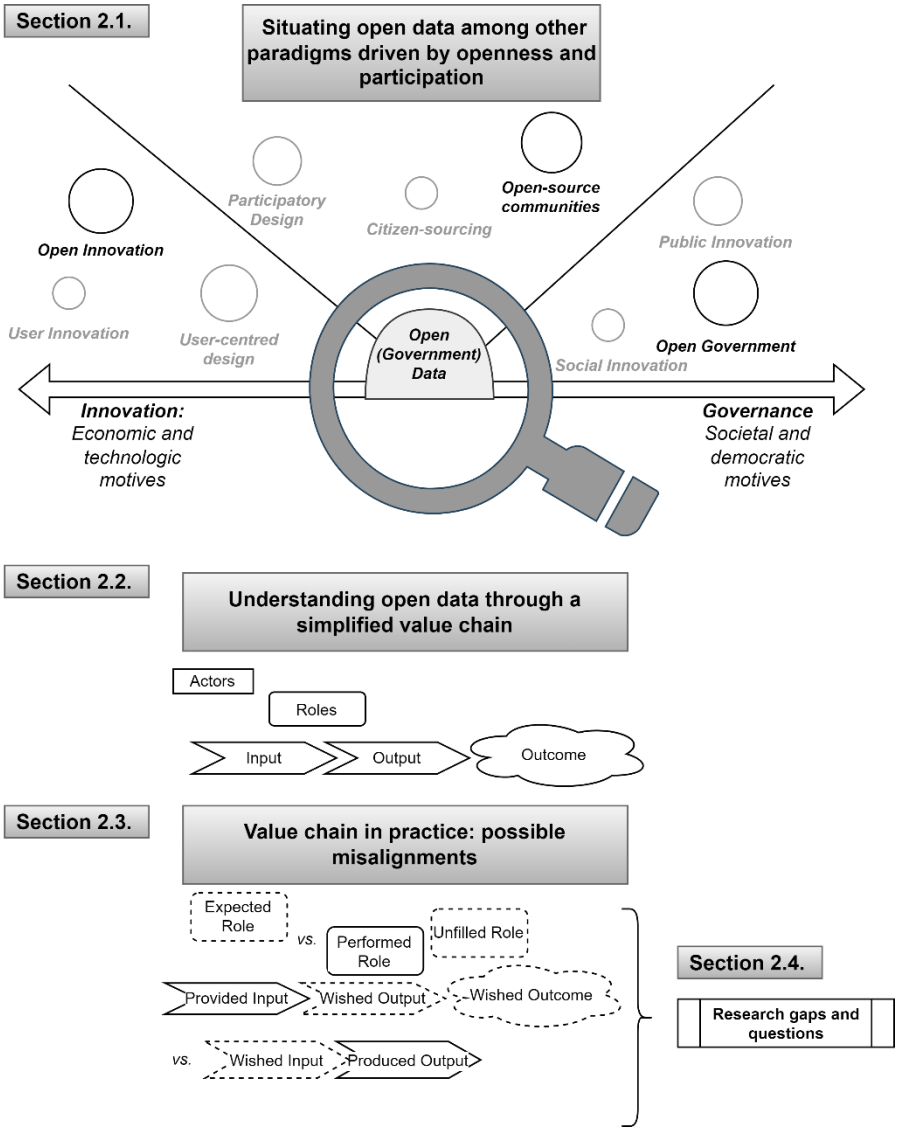


*Figure 4. Structure of the background chapter*

### 2.1.1. Opening innovation

This section presents the various meanings and forms of innovation, which can have purposes and outcomes bearing different values (economic, social, and public). However, one common trend is the increased collaboration and involvement of third parties. I introduce the main innovation approaches underlying this involvement with their rationales, the roles of the third parties (specifically, the users of the innovation output, a service or product), and an overview of the methods commonly used. These roles and methods are important to understand because they influence open data practices.

*Definitions and forms of innovation*

Innovation, from an etymological point of view, comes from Late Latin *innovatus*, the past participle of *innovare*, "*to renew, restore*". It is a compound word from *in-* which means "*into*" and *novus*, which means "*new*". Lately in English, from the 1540s, innovation takes the meaning of "*a novel change, experimental variation, new thing introduced in an established arrangement*". From the 1590s, the verb *innovate* means "*bring in new things, alter established practices*" (Online Etymology Dictionary, n.d.).

Nowadays, the Cambridge Dictionary defines the word *innovate* as "to ***introduce*** *changes and* ***new*** *ideas; to* ***develop*** *a* ***new*** *design, product, idea.*"(Cambridge Dictionary, n.d.-a). In the business sector and for the European Commission, the definition used to measure innovation and develop policies is stated in the Oslo Manual, a joint publication of the OECD and Eurostat (*Oslo Manual, Guidelines for Collecting and Interpreting Innovation Data*, 2005). It claims that "*an innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method, or a new organisational method in business practices, workplace organisation or external relations.*"

Behind those commonly accepted definitions, there is the shared idea that innovation happens when *something new is brought* or *developed*. Innovation is about novelty and has become a pervasive term, used, and over-used by organisations in their visions, mission statements, or marketing campaigns. Even politicians and public administrations want to become more innovative and foster innovation in their territory. Like the aforementioned definitions, the scientific literature, as synthesised by Kahn (2018), makes a difference between innovation as an *outcome*, as a *process*, and additionally, as a *mindset*.

Innovation ***as an outcome*** includes product innovation (e.g., new products, services, programs by introducing improvements, line extensions, new uses, or "new-to-the-world" products), process innovation (e.g., changes in a methodology to achieve efficiency, cost reduction), marketing innovation (e.g., new ways to connect with consumers to achieve awareness, brand recognition), business model innovation (e.g., new value chain, revenue model), supply chain innovation (e.g., changes in supply chain technologies, processes) and organisational innovation (e.g., new structures, forms of management).

Innovation ***as a process*** is the way in which innovation is and should be organised to reach the desired outcome. It is divided into different phases, such as the discovering phase, the development phase and the delivery phase. Within that perspective, a significant research stream is the New Product Development (NPD) Process, which can include idea generation, scope and business cases, development, test and validation, and market launch (Cooper, 2018).

Innovation ***as a mindset***, finally, addresses the way innovation is internalised by individuals to support an innovative culture in an organisation. The focus is on the required soft skills and the attitudes that enhance creativity and innovation. For example, Design Thinking, a human-centred approach to innovation, is not only about methods and design processes, but also about embracing an empathetic mindset towards the needs of the users (Brown, 2009).

Table 2 summarises the three conceptualisations of innovation that dominate the business and management scientific literature.

*Table 2. The three conceptualisations of innovation (Kahn 2018, p. 459)*

| Conceptualisation | Strategic focus | Strategic question | Considerations |
|---|---|---|---|
| Innovation as an outcome | Ends | *What* do you want to happen? | - Product innovation<br>- Process innovation<br>- Marketing innovation<br>- Business model innovation<br>- Supply chain innovation<br>- Organisational innovation |
| Innovation as a process | Ways and Means | *How* will you make it happen? | - Innovation process<br>- Product development process |
| Innovation as a mindset | State | What should be instilled and ingrained to prepare for the *what* and the *how*? | - Individual mind-set<br>- Organisational culture |

Nevertheless, innovation is no longer limited to the private sector and a large variety of concepts of innovation is flooding research and practice (Edwards-Schachter, 2018). The change, a common characteristic of all innovation projects, may originate in the need to solve a problem or a societal need, that organisations, any for-profit or non-profit oriented entities (public, private, third sector, hybrid firms), need to overcome to survive in rough times. The change can be planned, provoked or unintentional, but tackling it aims at improving a situation (Edwards-Schachter, 2018). For example, innovation can come in response to grand challenges, such as the climate crisis (responsible innovation). Innovation can also contribute to a greater good to drive changes in social practices and solve social problems (social innovation) (Cajaiba-Santana, 2014). When it deals with public sector efficiency (e.g., innovation in public organisations, processes, services and governance), innovation is referred to as public innovation (De Vries, Bekkers, & Tummers, 2016). Sometimes characterised as "dark" or "hidden", innovation can come from the third sector, public institutions, multi-stakeholders and cross-sectorial collaborations (Edwards-Schachter, 2018).

However, there is a fundamental difference in the purpose of these innovations. Innovation in the public sector does not aim at creating *economic value* but *public value*, the value that results from the public sector's activities, which must serve society as a whole, and not special interests (Moore, 1995). Social innovation, often associated with grassroots innovation in the third sector, aims at creating social value. *Social value* is the value created when resources, inputs, processes or policies are combined to generate improvements in the lives of individuals or society as a whole (Emerson, Wachowicz, & Chun, 2000).

In sum, innovation encompasses several concepts and can be studied from the perspective of the process, the outcome (economic, social, or public), or the mindset. In my thesis, the timid outcomes of open data are the motivation to investigate the innovation process.

As innovation is spreading across sectors, it is inspired and motivated by different values and purposes. One common trend, however, is the **increased collaboration** in the innovation **processes**: innovation becomes **more open and participatory**, involving diverse actors and sources of knowledge. One stakeholder, especially, is the focus of this thesis: the consumer, customer, beneficiary, or more generally, "user" of the innovation output.

In the next subsections, I elaborate on the increasing interest for user involvement in innovation from the private sector to the public and third sectors, and the different rationales, methods, and users' roles standing behind the innovation paradigms. I introduce these paradigms and their views on the user role as I identify them as a possible source of misalignments in the open data value chain (see later 2.3.).

### *Participation in innovation: a brief overview of rationales and user roles in private, public, not-for-profit, and hybrid sectors*

#### Participation in innovation in the private sector

In the last century, the leading innovation paradigm and the practice assumed that innovation occurred inside the firms and was controlled and developed inside the R&D department before being launched on the market, in a technology-push approach.

New perspectives bumped into this traditional approach, especially with the involvement of the user. User participation is claimed to have numerous benefits. It helps the designer or developer to understand user requirements, access useful information, and new ideas, and define the scope of a project. It increases the user's satisfaction and acceptance of the designed product, by improving it all along the process. It has positive effects on the whole product or service success (Kujala, 2003).

Nowadays, a broad landscape of approaches and practices exists in several research streams. I focus on and introduce four approaches, according to the degree of user involvement (from the lowest to the highest): the user-centred design (Abras, Maloney-krichmar, and Preece, 2004), participatory design (Schuler & Namioka, 1993), user-driven innovation (von Hippel, 2005a), and open innovation (Chesbrough, 2011). Each approach can be differentiated by the emphasis on the methods used and the rationale behind them. They have evolved and influenced each other over time.

***User-centred design*** was first mentioned by Donald Norman from the University of California San Diego in the '80s, in the area of Human and Computer Interaction (Abras et al., 2004). The approach of development is centred on the user's needs, wishes and limitations, all along the design process. Despite the user's needs dominating the design process at all stages in an iterative process, **the user** was originally either a passive **object of study** or an **informant of its requirements,** the latter including more interactions with the researcher. The emphasis of the approach is on the usability of the product. The user can be involved all along the innovation process or observed, and analysed, with some degree of empathy (e.g., empathic design; Steen, Kuijt-Evers, & Klok, 2007, and design thinking; Brown, 2009). The researchers, designers or developers have the lead on the project and the responsibility for the result.

***Participatory design*** goes a step further in user involvement. The school emerged in the '70s in Scandinavian countries. At that time, trade unions saw the offices becoming automated by computers and they sought to increase the democratic values and the workers' voice in the workplace (Steen et al., 2007). The "democracy at work" started by involving the workers in the design of solutions and tools to improve their work experience, joining skilled workers together with designers in the search for solutions. It evolved into a set of methods and principles created to facilitate the integration and participation of the users in the innovation process of service design (Schuler & Namioka, 1993). In this approach, users are partners with the designers; their involvement, to different degrees, is considered central in the process. It is believed that the solution should serve the people it is designed for (fit their tasks and environment) and that people affected by a change should be allowed to influence it (Karlsson, Holgersson, Söderström, & Hedström, 2012). Therefore, democratic participation in the design process is emphasised. The political influence in participatory design progressively disappeared, and the second generation of participatory design came. Nowadays, it is still moved by a logic of empowerment and acknowledgement of each other's competencies and knowledge. The difference in knowledge and language between participants (users) and experts (designers, developers, and scientists) is a challenge. It has made the development of tools and techniques to reduce the cultural and knowledge gap and enable communication and collaboration between the parties, a full part of the design process. For instance, the co-creation and generative tools of Sanders & Stappers (2008) combine three perspectives, what people say, do, and make, to access users' feelings, tacit knowledge, and dreams. The users are partners and experts in their needs.

In ***user innovation*** (von Hippel, 1986), the idea is that relevant and useful information exists outside the firm, and cannot always be acquired by the traditional vertical integration strategies. It is the "sticky information" of the user (von Hippel, 1994). The author claims that "*users of products and services, both firms and individual consumers, are increasingly able to innovate for themselves*" (von Hippel, 2005b, p.1). As an expert on his environment and needs, the user is considered the more aware of the solution required to fulfil his needs. This type of innovative user is called the "lead user". Collaboration with lead users can help a firm to anticipate market needs and develop innovative user-oriented solutions. Innovation, leaving behind a paradigm of internally managed processes, is said to be "democratized" in the sense of being open to external innovators with the help of toolkits (von Hippel, 2005b). The user is responsible not only for identifying problems but also for developing solutions.

Similarly, in the early 2000s, Chesbrough (2003) introduced the ***Open Innovation*** paradigm. His main thesis is that inbound or outbound knowledge flows can benefit the firm and increase its profits. With the release of internal knowledge to external innovators or the importation of external knowledge inside internal processes, a firm can maximize innovation effectiveness. In that way, the firm can find new ways to commercialize externally developed solutions or reinvent its methods with fresh looks.

User innovation and open innovation can also be called ***distributive innovation*** (Bogers & West, 2012) since the innovation process and value creation are distributed between actors inside and outside the firm's borders. The rationale behind these approaches is that participation can increase the firm's profit and market shares, with a focus on the efficiency of the process in open innovation (innovation as a process) and the user value for the user innovation (innovation as an outcome). The user is an innovator.

However, it implies that the innovator has the skills and resources to work on innovative solutions. One of the major challenges for the firm is to get access to the right user or innovator. Meanwhile, the opportunities brought by the Internet and the easiness to work and communicate at distance created new opportunities. Innovation toolkits (Franke & von Hippel, 2003), open innovation contests (Terwiesch & Xu, 2008), and crowdsourcing (Piller, Ihl, and Vossen, 2010) are typical examples of methods and incentives used to raise interest and attract talents to solve specific market challenges. Companies organise idea competitions, where the community members can be idea generators, developers or designers, and users of the winning solution (Steen et al., 2007).

Nowadays, the different schools of thought influence each other (Sanders & Stappers, 2008). There is not a clear border between the different approaches and the combination of methods and practices from different schools of thought is common. They have all, to some degree, influenced the open data movement with the idea that the data should be available for innovation (see later 2.1.3.).

Table 3 presents a summary of the key characteristics of the abovementioned innovation approaches, originating in the private sector. It shows the multiple ways to involve the users of a service or product in its development. User-centred and participatory design are already popular and recommended for digital service innovation (Rose et al., 2019), which opens perspectives for an active role of the end-user in the open data value chain.

*Table 3. Summary of the characteristics of the participatory innovation approaches*

| Innovation Approach | Rationale | User's roles | Methods and techniques | Leading authors |
|---|---|---|---|---|
| **User-centred Design** | Needs and interests of the user<br><br>**Usability** of the thing being designed (design of products, technology fit with user's tasks and environment)<br><br>**For** users | User as a passive object of study (what can be learned about him/her)<br><br>A respondent, subject in lab experiments | **User analysis**<br>- Personas<br>**Task analysis**<br>- Scenarios<br>- Background Interviews and questionnaires<br>- Sequence of work interviews and questionnaires<br>- Focus groups<br>- On-site observations (ethnographic fieldwork, contextual design)<br>**Prototyping**<br>- Role-playing, walkthroughs, and simulations<br>- scenarios<br>**Evaluation**<br>- Usability testing with test and feedback, videotaping, think aloud<br>- interview and questionnaire | (Norman, 1986) |
| **Participatory Design** (open innovation) (Co-design) (Co-creation) | **Democratic participation**<br>Skill enhancement (design for a purpose)<br><br>Use of each other's knowledge<br><br>**With** users | User as an expert of his/her experience<br><br>Advisors, representatives of a group, partners | **Co-design Workshops**<br>**Prototyping**<br>- Mock-ups, paper-based outlines<br>Design of **mediation tools, techniques**<br>- "make tools" & "generative tools" as a design language | (Schuler & Namioka, 1993)<br>(Muller, M. J. & Kuhn, 1993)<br>(Sanders & Stappers, 2008) |
| **User Innovation**<br><br>**Open Innovation** | User innovation<br><br>**By** users | User as a capable innovator, idea generator, developer, designer | **Innovation toolkits**<br>**Crowdsourcing**<br>- User communities<br>- Innovation contests<br>- Innovation toolkits | (von Hippel, 2005a)<br>(Chesbrough, 2011)<br>(Piller et al., 2010) |

### Private-collective innovation: the hybrid model of open-source communities

Before continuing to the public and third sectors, I want to highlight a collaborative and innovative movement that is neither fully private nor really grassroots and had a significant influence on developers at the origin of the open data movement (further developed in Section 2.1.3.). The phenomenon of ***open-source communities*** applied the concept of openness and knowledge sharing long before the emergence of the lead-user and open innovation theories. Especially in the development of open software, open-source communities implemented peer production and developed community-driven innovation.

As reminded by von Hippel and Krogh (2003), to understand the roots of the open-source movement, we have to go back to the '60s and '70s, when most of the software development was carried out in corporation labs and universities by scientists and engineers. For them, it was normal and part of their work culture to exchange code with their peers, modify, and build upon each other's contributions individually or collaboratively. This behaviour became central in the "hacker culture". The scientists, engineers, and developers got access to new possibilities of collaboration with the creation of the first computer network, and later the Internet. In the '80s, when a commercial company, in accordance with its practices, privatized the work communally developed by developers and hackers, it hurt the values and morals of part of the community. In reaction, Richard Stallman, a programmer at the Massachusetts Institute of Technology founded the Free Software Foundation in 1985. He paved the way to open licences and the open-source movement.

In open-source, software are developed with the help of the user community. The users can have access to the code to get inspiration from the way the contributors implemented the solution, for example, the algorithms that were used (users as readers), report problems (bug reporters), or improve the system with variable levels of involvement (bug fixers, peripheral developers, active developers, core members of the project) (Nakakoji, Yamamoto, Nishinaka, Kishida, & Ye, 2002). Von Hippel and Krogh (2003) define it as a ***"Private-Collective" model of innovation***.

### Participation in innovation in the public and not-for-profit sectors

Finally, these approaches of innovation and new methods have spread across the public and not-for-profit sectors. The shift in the rationales and locus of the value, previously mostly captured by the firm, then extended to and shared with stakeholders or the community (e.g., the users), have blurred the borders between the actors creating and the actors capturing the value of innovation. It has accelerated the adoption of these innovation practices by the public and not-for-profit sectors in which they found some common ground and values.

For example, the concept of ***participatory innovation*** is an attempt to combine the user perspective with business orientation to maximize the value captured by all the parties involved (Buur & Matthews, 2008; Buur & Larsen, 2010). Participatory innovation can, in some cases, evoke a multiplier effect that enables each individual user-member to reap benefits that would have been unattainable without collective organisation (Borgen & Aarset, 2016).

In the public and third sectors, user involvement is used to solve socio-economic development and territorial challenges (Bhatt, Ahmad, & Roomi, 2016). Public and social issues are often considered wicked problems that involve multiple stakeholders, whose mutual understanding and collaboration can be the first step towards a solution. The seminal work of Brown & Katz

(2011) has introduced design thinking methods into the public (McGann, Blomkamp, & Lewis, 2018) and not-for-profit sectors (Brown & Wyatt, 2010). The use of participatory approaches combined with new technologies has become popular to try to solve "old" problems in a new way. For instance, in Living Labs, user-centred and participatory design are used to co-create technological innovations with the beneficiaries of the solutions and stimulate bottom-up innovation (Pallot et al., 2011; Dell'Era & Landoni, 2014). With this method, researchers, public actors, and citizens in Malmö, Sweden, tried to tackle immigration and social integration (Björgvinsson, Ehn, & Hillgren, 2010).

The public sector has also integrated the idea of involving citizens in innovation. Citizens' participation is more and more solicited to develop new public services, in what is called ***collaborative innovation*** (e.g. Bommert, 2010; Sørensen and Torfing, 2011). Crowdsourcing principles can also be applied in the public sector to be turned into ***citizen-sourcing***. Citizens are involved to give ideas, improve or develop policies, and co-produce services (Nam, 2012). Finally, in ***social innovation,*** the participation of the stakeholders is also a core component since the aim is to create value for the communities.

In conclusion, the idea of involving users, citizens or beneficiaries is commonly accepted across the sectors. Several methods and tools developed in the private sector, where innovation is vivid but motivated by profit, are imported and reinterpreted by the actors of the public and third sectors. In appearance, the practices can look similar, but they are driven by different purposes and values.

Similarly, the release of open data can be motivated by innovation or governance and transparency. The involvement of citizens, citizen participation, takes a completely different meaning when anchored in public and democratic values. The next section explores and develops the evolution of recent paradigms in public management going towards the increasing importance of citizens' involvement and openness. They are the other side of the coin that constitutes the foundations of open data.

### 2.1.2. Opening the governments: participation in democracy, transparency, and governance

As mentioned before, the aim of the public sector is to create public value. More specifically, a government can be defined as a mechanism for collective action. By making laws and policies, collecting taxes, and building institutions, administrations and public agencies, it manages problems that are too large for the individuals, and provides services that are too expensive to be attractive or profitable for the private sector but that are essential because they serve the common interest (O'Reilly, 2011).

Moreover, in democracies, public priorities and services are subject to the citizens' oversight. To serve the common interest, all public organisations deal with questions of efficiency, accountability, and equity. Several institutional paradigms co-exist to tackle them. They have different views about what is at stake and bring different answers on how to achieve these objectives (Stoker, 2006). Throughout history, dominant paradigms successively appeared and influenced the way governments conceive their role, their responsibilities, and the way to interact with the citizens. New paradigms appeared in reaction to perceived governance deficiencies, changes in values, and the evolution of society. Accordingly, the philosophy of governing, the citizen's roles, and the mode of interaction with the government evolve in a

spectrum (Vigoda, 2002), where open data can be used as a tool to achieve the governments' objectives and renew the interactions with the citizens.

I limit this brief overview to four paradigms that have a role in the increase of citizen participation and later the rise of open data in cities.

***Authoritarianism*** was the governance model that predominated for ages, until the late 19th century, and still today in dictatorial regimes. Bureaucratic structures hold high power, and the system is mostly non-democratic. The public sector monitors many aspects of the citizens' lives (education, socialization), dominates by force, and provides a limited number of public services. The citizens are dependent on the government's will and either accept to not speak their voice or, given a context of violence or pressure, for instance, do not dare to do otherwise.

The voting system is the main and first participatory mechanism that allowed the citizens to give input into public management. Representative democracies spread in western countries. At regular intervals of times, citizens can vote and elect their representatives who manage public affairs, with respect to the values and vision of their party. However, that model called "***old administration***" (Vigoda, 2002) showed its limits as there are few instruments or mechanisms to allow in-between citizens' influence. One renowned scholar who emphasized the problem is Arnstein (1969) and her famous model, the Ladder of Participation. The model attempts to increase the "true" citizen participation, initially in urban planning projects. However, over time, research showed that not all citizens are able or inclined to participate. A form of cynicism and critics against the politicians, their actions, and the representative democracies increased and is still ongoing today (Papadopoulos, 2013). The trust the citizens have in the institutions, political parties, and elected representatives is in crisis (Armingeon & Guthmann, 2014).

In reaction to the situation, scholars (e.g., in the seminal book "Reinventing Government" by D. Osborne and Gaebler, 1992) and citizens called for a reform of governments. The ***New Public Management*** was presented as an alternative to gain efficiency and increase citizens' satisfaction by importing and using business practices in public management. Within that New Governance Paradigm (S. P. Osborne, 2006) further elaborated as the New Public Service or Public Service Logic (S. P. Osborne, 2018), the creation of public value is believed to be no longer the prerogative of the administration. The idea is that public services can be created by any actor, as long as there is a focus on public value. This public value emerges via interactions between service users and service providers, not only between the citizens and governments (S. P. Osborne, Radnor, & Strokosch, 2016). The government can increase efficiency by opening public service provision to the private sector, not-for-profit actors, and citizens. The higher purpose is "to meet the needs" of the citizens (customer-oriented approach) and create a better quality of life for all (the public value of equity). In addition, it is believed that public services need to be reformed in a way that empowers individuals and communities to have greater choices over the packages of services they receive.

Despite these late developments toward more citizen participation, New Public Management was criticized for not encouraging enough collaboration or partnership with the citizens and causing instead more passivity (Vigoda & Golembiewski, 2001), and for challenging democratic principles and values (Farnham, David & Horton, 1995). In response, the ***Open Government Paradigm*** aims at improving the public value generative mechanisms of participation, collaboration and transparency (Harrison et al., 2012). The open government paradigm has

been defined as the extent to which citizens can monitor and influence government processes through access to government information and decision-making arenas, combining the availability of information (transparency) but also interactions between government and citizens (participation and collaboration) (Meijer et al., 2012). The higher principles of Open Government are citizen engagement, transparency, accountability, and integrity. To implement these principles, governments develop policies for more inclusive decision-making processes, access to information, and access to public services such as open online portals and service to citizens initiatives among others (O'Connor, Janenova, & Knox, 2019). New technologies are a means towards transparency, with the better quality of public services as a direct short-term impact. However, open government cannot be exclusively linked to service improvements as it includes also better training for the civil servants, better salaries to avoid corruption and customer orientation in the public management. In the long run, it aims at improving the quality of democracy, inclusive growth, trust in government, and rule of law (O'Connor et al., 2019).

Table 4 provides a summary of the paradigms, rationale (including the rationale for open data in the two last paradigms), and roles of the citizens.

*Table 4. Main public management paradigms with their corresponding rationales and citizens' roles (based on Vigoda, 2002)*

| Public management paradigm | Rationale | Citizens' role | Type of interaction |
|---|---|---|---|
| Authoritarianism | The leaders and administrations have absolute power and control over the citizens and public affairs. | Subject | Coerciveness |
| "Old" administration | Citizens' representatives, with their experience, wisdom, and civic values are entitled to manage public affairs. | Voter | Delegation |
| New Public Management | The citizens have needs (demand) and the government can better meet them (supply) by creating mechanisms and structures inspired by business management practices.<br>→ Open data for innovative and outsourced services, efficiency | Client or customer | Responsiveness to citizens' needs |
| Open Government | Public values generating mechanism (participation, collaboration and transparency)<br>→ Open data for transparency and citizen participation | Partner | Collaboration |

In conclusion, different paradigms co-exist and influence the way politicians define their public management style, their role, the way to interact with the citizens and make use of the opportunities offered by information and new technologies, including open data.

In innovation or governance, once a form of involvement or participation is expected, the means to enable that participation must be provided, with tools, methods, or mechanisms. Information and communication technologies offer new opportunities to interact, solve public challenges and improve the public actor's efficiency and transparency. Internet, digital

platforms, and toolkits also took a central role in the implementation of public or user innovation processes.

Open (Government) Data became one of these new means promising innovation, transparency, and participation. It is a new means to reach certain objectives of a government, and in that way, can be considered "innovation as a process" (Martin, 2014), specifically an open innovation process (Smith & Sandberg, 2018; Corrales-Garay, Mora-Valentín, & Ortiz-de-Urbina-Criado, 2019; Corrales-Garay, Ortiz-De-Urbina-Criado, & Mora-Valentin, 2020). The recent developments in public management paradigms and innovation practices created a fertile soil for its rise. I discuss them in the next subsection.

### 2.1.3. Open data at a crossroads between innovation and governance means

#### *The rise and importance of open data*

Open data increased in importance in the last decade, and more precisely from the year 2009 (Zuiderwijk, Helbig, Gil-García, & Janssen, 2014). A more recent state-of-the-art of open data research confirms this significant rise in research publications. According to the study of Zhang et al. (2017), which covers the 1998 to 2016 range, the number of research outputs about open data has increased from 94 papers during the period 1998–2008 (9%) to 951 during 2009–2016 (91%).

This time coincides with the famous Obama's Administration Memorandum on Transparency and Open Government, published in 2009 (The White House, 2009). One of the resulting directives was that governments should open their information, including data, to the public.

However, the federal administrations are not the only actors reinforcing the pressing call for open data release. I identified three important streams that have pushed and advocated for open data, from top-down, bottom-up, and academic voices. They are anchored in the trends introduced in the previous sections (especially, *open innovation*, *open-source communities*, and *open government* that I highlight in the next paragraphs). I briefly develop them, in connection with the participatory innovation approaches and public management paradigms, to point out the diversity of meanings and purposes open data can have for its stakeholders.

The institutional push (top-down): open data for transparency and economic development.

Access to public information has a long history. Traditionally, access was given upon request following the principles of the Freedom of Information Act in the United States, or its equivalent in other countries. The right to access information is related to the Human Right of freedom of speech, Article 19 (United Nations, 1948; Charalabidis et al., 2018a). To move that access a step further, reflections about facilitated and unlimited access to information increased in the public institutions and national governments.

In the United States, Obama's Memorandum of 2009 focused initially on the principles of *Open Government*, to create a more open, transparent and collaborative administration, and was accompanied by the launch of the national open data portal Data.gov (U.S. Federal Government, 2009) the same year. In 2013, an executive order specified that public information should be made available in a more raw and workable format, in other words, in a data format, and the data should be machine-readable by default (The White House, 2013). It is an important step towards the systematic release of open data. Since then, open data

initiatives have spread beyond the federal level. Open data was adopted by states and local administrations, also outside the United States.

In Europe, the reflections on the regulation and access to public information started in 2003 with the adoption of the Freedom of access to information Directive (European Parliament & European Council, 2003) regarding environmental information. Later, the European Commission specified its scope to the Public Sector Information (European Parliament & European Council, 2013), which aimed at making available information produced by the public sector at all levels (national, regional, local) and organisations partially funded and controlled by public authorities (e.g., public transport companies, weather agencies). The information comprises data, and the European Commission's perspective is clearly more economical than Obama's: data release is encouraged for *(open) innovation* and economic growth. The directive was further revised in 2019 to ensure better reusability of the open data and its application by the state members (European Parliament & European Council, 2019). It highlights transparency and fair competition, giving more details about data availability, formats, licences, and highly valuable data. To support the data release, a European open data portal was launched in 2012 (Publications Office of the European Union, 2012), and impact studies, national reports, and e-learning tools are published on a regular basis.

Other supra-national institutions, such as the OECD (Ubaldi, 2013) and the World Bank (The World Bank, n.d.), together with organisations including among others the Open Government Partnership (The Open Government Partnership, n.d.), claim the power of accessible data for transparent governments (*open government*) and economic development (*innovation*). They advocate for open data, giving more weight to the narrative surrounding open data and pressuring smaller institutions and administrations to follow.

In other words, leading institutions and governments encourage the release of reusable and machine-readable data instead of printed or scanned documents and archives, the "old way" of transparency. This information was historically accessible but in inflexible formats and upon request. Nowadays, the assumption is that raw data opens more possibilities for commercial reuse, transparency, and participation in governance. Depending on the political agendas and paradigms, data and ICTs are used by the governments to experiment, learn, and collaborate with the citizens in terms of public service provision (*New Public Management*) and governance (*Open Government*), or enable the citizens to monitor the public action and develop services on their own (*Open Government*) (Gray, 2014).

The bottom-up push: open data for open access, collaboration, and activism.

Open data is also clearly inspired by the *open-source philosophy* (Harrison et al., 2012). Open access and civic hacking communities started to use the term "open data" as a distinctive concept around 2006-2007. Before that, the interest in open data was especially important in geo and geospatial data (Gray, 2014). For example, Open Street Map was created by Steve Coast in 2004. Members of the open-source community increasingly collected and released data to produce free street maps of the world. Open Street Map is still a leading project that attracts many developers into the world of open data, largely influenced by the principles and practices of the open-source movement and using *crowdsourcing*.

Nowadays, the figurehead of the open movement is the Open Knowledge Foundation, established in 2004, which is also heavily influenced by the open-source movement, democratic values and civic activism. The foundation states that openness should enhance "universal

participation": anybody should have access to data and information to create and share knowledge without restriction, for self-empowerment, improved efficiency of products, services or governments, and democratic control (Open Knowledge Foundation, 2015). Open data is seen as a means to empower the people for democracy, universal access to information, and freedom of choice in terms of digital service providers. As such, Open Street Map is seen as a tool of counter-power against the monopolistic giant Google and its application Google Maps. In their philosophy, people should be allowed to choose or develop themselves their services with public, transparent, licence-free, and community-based data. Open data meet that ambition. As emphasised in their definition, open data is data that can be freely used, re-used, and redistributed by anyone for any purpose - subject only, at most, to the requirement to attribute and share-alike (Open Knowledge Foundation, 2015). As such, open data contributes to a form of digital activism, where community members, initially interested in using data to power their own websites and applications, turned into civic hackers and engaged in opening data for transparency and self-developed civic apps (Gray, 2014). The Foundation has spread over the world and operates as an umbrella organisation with "local chapters", i.e., regional offices across countries. They implement their mission: "*an open world, where all non-personal information is open, free for everyone to use, build on and share; and creators and innovators are fairly recognised and rewarded*" (Open Knowledge Foundation, n.d.)

Last but not least, I should mention the influence of data journalism, which together with the civic hackers but in their own way, advocates for transparency and universal access to information (Gray, 2014).

The academic push: new models of public governance and management

In research, the advocacy for open data mixes economical perspectives (*open innovation, New Public Management*) and democratic perspectives (*Open Government*). An important stream of research argues that governments should provide data not only for transparency but also more actively for others to develop public services and engage in governance. They suggest new models such as the Wiki Government (Noveck, 2009), Government-as-a-Platform (O'Reilly, 2011; Linders, 2012), or Do-It-Yourself Government (A. J. Meijer, 2012; Linders, 2012). They share the idea that technology and data can allow the citizens and the private sector to develop their own services of public interest, step in the governance process, and produce their own public value. To support that argument, O'Reilly (2011) uses the metaphor of the state as a vending machine: against the payment of taxes, the "old administrations" provide a limited number of standard services. If the client/citizen does not get what he expected, his "participation" is limited to protesting and shaking the vending machine. O'Reilly (2011) sees the platform model (i.e., the provision of infrastructures instead of services) as a way to open the development and supply of better public services thanks to the innovativeness of third parties while driving the costs down for the public sector. When a public actor goes for a platform thinking, it is believed that it opens to possibilities for new forms of online or offline collaborative arenas. Platforms create intersections and opportunities for collaboration between local policy-makers, urban activists and digital way of living, wherein open data is a pillar (Anttiroiko, 2016). Living labs, using *user-centred and participatory design* methods can be testbeds for new forms of collaboration enabled by ICTs and open data (Ruijer & Meijer, 2020).

O'Reilly (2011) emphasizes the economic benefits. Meijer (2012) emphasizes democracy and enabled participation, while Linders (2012) sees the Do-It-Yourself government as a model

where the citizens self-organise to supply public services (in a form of self-service) and exert democracy (in self-monitoring). All agree on calling for the opening of data rather than the provision of services or information. They have paved the way for open data.

## 2.2. Understanding open data through a value chain

Knowing the diversity of perspectives and motivations of the actors involved in open data, we can foresee the potential complexity of the open data value creation process. For the same object, the actors can have their own definitions of value and expected outcomes. An outlook on the technical characteristics is also important to grasp the phenomenon.

I introduce this complexity in a value chain, with a focus on roles.

In the next subsections, I **position** the choice of the **value chain** against a process perspective and the focus on **the role** against the focus on the coordination of tasks (Section 2.2.1.).

Then, I present open data by simplifying and breaking down the main roles and components of its value chain for explanatory purposes. First, I define **the input** (the data and its formats, Section 2.2.2.), then, **the actors and roles involved** in the data value chain (Section 2.2.3.), **the output** (what the reuses of data, in their multiple forms, can be, Section 2.2.4.), and finally **the outcome** (expected benefits and impact, Section 2.2.5.). The outcome is related to the values and initial motivations of the actors, in other words, the higher purpose of open data.

These components are finally sketched into a **conceptual framework** (Section 2.2.6.): a simplified open data value chain that is used to emphasise the possible misalignments, research gaps and contributions of the studies.

### 2.2.1. Why the roles in a value chain?

**A process** is a model that sums up the activities carried out to design and produce a product or service. It presents the transformation stages of an input to become a desired output. It can be descriptive and show *what is done or happens* or be prescriptive and conceptualise *what should be done*. A process is an abstraction subject to choices regarding the level of details and the modelling (what activities and properties to present and how to display them), and biases regarding what is considered significant (Eckert & Stacey, 2010). Process models are common tools used in Business Process Model, engineering, and design, amongst others.

While a process focuses on the activities to transform an input into an outcome, **a value chain** is a type of process that focuses particularly on the *value generated by these activities*. Originally, it was introduced by Porter (1985) as a strategic tool for firms to increase their competitiveness in a specific industry. A value chain analysis differentiates the value-adding activities, the necessary activities (support activities), and the unnecessary activities (evaluated inefficient), in addition to the cost of producing a service or product, to isolate the margin made by the firm. To a larger extent, it allows reflecting on how a company can gain a competitive advantage in a production line that goes beyond its borders, at the scale of an industry (Harmon, 2011). The value chain has been used in other contexts, for example, to challenge the effects of globalization and the revenue of small producers as an effort to encourage equity (Kaplinsky & Morris, 2001).

I see the value chain as a relevant framework for the following reasons:

- The focus of the research project (Wal-e-cities) is on value creation and citizen participation, which a value chain can highlight. It can raise critical issues regarding the value-generating activities and (lack of) value captured by the actors involved (Kaplinsky & Morris, 2001), a concern in open data.
- The value chain is customer-oriented: it breaks down the internal and external processes into small processes that should define what inputs are of value for its internal customers (e.g., other companies in the chain, other employees in the production line), and outputs are of value for its final customers (e.g., end-user of a solution) (Harmon, 2011). The end-user value is also a concern in open data.
- Finally, a value chain is about the actor's roles in value creation. One important dimension of a value chain analysis is the governance of the actors involved in the value creation: it implies that value creation is interactional and should be coordinated to succeed, as each actor needs and creates value to enter the chain and to capture some kind of value to stay (Piboonrungroj, Williams, & Simatupang, 2017). This assumption challenge one assumption in open data: the one that value is created by the spontaneous and independent involvement of people.

However, in this thesis, I limit my use of the value chain as a frame, a conceptual model that I believe is more comprehensive than a process model type "Input – Process (transformational activities) – Output", to articulate the issues important to my research. It is an abstraction or metaphorical presentation of reality. Hence, I do not conduct research to develop a comprehensive open data value chain. Instead, I use the value chain to keep the value creation in the background of my research questions, to put the role of the actors in the foreground, and to present misalignments between its main elements: actors, roles, input, output, and outcome.

These elements are further described in the coming sections.

### 2.2.2. The input: open data, definitions and principles

**Data** can be defined as a representation of objective facts that can be stored and manipulated. Once processed, it can become information, which, accumulated and properly used for specific purposes, can create knowledge (Hey, 2004).

How could we define **open data,** then? As discussed in Section 2.1., there are many perspectives at stake. In search of a definition, one should think of the background and orientations of the definition's writer. For a "neutral" definition of open data, we could focus on the **ideal characteristics** that should make it "open". However, the intent to regulate and standardise is also motivated by an agenda and political orientations. For instance, an Open Government working group, held in Sebastopol, California, in 2007, worked on defining general principles that would make the data release, and the publishers (the governments), more open. They ended up with eight principles since then updated and augmented to ten principles by the Sunlight Foundation (an American not-for-profit organisation that advocates for open government) (Sunlight Foundation, 2010).

**Ten Principles for Opening up Government Data** (Sunlight Foundation, 2010)

The government data shall be considered open if it is:

1. **Complete:** all public data that is not subject to valid privacy, security or privileged limitations should be available.
2. **Primary:** the data should be in its original form, as collected by the government and published with original source documents and method of data collection.
3. **Timeless:** the data should be made available in a timely fashion, as quickly as collected and at best, with real-time data.
4. **Easy to access physically and electronically:** access requirements, procedures, forms or systems that require specific technologies should be avoided. By contrast, interfaces to find and download datasets easily, and Application Programming Interfaces (APIs[2]) should be privileged.
5. **Machine-readable:** data should be structured to allow automated processing (which a PDF document is not).
6. **Non-discriminatory:** data should be accessible without registration or membership requirements.
7. **Non-proprietary (commonly owned or following open standards):** data is available in a format over which no entity has exclusive control, and do not require licensed software to be processed (e.g., Microsoft Excel).
8. **Licence-free:** data is available without restriction on use (e.g., terms of use, attribution requirements, copyright, trade regulations).
9. **Permanent:** data released online should be available in perpetuity, with appropriate version-tracking (no alteration should be made without notification).
10. **Bearing no usage cost:** as most data is collected for governmental purposes, the existence of fees has little effect on whether the government collects the data in the first place but does create a barrier to reuse.

Those principles are commonly used as a reference in research (e.g., Ubaldi, 2013; Charalabidis et al., 2018a), although they are not the only ones. Each organisation advocating for open data can develop its own best-fitting principles and charters. For instance, the International Open Data Charter, a collaboration between governments launched at the margins of the 2015 United Nations General Assembly, published six open data principles for the governments (The Open Data Charter, 2015). In Belgium, the 20 principles of the Flemish Open Data Charter were approved by the 13 leading cities in Flanders, in the summer of 2018 (Smart Flanders, 2018).

The purpose of open data principles is to enforce data release in a way that should make it reused, serve the expected outcomes of the publishers, and comply with the law. The principle of **anonymity** is declined among the different versions of charters and guidelines. Anonymity is

[2] Application Programming Interfaces (APIs) are sets of functions or intermediary software that allow two applications to communicate with each other, for example, to access data and interact with external software components, or micro-services. They ease the reuse of open data.

understood as a way to comply with the GDPR[3] while publishing data (*anonymity of the data*), and a way to encourage reuse (*anonymity of the re-user* that should access the data without registration).

In practice, open data can still have many shapes: from a PDF of a scanned council report, an excel spreadsheet listing the local events of a municipality with limited metadata, to the statistics of demography, as a formatted dataset on a data portal, with extra documentation to help the re-user in its labour. It is not guaranteed that the publishers will follow the principles, depending on their capabilities, resources, or political choices. This said, the lack of standards among publishers does impact the easiness to reuse and data quality.

The famous Five-Star Model of open data (Figure 5), proposed by Tim Berners-Lee (2012), the inventor of the World Wide Web and director of the World Wide Web Consortium, is a tool that attempts to guide the publishers of open data towards better data quality and can be used as an indicator of the data maturity. In that model, the five-star data is the linked open data, a form of data that eases automatic integration (for combination and aggregation) even for a large number of datasets.

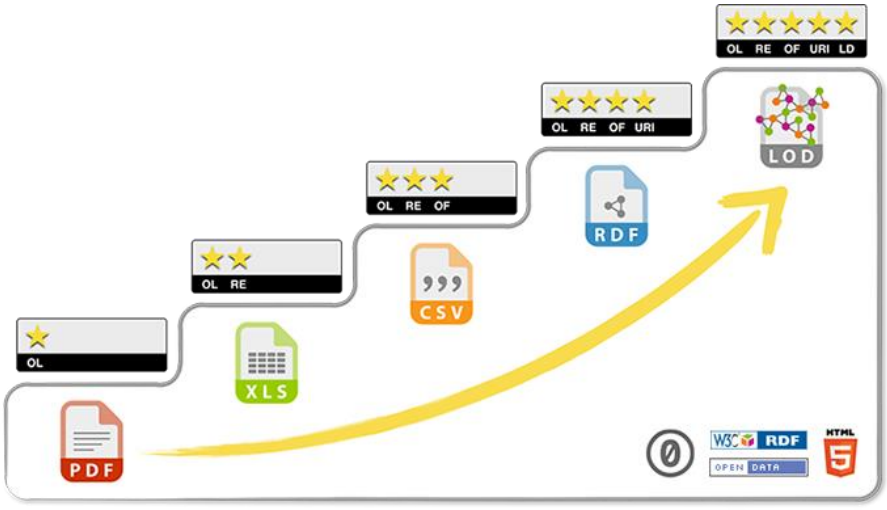


*Figure 5. The Five-Star Model of Tim Berners-Lee (Illustration from Berners-Lee, 2012, available under CC0 Public Domain Dedication)*

The technical characteristics are therefore a key aspect of open data, which can enable or impede its reuse and resulting outputs. The reuse of data creates its value, as data, out of context and interpretation, has little meaning. The actors involved in the process play a significant role in this value creation.

[3] The General Data Protection Regulation (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information from individuals who live in the European Union.

### 2.2.3. The main roles and actors

As previously explained in the introduction (see Section 1.2.), actors and roles are two different concepts that can be confusing because used interchangeably and context-specific.

In open data, the discipline of the researchers and the purpose of the represented process have an impact on the roles identified. For example, Attard, Orlandi and Auer (2016a) create a typology of six roles based on a data life-cycle that ignores the potential existence of an end-user. The so-called consumer is the consumer of data, not service. From an economic perspective, Berends, Carrara, Engbers, and Vollers (2017) acknowledge the end-user, but circumscribe the task of reusing the data to the public and private sectors. Lassinantti, Ståhlbröst and Runardotter (2018) are more comprehensive in their understanding of who makes use of the data, including civil society and citizens, while Safarov et al. (2017) point out that the citizens are both supposed to use data and finished solutions while their "participation" has to be proven. That confusion is the result of two observations:

- In open data, one actor can take several roles (roles are interchangeable and not limited), which generates complexity (Attard et al., 2016b).
- Roles' definition and scope are influenced by the discipline and perspectives of the people creating and using them. There is no agreement in the literature about the use of specific labels for specific roles or actors. For example, "citizen" is used as a role by the proponents of open data for democracy and open government (people in the role of using open data-based information to participate in a democratic process; Ruijer et al., 2017). For others, the citizens are considered actors in the open data reuse process (the citizens are developers, users of open data-based services, or both; Abella et al., 2019).

To solve this problem, I conceptually make a distinction between actors and roles. Figure 6 builds upon the concepts and definitions brought in the introduction (Figure 1) to add the influence of the paradigms presented in Section 2.1. The paradigms and the distinction between roles and actors are key to capturing the misalignments in open data. Hence, I integrate them later into the conceptual framework (Section 2.2.6.).

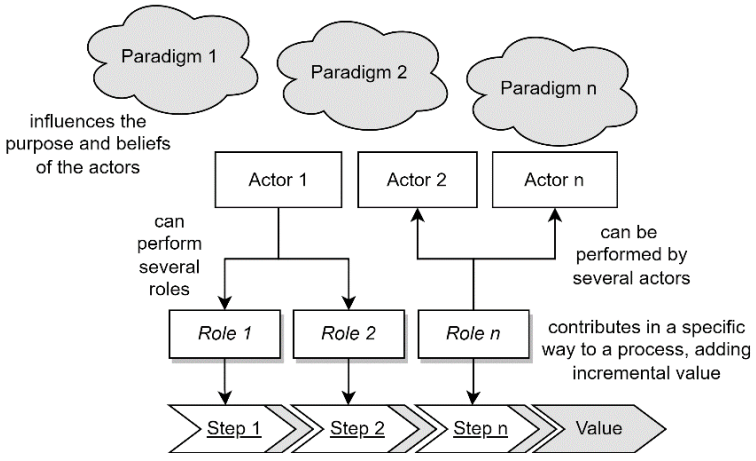


*Figure 6. Roles, actors, and their paradigm in a value chain*

In open data, I identify three main roles that are inevitably a simplification of reality but are suggested as general categories to understand the basic activities required to create value.

The first role in the value chain is the **"publisher"**, the people or organisations that publish data. Any organisation (e.g., governments, businesses, or NGOs) or individual (e.g., researchers, citizens, or developers) can create and publish its data on the Internet for others to reuse. Specifically, the data produced and published by governments, public institutions or agencies is called "Open Government Data" (OGD). This role is also found in the literature under the label "supplier" (Berends et al., 2017) or "provider" (Zuiderwijk, Janssen, et al., 2014), with a more economic connotation. Some authors dissociate the data "producer" from the data "publisher" (Attard et al., 2016a), as the first might need the expertise and services of the second to get its data online. Indeed, publishing data implies a series of preliminary tasks, including data creation, cleaning, compliance with legal requirements and licences, and publication (Attard et al., 2016b).

The people or organisations that reuse the data are referred to as the "**re-users**" (Berends et al., 2017), "**reusers"** (Abella et al., 2019) or "**users**" (Susha et al., 2015). The actors behind these roles can belong to different social groups, for example, citizens, developers, public organisations, businesses, NGOs, start-ups, or journalists (Lassinantti et al., 2018). They have different paradigms, motivations, and capabilities to reuse data (Lassinantti et al., 2018). Through a series of phases and tasks, that include the identification, acquisition, and enrichment of the data (Crusoe & Ahlin, 2019), their work adds value to the data. The re-users can also be **intermediaries** for others to enable them to benefit from open data.

When the re-users work for others, not able or willing to manipulate data by themselves, they address their solution to potential **end-users.** The end-users are in fact **information seekers** (facts, data applied to questions, contexts, to become information), or **users** of finished solutions (visualisations, products or services). There is confusion in the literature as the label "users" and "end-users" can be used for the same position in the process (see e.g., Abella et al., 2019). Depending on the output of the data reuse and its context, the end-users can be associated with categories of actors such as citizens if the solution is a service of public interest, or consumers, clients (B2C), and companies (B2B).

The term "user" is confusing in the literature, as authors give it a meaning specific to a study (a research paper) or discipline (innovation, business management, public administration), sometimes encompassing different roles. For example, the user in the innovation literature is a user of a product or service, possibly an innovator, or co-creator (see Section 2.1.1). The user in the open data literature is mostly the user of data, sometimes the user of a solution. In this thesis (Introduction, Background, Discussion, and Conclusion), I name the three main roles as the ***publisher*** (to emphasise the action of publishing data) – ***re-user*** (to emphasise the action of reusing the data for an output) – and ***end-user*** (to emphasise the final position of this role in the data value chain and its externality to the data reuse process).

This standard was not followed in the published studies (Chapters 4, 5, and 6), given that the collaborations and reviewing process led to specific compromises. However, each paper redefines the concepts and terms used in context to support the reader's understanding.

### 2.2.4. The output: the multiple forms of open data reuses

The result of the reuse, the output, depends on the re-user's intentions and capabilities. There is no standard way of using data and no standard output. Unlike in open software communities, the purpose of the open data re-user is not always to develop a finished solution, software, or application, and the project of reuse is not always collective. Re-users can be hobbyists who reuse data for fun, to learn new skills, or solve problems in their leisure time: they do not care about finishing projects (Smith & Sandberg, 2018).

According to Davies (2010), open data reuse can lead to:

#### Data to Data

The re-user uses open data to upload back more qualitative data for use by other re-users. For example, by converting formats, augmenting or combining datasets to create new datasets, and providing new APIs on open-source platforms or repository services such as, for example, GitHub[4]. The re-users create value for their own project and other re-users, their peers, and their community (Lassinantti, 2019).

#### Data to Facts

The re-user manipulates data to find facts of interest. As explained by Nicholas & Herman (2010), finding facts is a recurrent and everyday need that is relatively precise, well defined, and requires a limited amount of data. The need can be fulfilled by answering straightforward questions (e.g., how, what, why, where, who, when?). The facts can be, for example, names, dates, or more complex statistics. The re-user does not necessarily report or share with others the found facts, since the need can be unique to the person, its interests and current knowledge. For example, an aficionado of local history who would like to know about the voting history of his municipality just for the fun of it. Facts can be shared on simple digital solutions that visualise the answers to simple questions to support the re-users in their civic engagement and advocacy (Davies, 2010). For example, how many streets in Brussels are named after a man, a woman, or a trans person? (Open Knowledge Belgium & Noms peut-être!, 2020).

#### Data to Information

The re-user manipulates data to generate information. Unlike facts, information is not just extracted from data: it is contextualized and reported in texts or graphics. Hey (2004) describes information as a resource that can exist in abundance and become overwhelming. By combining pieces of information, people can make *informed* decisions and build knowledge. For example, the re-user can be a student and creates a visualisation for an assignment, building up its own understanding of a situation with data. A local politician could write a report to defend a project at the next council, based on the interpretation of several data sources. Data journalists put in perspective data to raise awareness. For example, the YouTube channel *Data Gueule* (Poulain, Goetz, & Lapoix, n.d.) explains unknown mechanisms in society and challenges political statements with data.

[4] GitHub is a website for developers and programmers to collaboratively work on code and open-source software.

#### Data to Interface

The re-user combines, aggregates data, writes code and develops an interactive presentation of the data. The interface can be used by others (users of the interface) and provides information customized to their input. For example, where is the closest recycling centre to the user's position visualised on a map, or a heat map showing activity by neighbourhoods, filtered by type (traffic accidents, robberies, air quality, population density). The interactive presentation of datasets through customizable maps or charts is a typical feature of open data portals that try to become more "user-friendly" and accessible (e.g., see the portals developed by OpenDataSoft, 2021, implemented by the municipality of Namur; Ville de Namur, n.d.-a).

#### Data to Services or Products

The re-user integrates the data into existing products, and services, or creates new ones. For example, a service can be an application to calculate the best routes with alternatives to public transport and car-pooling. The solution is not necessarily exposed to external end-users, for example, if the data is used to optimize the trash collection in public bins the end-user is within the public organisation.

In sum, the open data's outputs can be digital or not, tangible or not, sharable and shared or not.

### 2.2.5. The outcomes of open data

The outputs resulting in shared information, interfaces, services or products create what Lassinantti (2019) calls a "broader value of open data". It can be captured by the re-users, any type of end-users (people or organisations), including the public sector. I define that broader value as the **outcome**, as it is generated when open data or solutions using it are consumed and complete open data's higher purpose. This outcome has a positive connotation in previous research, as it is also referred to as benefits (Janssen et al., 2012; Zuiderwijk et al., 2019) or impact (Carrara et al., 2015; Neves et al., 2020; Worthy, 2015).

A publisher that relies on a re-user to reuse data always expects an outcome. The expected outcomes are the most influential factor for a public organisation to engage in open data in the study of Wang & Lo (2016). Data release requires resources (work, time, infrastructures) and would not be undertaken without a higher purpose. Zuiderwijk et al. (2019) reviewed the scientific literature and policy documents and classified the delivered benefits, as mentioned in the reviewed sources (publishers from the public sector), into three categories.

*The economic benefits.* For instance, beyond the creation of services, the expected outcome is that the new services would create new jobs or efficiencies in public service supply. Other economic benefits can include stimulated competitiveness, innovation, growth, reduced spending, and improvement of services.

*The political and societal benefits.* The available public data about public budgets and expenses could create transparency, and the traffic data could improve the lives of the citizen by allowing them to save time and avoid dangerous areas. Benefits in this category can also include an improved policy-making process, increased collaboration, participation, trust in government, and democratic accountability.

*The technical and operational benefits.* An administration can improve its internal processes by reusing its data (in terms of decision-making and efficiency), access to the wisdom of the crowds to solve problems, and easy access to data can increase the ability of the stakeholders to reuse data. The outcome is not directed towards end-users, but re-users (whether the administration or stakeholders), to create, in the end, more public value.

In sum, the **outcomes** can regard the **creation of economic, public, and social value** (Attard et al., 2015).

### 2.2.6. Framework: a simplified open data value chain

To conclude, the following framework (Figure 7) is suggested as a summary and a simplistic value chain that tries to articulate and put together all the elements introduced in the previous sections. It is similar to Figure 2 in the Introduction but adds a box for the actors influenced by their paradigms. The actors' diversity adds an extra layer of precision to the framework, to differentiate the influencing paradigms or rationales and actors' roles in society that influence the form of the input (data), output (reuse) and the expected outcome (higher purpose). The singular actors' intentions and other roles in society shape unique outputs and outcomes. They are important in the conceptualisation of misalignments.

This framework represents value creation as a sequenced process where actors take roles and their respective tasks come one after another to deliver an output, and potentially an outcome. In actuality, things do not happen so smoothly. I use the framework to highlight complexities in the tasks and definitions of roles, diverging actors' objectives, that can create possible misalignments between the actors, the roles, or the components of the process, and in the end, hinder the realisation of the expected outcomes.

Those possible misalignments are suggested as a novel approach in open data research, and they unravel the knowledge gaps that the present thesis aims to fill.

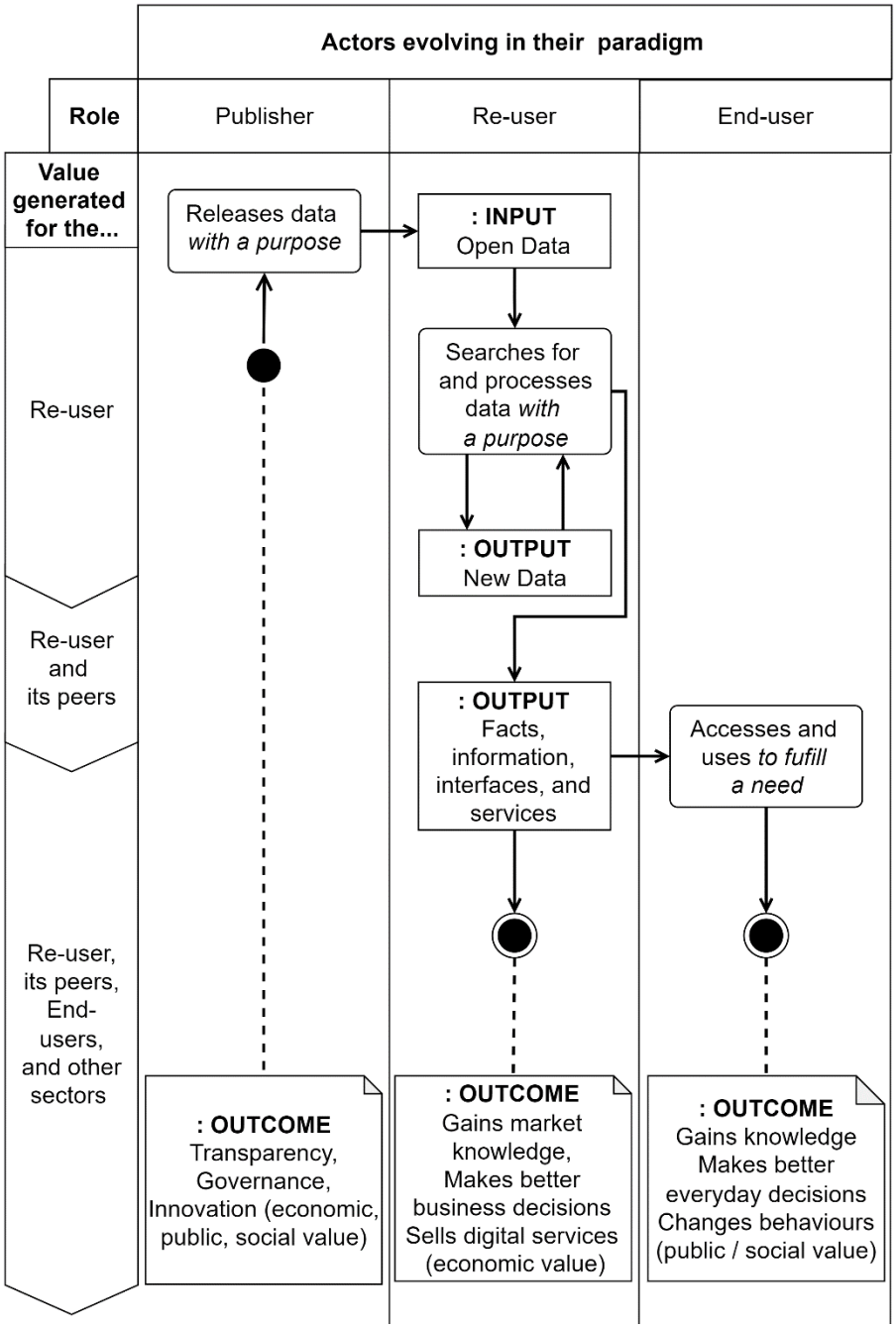


*Figure 7. A schematic open data value chain*

## 2.3. Possible misalignments

In this section, I situate four possible misalignments in context and show their relevance for practice and research. I follow the following reasoning, as represented in Figure 8.

I first show that ***open data is not reaching full fruition***: the expected outcomes are not always realised, which is a source of criticism against open data in part of the literature, and difficulty for the practitioners. It is also the main motivation and starting point to undertake this research. I assume that if a broader value is not achieved, from a process perspective, the ideal value chain is possibly hindered by misalignments.

Secondly, to understand the misalignments in the chain, I present a selection of ***possible sources of misalignments*** (in reference back to the definition of misalignments presented in the introduction). These sources articulate and problematize elements introduced earlier in the background without being too specific and redundant which the background sections of each study.

Thirdly, I formulate ***four possible misalignments*** that can cause problems and hinder the value creation and realisation of the expected outcomes. The misalignments conceptualise the way certain components of our ideal process are not fitting together, which can hinder its completion and outcome. I focus on misalignment related to actors and roles and see them at three conceptual levels. They are not exhaustive but are a personal expression of my understanding of the subject.

Finally, to highlight the practical significance of these misalignments, I point out a ***few possible specific consequences***. It closes the loop, bringing us back to the unrealised outcomes introduced as the starting point of my reflection.

In the last section (2.4.), each possible misalignment is used to pose the research questions in line with identified theoretical problems and research gaps.

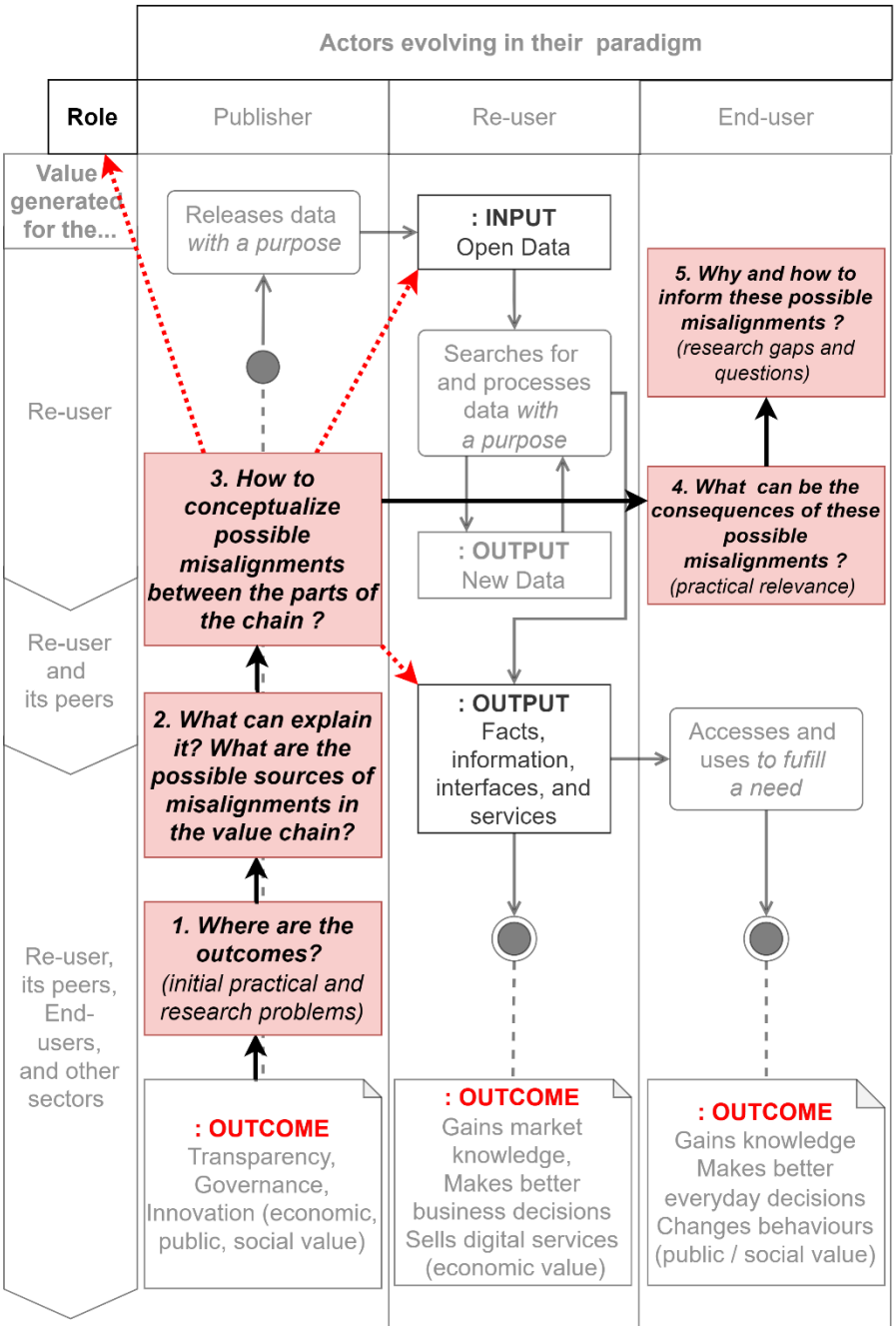


*Figure 8. Development of the possible misalignments and research questions*

### 2.3.1. Where are the outcomes?

I qualify the outcomes (higher purpose) as *expected*, since the literature is not agreeing about their realisation. What Zuiderwijk et al. (2019) present as "*benefits* delivered *by open government data initiatives as mentioned in the literature and policy documents*" in their recent paper relies on early literature. Of the 20 included papers, 2 papers date from 2002 and 2004, before the rise of open data, and 16 papers date from 2009 to 2013, when open data just started to draw the interest of research and practice. These studies might lack a critical distance to assess the impact of open data. Recent studies (e.g., Abella et al., 2019) stress that the impact of data reuses on society and the mechanisms to succeed are not clear. Worthy (2015) argues that impact is unpredictable and political.

*In terms of public and social value*, for instance, later studies have criticised the lack of empirical data to prove citizen participation (Safarov et al., 2017). In policy-making and governance, mechanisms to encourage and process citizens' feedback based on open data are weak, infrequent, completely missing (Reggi & Dawes, 2016), or research-driven (e.g., Ruijer et al., 2017). Moreover, there is no guarantee that the release of open data leads to increased transparency since all the data is not published, all the citizens are not capable to reuse it or aware of its existence, transparency can be reached by other means, and OGD can exist in a context of low trust (Yu & Robinson, 2012). For example, O'Connor et al. (2019) study the national case of Kazakhstan, an authoritarian state, where OGD is released for international strategy and mimicry. Still, practitioners cite the increased transparency as the first benefit realised while releasing OGD (Zuiderwijk et al., 2019).

*In terms of economic value*, studies funded by the European Commission for the European Open Data Portal are positive in terms of market size, employment, and perspective of growth (e.g., Berends, Carrara, Engbers, et al., 2017; Huyer and van Knippenberg, 2020). Other research papers, however, can be more critical. The creation of commercial value is defined as a wicked problem that needs more policy support (Zuiderwijk, Janssen, Van De Kaa, & Poulis, 2016), and capabilities (infrastructure with high quality and demanded data) to support the development of sustainable business models (Ahmadi Zeleti, Ojo, & Curry, 2016).

The realisation of the open data outcomes is not certain. What can explain this difficulty? What can be the possible sources of misalignments in the process, specifically regarding the actors, roles and their respective position and contribution to the value chain?

### 2.3.2. Possible sources of misalignment in the value chain

I identify six possible sources of misalignment that can become problematic when it comes to realising the open data outcomes. They come from my understanding of open data, based on my previous knowledge developed in the background (Chapter 2.) and the chosen focus on the actors and their roles. They problematize the concepts and paradigms introduced earlier in the perspective of value creation.

#### *1. Actor's coexisting paradigms and rationales*

Each actor evolves in its own environment and looks at open data with its own perspective, shaped by its paradigm and value system. Section 2.1. has emphasised the multiple perspectives that can exist on open data, even inside a single sector. The way an actor defines the purpose of open data, its expected outcome, and the way to realise it (e.g., what roles, responsibilities, resources involved, and output delivered) are influenced by its paradigm. This

situation can create misalignments between the actors' perception of other roles (publishers, re-users), purposes, and outcomes.

Publishers from the public sector, for example, municipalities, can be driven by an open government paradigm (data for transparency, increased participation and collaboration with the citizens). For that, it can, for example, release data about demography, public budget, and expenses. In a New Public Management or Do-It-Yourself Government paradigm (see Section 2.1.2.), open data is expected to foster innovation and the development of services or solutions, thus public value, by the stakeholders. In that case, the idea is that the provided data will sparkle ideas, the development of innovative services for the citizens, and a better quality of life, with the involvement of external problem solvers (e.g., citizens, developers) (McBride, Aavik, Toots, Kalvet, & Krimmer, 2019). The municipality acknowledges its lack of resources and innovativeness, and therefore, opens its data to compensate by attracting external innovators in an open innovation-like approach (Chesbrough, 2011; Chan, 2013).

However, it is not sure that the re-user is in the same mindset and will use open data as the publisher expects him to do (Gonzalez-Zapata & Heeks, 2015). Re-users belong to several social groups, with their specific motivations. Lassinantti et al. (2018) identify five rationales for re-users to reuse data. First, is the exploration of creativity. Typically, developers, students, and members of the civic tech or open-source community participate in hackathons, meet-ups, or try out at home to develop apps. Second, is the creation of business value. Start-ups and companies of all sizes can enhance their existing solutions or create new ones with open data. Third, enabling local citizen value. Civil society organisations, developers, and city managers might try to engage citizens to develop new services and speak their minds about local issues and governance. Fourth, is the will to address global societal challenges. NGOs, entrepreneurs, and researchers use open data to understand problems, spread information, inform policies, and monitor governments. Finally, the advocacy for open data agenda. Developers, journalists, citizens, and civic tech communities advocate for the principles and expected outcome of open data: co-creation of open government mechanisms, triggering transparency by opening more data and reinforcing data infrastructures. Their motivation shapes their contribution to the process and output.

*Possible source of misalignment 1: Actors´ vision on open data. Open data brings many promises, but the actors involved can look and go in different directions due to their singular worldview and purpose.*

### *2. Defining the roles: diverse frames of reference and scopes*

Clarity in the definition and scope of the roles is important for the performance of a process. However, given the diversity of the publishers' and re-users' purposes and backgrounds, we can foresee that the roles can be made based on different frames of reference. In previous literature, the roles and their position in the value creation process can be based on two levels. Either, the roles are defined at the level of the transformation of data into an output (the data-related roles, from a technical perspective). Or, taking a larger perspective, the roles include the use of the data for an outcome (for a higher purpose in the "real world", roles involved in the creation of a broader value)(Lassinantti, 2019). It is a source of confusion in terms of positioning for the actors and terminology in research.

At the level of the data transformation, the literature does not agree on the terminology of the most basic roles and their tasks (see Section 2.2.3.). We also know that a narrow definition of

roles, limited to the data release and reuse process, for instance, generates roles that can be interchangeable, cumulated, and not limited to one actor (Attard et al., 2016b). For example, a municipality can be a publisher and re-user at the same time (Mergel et al., 2018), and many intermediaries can be needed to deliver one output. I argue also that narrowing down too much the definition of a role does not provide a holistic understanding of reality. For example, to understand the role of a publisher of open government data in cities, I think that we need to acknowledge that it is also, and first of all, a municipality with a public mission.

On the other hand, practitioners and researchers tend to make "all-in-one" roles, merging tasks and expectations from several frames of reference. The best and most recurrent example is the role of the citizen, often expected to be the re-user/developer and use data for civic purposes (Safarov et al., 2017). It results in roles lacking clarity and being difficult to fulfil by anyone.

When the roles are too vague or too narrow, it creates grey zones, overlaps or gaps between them in the value chain, and in the end, it can hinder the realisation of the outcome. The actors can have a hard time knowing what they are supposed to do in a role to fit with the previous and following roles in the chain. It could also mean that an actor expects others to do a certain work, while the others expect the first role to do it.

*Possible source of misalignment 2: roles' frames of reference. Roles are difficult to delimit in open data, the literature does not agree on terminology or set of tasks for one type of role, switching from a too narrow to a too large definition. It is not clear what is to be a publisher, or what is to be a citizen in an open data value chain.*

### *3. Reliance on third parties to realise "any" valuable output*

A third possible source of misalignment, inherent to the open innovation paradigm and the open-source movement, is the reliance on third parties to contribute and create value. The openness implies that the innovation arena, open to anyone, will attract someone to do a part and create anything (the output). However, the activity of releasing data requires an initial investment (work, time, resources, and infrastructure). Because of that, the publishers of open data do have at least a vague idea of what they want to get in return, which influences the way they see their position and role in the process, their degree of involvement to make it happen, and the input they deliver. Logically, the rest of the process and the development opportunities depend largely on the input provided (the type and quality of data). It means that the publishers, from the beginning, influence the outputs and their future outcomes that can drastically limit the possibilities of the objective of "any reuse for any purpose".

Open data is neither neutral nor universal. The realisation of "any" valuable output is subject to variables such as a match between the input provided, which is shaped by the work and purpose of the publisher, and the multiple and sometimes opposite motivations of re-users (Gonzalez-Zapata & Heeks, 2015; Lassinantti et al., 2018). To succeed, a process relying on the involvement of third parties cannot ignore their requirements. Without a minimum understanding of each other role's scope and purpose, the third parties might not show up or produce anything valuable.

*Possible source of misalignment 3: the gap to fill in the process (one size does not fit all). The underlying "openness" assumes that someone will do something wherein the publisher would have little control* (Janssen et al., 2012)*, while the publisher's input is not neutral and universal by nature. The publisher has generally an agenda. A lack of agreement between the actors*

*regarding their purpose and role's scope can make the creation of real value and publishers' agenda (wished outcome) a gamble.*

#### *4. Appropriateness of the open-source model*

Elaborating on the idea of interdependencies and initial investment, we can question the appropriateness and applicability of the model at the origin of the open data movement. Open data is influenced by the open-source communities and the hacker's values (sharing and contributing, see Section 2.1.1.). However, open data is different from open-source communities on at least three points.

(1) The publisher rarely makes money by releasing data. Instead, it costs. In open-source, on the contrary, the company releasing code makes savings and benefits directly from the contribution of the community (software development, knowledge exchange) (Lindman & Nyman, 2014). This cost supported by the publishers can be one of the reasons why they tend to publish data following their own priorities, starting first with data that is identified as low-risk (in terms of transparency, GDPR), easy to release, or considered as a thematic priority by the government (Ham et al., 2019). The release of data is supply-driven oriented and publishers lack feedback mechanisms (Evans & Campos, 2013).

(2) The open-source community contributes to developing a finished software based on open code, while in open data, at best, when the community collaborates with the publisher, it is around the data and its quality. After that, there is a long way to go to have a finished product and the re-user can limit its contribution to intangible output for its personal use (facts, information), or improved input for others (better data, APIs, prototypes) (Lindman & Nyman, 2014; Davies, 2010). Since publishers and re-users do not share a common goal, and since the reasoning is "provide open infrastructures instead of services" (O'Reilly, 2011), it opens the door to unlimited possibilities. At the same time, the value captured by each actor is not obvious.

(3) Open data is not fully open, unlike open code. The publisher keeps on being the owner and deciding what and when to publish, update, or withdraw. It creates dependencies and power issues along the value creation process: the publisher is dependent on the re-user for the development and maintenance of data-based services, and the re-user is dependent on the publisher for data access, accuracy, and stability provision (Sangiambut & Sieber, 2017).

*Possible source of misalignment 4: dependency and value captured by the actors. The value created and captured is not as obvious as in open software communities, which model influences the open data movement. The actors are interdependent and control different parts of the process while the open data's ideals and values tend to believe otherwise.*

#### *5. Nature of open data*

Despite the attempts of open data principles, charters, and other Five-Star maturity models, the provision of open data keeps on following no standards. The publishers face singular re-users with singular needs (van Loenen, 2018), using data to produce multiple types of outputs. In that context, standardisation could be just a dream.

On the side of the re-user, it raises another issue: task complexity. The re-users face multiple barriers and difficulties in their labour, abundantly discussed in the literature (Janssen et al., 2012; Zuiderwijk et al., 2012; Ma & Lam, 2019; Martin, 2014; Smith & Sandberg, 2018; Toots,

McBride, Kalvet, & Krimmer, 2017). Reusing data is complicated by nature, especially when a low quality and lack of metadata, above the lack of standards, create difficulties in accessibility, usability, compatibility, and combinability of datasets. Like a research process, data, which is information to be, can open many paths of reuse and solutions *as long as* the re-user has a question to answer or a need to fill (field knowledge), access to the right data, and the technical skills to make use of it.

*Possible source of misalignment 5: task complexity inherent to open data. Reusing successfully open data is not obvious and requires technical skills and field knowledge.*

### *6. Ghost end-users and misunderstood needs*

Finally, given the complexity and the technical aspect of open data, the lay end-users tend to be omitted in the process by the publishers and re-users, and generally in research.

The publishers are aware of the barriers and lack of reuse. The most engaged and resourceful try to stimulate the demand and reuse with mechanisms that should increase the "user involvement" or "citizen participation". In fact, they use often technology-driven tools and incentives (e.g., data training, hackathons, open challenges; Gascó-Hernández, Martin, Reggi, Pyo, & Luna-Reyes, 2018) which illustrates a certain technological determinism in the publishers approach to open data (Evans & Campos, 2013). Open data challenges are addressed to and appeal to like-minded participants: users of data, not end-users (Gascó-Hernández et al., 2018; Hivon & Titah, 2017; Susha et al., 2015; van Loenen, 2018). In such open data challenges, developers and hobbyists show up for the fun of challenging themselves, developing new skills, meeting their peers, and getting a reputation, before satisfying end-users' needs (Juell-Skielse, Hjalmarsson, Johannesson, & Rudmark, 2014). The hobbyists' purpose is not even to finish a solution for any end-users (Smith & Sandberg, 2018). In previous research, the concept of "user engagement" is also understood as the engagement of the re-user of data and omits the end-user (Susha et al., 2015).

The end-users' needs are neglected or interpreted through the perception of the re-users, who should perform a mediatory role. Unfortunately, seizing the end-users' needs is a complex matter as we are talking about information needs. A lay end-user does not need data but solutions that answer his daily questions and information needs. Those needs are difficult to capture because temporary, personal, dependent on previous knowledge, and can be fulfilled by many competing information sources (Nicholas & Herman, 2010). When an output does not match the needs of end-user groups, it might interest only few people beyond its developer.

*Possible source of misalignment 6: misunderstood end-users' needs. Due to its technical aspect, lay users are not appealed to open data. They do not need data, but information. These needs especially challenging to capture for the re-users, supposed to play an intermediary role.*

### 2.3.3. Possible misalignments

I conceptualise four possible misalignments at three levels. These misalignments are a personal expression of my understanding of open data value creation issues, through the lens of roles. Other misalignments can exist but go beyond the scope of this thesis.

Misalignments 1 and 2 take the actors as a starting point, in regard to other actors' roles, and outcomes. Misalignments 3 and 4 take the roles as a starting point, in regard to their fit in the process (with each other contributions -input/output- and needs in the value chain).

At the Actors ↔ Roles' level:

(1) A possible misalignment between the actor's perception of other roles (in terms of definition, scope, and purpose);

At the Actors ↔ Outcome's level:

(2) A possible misalignment between the municipalities' expected and realised outcomes;

At the Roles ↔ Process' level:

(3) A possible misalignment between the publisher's input and the re-user's requirements and purposes;

(4) A possible misalignment between the publisher's input, the re-user's output and the end-users' needs.

In Figure 9, I sketch and add the possible misalignments in the value chain introduced earlier.

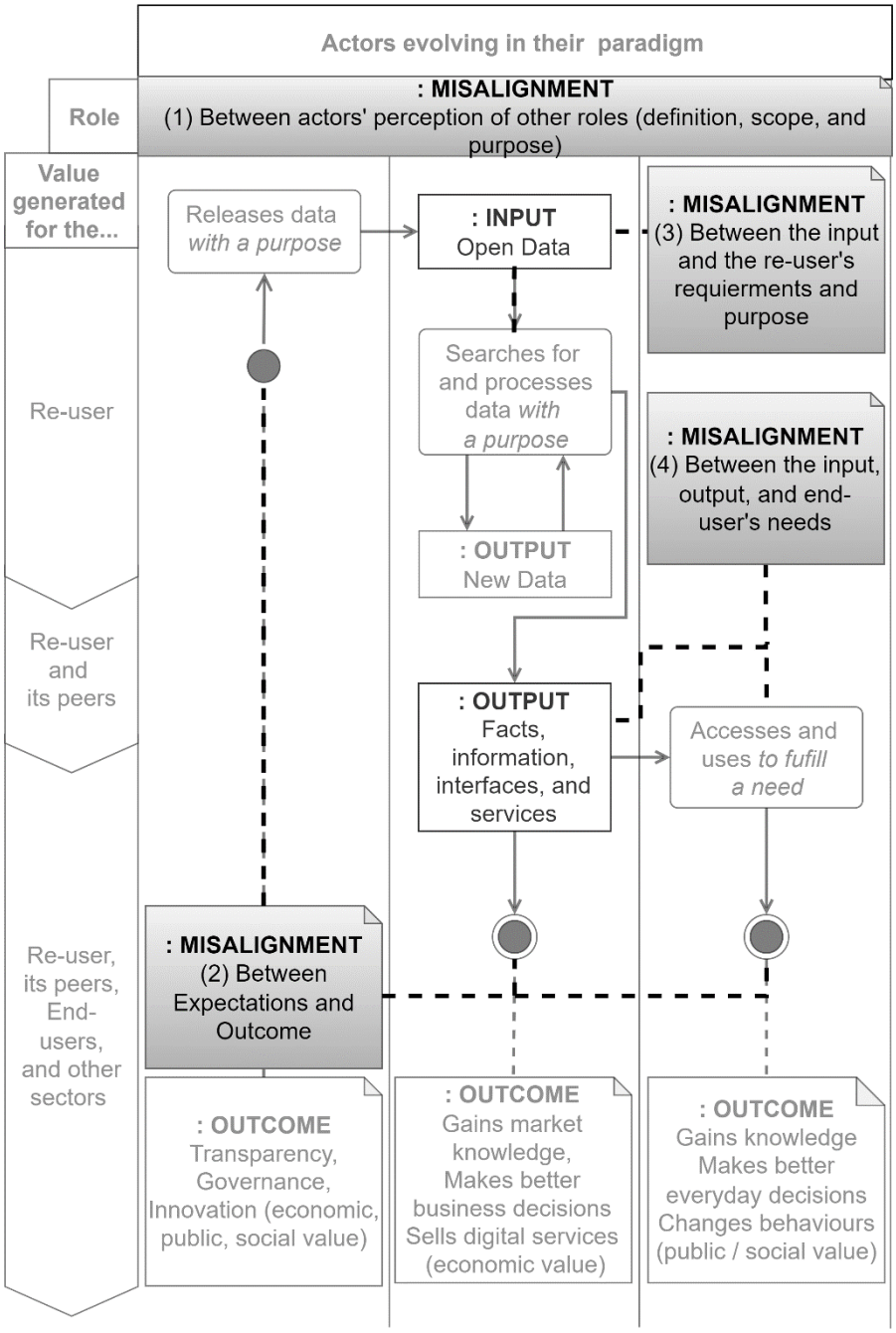


*Figure 9. Possible misalignments that can impede the value creation and outcomes achievement*

### 2.3.4. Possible consequences

Misalignments can hinder the successful completion of the value chain and the realisation of its expected outcome. I list here a couple of possible consequences per misalignment.

**Possible misalignment 1) When the actors do not align their roles (in terms of definition, scope, and purpose):**

- **Blurry value creation process**. Jetzek, Avital, & Bjorn-Andersen (2013) revealed four mechanisms of value creation with open data: transparency, participation, efficiency and innovation, later complemented with enabling factors (Jetzek, Avital, & Bjørn-Andersen, 2014). They however concede that we lack knowledge regarding what needs to be done to create value, who should do it, how, and for whom (Jetzek, Avital, & Bjørn-Andersen, 2014).
- **Unfilled roles**. Safarov et al. (2017) point out that we are lacking evidence to claim that the citizens would participate with open data. Attard et al. (2015) refer to the "90-9-1" rule (Nielsen Norman Group, n.d.), to recall that most users do not take the role of contributor in online communities: 90% of users follow and read, 9% contribute occasionally, and 1% are very active and account for most contributions.

**Possible misalignment 2) When the municipalities' expected outcomes do not align with the realised outcome:**

- **Tensions between the actors and opposite requirements**. The actors can have conflicting opinions about the value to be created, based on their rationales and paradigms. In the study of Martin (2014), some members of the re-user community perceive that private sector exploitation of open data risks undermining social and value creation, while others perceive financial exploitation of open data as either a core part of the agenda or an appropriate means of creating social value.
- **Outputs of various forms, unrealised outcomes**. Lassinantti et al. (2018) reveal that re-users focussing on the expected outcome (economic value, social value) stay vague about the way to reach it. Their main concern is data availability, which becomes the output they work on. Re-users focussing on creativity and exploration of open data's possibilities stay vague about the problem they address. Their interest is to use methods (e.g., hackathons), technologies, and data. The outcome of the re-users might not meet the expectations and ambitions of the publishers, who claim to publish data for the creation of a higher value.

**Possible misalignment 3) When the publishers do not align their input with the re-user's requirements and purpose:**

- **Unused data.** Publishers who go for a data release based on data availability or their own thematic priorities can face lower re-use of their data than those choosing a market-based approach (demand-driven) (Ham et al., 2019).
- **Unexpected or unrealised outcome.** Limited reuse of open data ends up with a limited impact on the end-users and society (Zuiderwijk & Janssen, 2014).

**Possible misalignment 4) When the publishers do not align their input and re-users, their output, with the end-users' needs:**

- **Unused outputs.** If the end-user does not see how a solution can satisfy its information needs, it might choose other alternatives available, as the information sources are numerous and easy access is determinant for the user when seeking information and selecting digital solutions (Nicholas & Herman, 2010).
- **Unrealised outcome.** Limited reuse of open data and the solutions using it ends up with a limited impact on the end-users and society (Zuiderwijk & Janssen, 2014).

A synthetic table (Table 5) summarises linearly the possible sources of misalignments, conceptualised in misalignments, and possible consequences before we move to the research gaps and research questions that close the chapter.

*Table 5. Summary of the possible sources of misalignments, misalignments, and consequences*

| Possible sources of misalignments | Possible misalignments | Possible consequences |
|---|---|---|
| • Actor's vision on open data (co-existing paradigms)<br>• Roles' frames of reference (in society >< data roles, interchangeable, unclear)<br>• A gap to fill in the process (Reliance on third parties to create value) | (1) Misalignment between the actors' perception of other roles (definition, scope, purpose) | Blurry value creation process (misunderstanding of the roles, tasks, and whole value creation process),<br>Unfilled roles<br>→ Unexpected or unrealised outcomes |
| | (2) Misalignment between the municipalities' expected and realised outcome | Tensions between the actors and opposite requirements<br>Outputs of various forms<br>→ Unexpected or unrealised outcomes |
| • Dependency and value captured by the actors (appropriateness of the open-source model)<br>• Task complexity (Nature of open data)<br>• Misunderstood end-users' needs (omitted in the process) | (3) Misalignment between the publisher's input and the re-user's requirements and purposes | Unused data<br>→ Unexpected or unrealised outcomes |
| | (4) Misalignment between the publisher's input, re-user's output and the end-users' needs | Unused outputs<br>→ lack of impact on society and end-users (resulting from the unrealised outcome) |

## 2.4. Research gaps and research questions

I use the possible misalignments to determine the research gaps and the research questions that address them. The research questions do not tackle all the issues introduced. Nevertheless, they cover the three main roles, the four possible misalignments (Table 6), and together aim at building up a consistent contribution (see Table 1, p. 22).

*Table 6. Possible misalignments and research gaps addressed by the three studies*

| Addressed misalignment(s) | Research gaps | Research questions |
|---|---|---|
| **(1) Actors ↔ Roles** (actor's perception of other roles)<br><br>**(2) Actors ↔ Outcome** (municipalities' expected and realised outcome) | **The citizens' role(s)** (scope, tasks) given the co-existing paradigms and rationales<br><br>The position of these roles in the value chain and value created (versus outcome) | **Study 1:**<br>*What are the innovation approaches taken by participatory data-driven initiatives in urban settings?*<br><br>*What can be the citizens' roles and patterns per innovation approach?* |
| **(3) Roles ↔ Process** (Publisher's input >< users´ requirements)<br><br>**(1) Actors ↔ Roles** (actor's perception of other roles) | **The municipalities' and their users' role(s)** (scope, tasks) in the release and reuse process<br><br>How do these roles relate to their primary role (a public actor) and re-users or end-users' roles (expectations towards their "users") | **Study 2:**<br>*What are the possible municipalities' roles within OGD release and reuse?*<br><br>*What expected users' roles are implied by the municipalities' roles?* |
| **(4) Roles ↔ Process** (publisher's input, re-user's output >< end-users' needs) | **The end-users' information needs and ways to give them a more active role**<br>How to capture their information needs and involve them in the process | **Study 3:**<br>*What design principles could data providers and intermediaries follow to*<br>*design methods to identify the information needs of groups of users?* |

**The first research gap** focuses on the role of the citizens (as third parties) in open data initiatives.

In research and practice, there is a lack of knowledge regarding the citizens' roles and scope in the context of open data. Previous research uses the term "citizens" interchangeably for the re-user, end-user, and civic roles, which creates ambiguities about the meaning and rationale of its "participation" (what does it do, for what purpose?). Acknowledging the influence of the rationales and perspectives of the actors on the matter, and the difficulty for the lay citizens to engage with open data, I studied the initiatives claiming to be participatory in cities and involving the citizens as third parties. The first study was guided by the following research questions:

- *What are the innovation approaches taken by participatory data-driven initiatives in urban settings?*
- *What can be the citizens' roles and patterns per innovation approach?*

The contribution aims at understanding the innovation approaches at stake in the open data initiatives, the role of the "citizens" in the value chain and the type of value resulting from it. I observe, therefore, the innovation as a process and an outcome.

**The second research gap** focuses on the dual role of the municipality, being a public actor and holding a role in the data release and reuse. We know that the roles in the data value chain are

interchangeable, and that the public actor can be a publisher and re-user at once. However, no previous study has tried to understand how the first role of the actor (municipality) influences the way the actor undertakes its second role (publisher, re-user) and creates expectations towards the re-users or end-users to whom its input or output is addressed. In the second study, I examine the way the municipalities deliver data, portals, related activities and incentives, or data-based solutions, to create categories of roles for the municipalities and their expected users (a term encompassing the re-user of data and end-users). The research questions guiding the study are:

- *What are the possible municipalities' roles within OGD release and reuse?*
- *What expected users' roles are implied by the municipalities' roles?*

The contribution aims at clarifying the roles and expectations, and helping municipalities to understand how their role choice calls for a certain type of user that cannot be generalised as a "citizen". The study highlights the possible role-related issues that could impede the realisation of the outcomes. In this study, I see innovation as a process and an outcome, and focus on the role(s) of the municipality, at the beginning of the data value chain.

The third research gap focuses on the end-user. We know little about the end-users' needs, who do not handle data directly. Previous research tends to study the publishers and re-users in silo (Ham et al., 2019), trying to remove their respective barriers and understand the mechanisms of adoption and resistance (Martin, 2014; Smith & Sandberg, 2018; Wang & Lo, 2016; Janssen et al., 2012; Zuiderwijk et al., 2012; Wirtz, Piehler, Thomas, & Daiser, 2016). This research stream builds a well-needed knowledge regarding the publishers and re-users' requirements and challenges, but it misses the end-users' role who is, in the logic of a value chain, supposed to benefit from the sum of the value created in the chain.

Despite the huge scientific literature covering the methods and tools to innovate and involve the end-users in innovation processes (see Section 2.1.1.), research about end-user involvement in open data is scarce. In practice, by the publishers' side, little attention is given to the "how-to": how to solve societal problems with open data, how to create public value, how to generate benefits for the ordinary citizens (Zuiderwijk & Janssen, 2014).

Arguing that the end-users do not need data but answers and solutions to their daily problems, I tried to understand how to capture these information needs in the context of open data. The study's contribution is the development and implementation of a method, from which I retrieve design principles (prescriptive guidelines) to help the publishers and re-users to capture the end-users' needs and better align their work to them. The research question is:

- *What design principles could data providers and intermediaries follow to design methods to identify information needs of groups of users?*

In this study, data provider refers to the publisher, intermediary refers to the re-user and user (of solutions) refers to the end-user. The study focuses on the innovation process and is grounded in both innovation (service design, user involvement methods) and information systems (information needs) literatures, to approach a problem that was previously mostly discussed within the fields of information systems, computer sciences, and e-government.

# Chapter 3. Research Design

This chapter is structured as follows (Figure 10). I first introduce the research paradigm I follow, explaining the epistemological, ontological, and axiological assumptions and their impact on my research (Section 3.1.). Thereafter, I present and justify my general research approach (Section 3.2.), wherein I develop the principles of abduction and Design Science Research I applied in my qualitative studies (Section 3.2.1.). I tell my research journey and knowledge development: the way I developed the research questions, the role of the collaborative papers, and my publication strategy (Section 3.2.2.). It explains the multiple theoretical lenses used through the studies (Section 3.2.3). I finish this general overview with the empirical context of the studies and justify the choice for the national cases (Section 3.2.4.). Then, I go into the specific methodological choices and strategies per study to emphasise their consistency with my research paradigm and approach (Section 3.3.). For more details, the reader can read the method sections in the studies (Chapters 4, 5, and 6). Finally, I discuss the limitations due to the methodological choices (Section 3.4.).

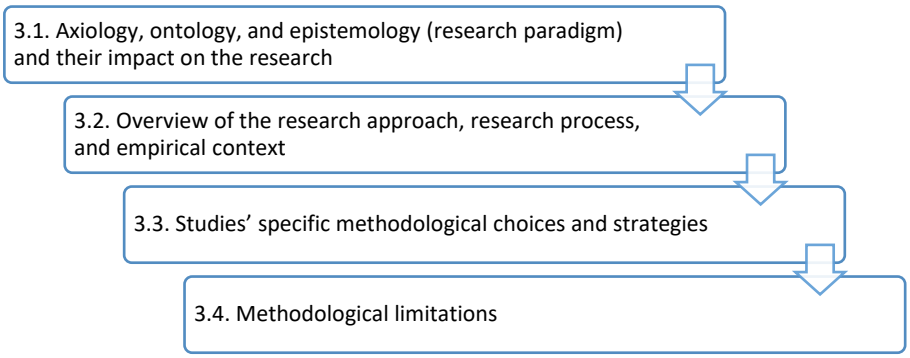


*Figure 10. Structure of Chapter 3*

## 3.1. Axiology, ontology, and epistemology (research paradigm) and their impact on the research

Before starting the research, I asked myself about my view on what knowledge building is and what research paradigm fits with my beliefs.

A research paradigm refers to a system of beliefs and assumptions about the development and nature of knowledge (Saunders & Lewis, 2018). The assumptions are axiological, ontological, and epistemological.

*Axiological assumptions* are about the importance and role of value and ethics within the research process. It means, for example, how we value, as a researcher, personal interactions, which can influence us to collect data via interviews or questionnaires rather than statistics for example (Saunders & Lewis, 2018). They are about the nature of ethics (Mertens, 2019).

*Ontological assumptions* are about the nature of reality, the way we see the world and define reality (Saunders & Lewis, 2018).

*Epistemological assumptions* are about the nature of knowledge, what constitutes acceptable, valid, and legitimate knowledge, and how we create and communicate knowledge (Saunders & Lewis, 2018). They are about the relationship between the knowing subject (understand respondent or researcher) and the "would-be known" (Mertens, 2019). They are shared by a community of research and rely on a consistent set of hypotheses (Avenier & Gavard-Perret, 2012).

Reflecting on our research paradigm is, therefore, essential to be aware of the roots of our inclinations for one or another research approach, choose an appropriate research method, and stay consistent in the way we assess our work to build acceptable, valid, and legitimate knowledge for a research community. Several research paradigms exist, the way they are called and their importance may vary between disciplines. Creswell and Creswell (2018) suggest four modern paradigms (Post-positivism, Constructivism, Transformative, Pragmatism). Specifically, in the discipline of management sciences, Saunders and Lewis (2018) suggest five current paradigms (Positivism, Critical Realism, Interpretivism, Post Modernism, and Pragmatism), and Avenier and Gavard-Perret (2012) suggest six of them (Positivism, Post-positivism including Scientific Realism and Critical Realism, Pragmatic Constructivism, Interpretivism, and Constructivism as defined by Guba and Lincoln). I will not discuss each of them in this thesis. Instead, I elaborate on the paradigm identified as relevant to my views in contrast with others and develop my understanding of it and its impact on my research.

### 3.1.1. Researcher´s background and assumptions

I have a background in management and innovation practices, with previous work experience in training and advising people on how to collaborate, learn, innovate, and grow as teams using design-thinking tools. I have been giving training and workshops for more than six years in various contexts (companies, citizen participation projects, not-for-profit associations), and took myself several lifelong learning training programs about teamwork and new management methods (e.g., sociocracy and participatory management). I believe in the power of the group and collaboration to solve problems, make balanced decisions, and develop better projects. I enjoy exploring and understanding people's perspectives, designing workshops and activities, and evaluating their impact on the group dynamics and participants' learning process.

I acknowledge and believe that I am influenced by my background and previous experience. I have discovered, experienced, tried out, and started to trust in the collaborative management tools and democratic values by being myself part of facilitated workshops and making decisions in councils in the scout association where I have been a volunteer for more than six years. This experience influenced my work choices and keen interest to learn more about participation and collaboration in theory and practice. I value working with people and collaboratively solving problems, with the hope to have an impact and be part of the change.

Consequently, this interest influenced the scientific literature I have read and the approach I have taken in my research projects. I do not believe that there is an independent, observable, reality (in a positivist way), but that the reality as we see it is shaped by our cultural background, education, and values. I think researchers are interpreters of reality, and that their contribution can never, in the absolute, be purely objective. Knowing that, I identified myself in a specific research paradigm.

### 3.1.2. Pragmatic constructivism: axiological, ontological, and epistemological assumptions

I place my research in the research paradigm of *pragmatic constructivism* as defined by Avenier and Gavard-Perret (2012). It should not be confused with the pragmatic constructivism of L. Nørreklit (2013), which is defended as a paradigm by its authors (H. Nørreklit, Nørreklit, & Mitchell, 2010) but is more used as a theoretical lens (analytical framework) especially in the field of accounting (e.g., Mattimoe & Seal, 2011; H. Nørreklit, Raffnsoe-Moller, & Mitchell, 2016).

According to Avenier and Gavard-Perret (2012), in social sciences, there are two constructivist paradigms: constructivism as defined by Guba and Lincoln and pragmatic constructivism. The first, as its name indicates, was developed by Guba and Lincoln (Guba & Lincoln, 1989, 1998). The second is anchored in the constructivism of Piaget, further developed by von Glasersfeld (1988, 2001) and Le Moigne (1995). The first paradigm is close to the interpretivist paradigm. Like interpretivism, it has an ontological hypothesis that stipulates that reality is relative, multiple, socially constructed, and not governed by natural laws (unlike positivism). The reality is said objective when intersubjective, which means that reality is said objective *if* the participants involved in a situation share a common interpretation and understanding of the experienced situation. On the contrary, the second constructivist paradigm, the pragmatic, neither has ontological hypotheses nor denies the possibility of an objective reality independent of the researcher (positivism) (Avenier & Gavard-Perret, 2012). Goldkuhl (2012) argues that the essence of a pragmatist ontology lies in actions and changes. He quotes Blumer (1969, p.71) to stress his ideas, who claimed that *"the essence of society lies in an ongoing process of action – not in a posited structure of relations. Without action, any structure of relations between people is meaningless. To be understood, a society must be seen and grasped in terms of the action that comprises it".*

In fact, the "pragmatic" paradigm is more concerned about how to solve a research problem and find applicable solutions, than following a determined philosophy of research. The axiological assumptions (ethics of research) focus on gaining knowledge in the pursuit of desired ends, which are influenced by the researcher's values and politics (Mertens, 2019). The researcher can use the method that best fits the needs and purpose of the research question (Creswell & Creswell, 2018). The pragmatic constructivism was first called "radical" constructivism by von Glasersfeld (1988), which created confusion among scholars. The term radical has been interpreted as if constructivism denied reality, when, in fact, it denies that we can rationally know a reality beyond our experience (Avenier & Thomas, 2015).

To present the constructivist paradigm in more detail, I summarise its **three founding hypotheses**, as suggested by Avenier and Gavard-Perret (2012). I supplement them with Morgan (2007), Goldkuhl (2012), and Mertens (2019).

**First**, ontologically, the constructivist paradigm postulates that what is knowable is the humans' experience of resistance to their actions (Avenier & Gavard-Perret, 2012). In other words, humans understand the world through their experience of it, which manifests itself through the resistance perceived by oneself to the actions we take. Each human knows *his own* experience of *one* reality. To make changes in the world, humans are driven by reason (purpose and knowledge) and actions. Therefore, human knowledge and actions are inseparable (Goldkuhl, 2012). There are two consequences to this ontological position: (1) no one can rationally claim to know "*the*" reality beyond one's perceptions (von Glasersfeld, 2001) and (2)

the hypothesis that states that "*the*" reality, an objective reality independent of the researcher, exists, cannot be proven (Avenier & Gavard-Perret, 2012). However, as mentioned earlier, pragmatists *do not deny* that things and events can exist independently of any observers. Instead, they take a middle position between the positivist and interpretivist ontologies (Goldkuhl, 2012). The nuance is that the pragmatic constructivist paradigm emphasises that reasons and thoughts create our perception of the world. It *only* posits that a flow of human experiences exists. This leads the theoreticians to let ontological assumptions aside. Goldkuhl (2012) calls this ontological position "symbolic realism": "realism" for the obvious realistic stance towards the external world, and "symbolic" for the meaning-orientation of knowledge, which role is to be useful for action and change.

The second hypothesis, epistemological, postulates that in knowledge building there is interdependence between the knowing subject and the object of study: what pertains to the reality being studied is inextricably built with what relates to the knowing subject. As a result, the knowledge that a researcher develops depends on the researcher, his knowledge project, background, personal history, previous knowledge, etc. (Avenier & Gavard-Perret, 2012). I understand that "the knowing subject", using the words of the authors (Avenier & Gavard-Perret, 2012), means the person (respondent or researcher) whose personal perception and understanding of an object under investigation are based on his previous knowledge and developing knowledge about it (in this context, knowing does not mean "knowing beforehand").

The third hypothesis is known as the teleological hypothesis. This hypothesis consists of recognising the intentionality or the finality of the knowing subject as the motor of the process of the construction of knowledge. The constructivist paradigm postulates that the project of knowing a certain reality (the research purpose) influences the way we experience it, therefore, the knowledge we develop (Avenier & Gavard-Perret, 2012). It is important to be aware of one's perspective on the subject as it shapes the research purpose. For example, a researcher convinced about the value of opening data could have the purpose to solve the lack of adoption by the citizens, while a critical researcher could have the purpose to demonstrate that is it worthless.

Morgan (2007) relates the two last hypotheses to axiology and argues that what shape the researcher's goals are his personality and intentions (who we are and how we act), which are influenced by his values and politics. Accordingly, in a pragmatic paradigm, the researcher recognises that his values, politics, and purpose (where he comes from) direct his attention to factors that have the most impact on *what* he has chosen to study and *how* he has chosen to do so (object, method, purpose of the research output).

The knowledge resulting from this paradigm aims at developing intelligibility in the flow of human experiences. In other words, the research output expresses how the researcher understands how reality works and is *one* representation of reality built by one person or group of persons. The findings and knowledge created are expressed in the form of symbolic constructions: representations, models, or frameworks (Avenier & Gavard-Perret, 2012). It is however not restricted to explanation and understanding, knowledge can be prescriptive (guidelines), normative (exhibiting values), and prospective (suggesting possibilities, how the world could work) (Goldkuhl, 2012). The criteria for legitimizing this knowledge are its functional adaptation (suitability to the concerned actors) and its viability to travel the world (adaptability to other contexts) (Avenier & Gavard-Perret, 2012). Thus, this knowledge should

offer the concerned actors' viable insights for *acting intentionally* in relation to the studied phenomenon. The main concern is to solve the research question and develop suitable solutions.

The paradigm has the following **methodological implications**. Any previous knowledge and method can be used to create new knowledge, as long as the researcher knows and can prove how it has been legitimized (e.g., the validity of the construct and the process based on previous research, published methods, consistency of the method with the research question, rigour) (Avenier & Gavard-Perret, 2012). The researcher builds on previous knowledge and can choose the method or combination of methods that best fit the research purpose and empirical situation (Goldkuhl, 2012). Nevertheless, he has the duty to show why the methods and previous knowledge are suitable to the situation and how they can help in the intervention (search for applicable solutions) (Avenier & Gavard-Perret, 2012). Despite the paradigm being called "constructivist", the knowledge creation is not limited to methods involving the co-construction of the knowledge with the participants (like the constructivism of Guba and Lincoln and to some extent, interpretivism). Qualitative methods (e.g., case studies, action research) play a major role, and techniques such as the Grounded Theory (Glaser & Strauss, 1967) are frequently used to create knowledge. Another research approach that is said to be appropriate to the constructivist paradigm is design science research (Goldkuhl, 2012).

Finally, in terms of the **connection between data and theory development**, the knowledge created in the constructivist paradigm is not generalised by replication, but by conceptual generalisation from induction or abduction (Morgan, 2007; Avenier & Gavard-Perret, 2012), which is consistent with the Design Science approach (Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007).

### 3.1.3. Influence of the research paradigm on the research project: choices of research approaches and alternatives

The pragmatic constructivist paradigm had an influence on my overall research project, as I show in Table 7.

It determined the research approach I chose for my studies.

*Table 7. Influence of the research paradigm and its assumptions on the studies*

| | **Pragmatic constructivism** (based on Avenier and Thomas, 2015; Goldkuhl, 2012; and Mertens, 2019) | **Overall impact on my research** (in terms of studies' purposes, theoretical frameworks, methods, and research outputs) |
|---|---|---|
| **Axiological assumptions** (Teleological hypothesis) | Gain knowledge in pursuit of desired ends, influenced by the researcher's values and politics. | I value interactions with the respondents (inclinations for methods that imply interviews, participatory workshops).<br>I aim at developing research outputs that have an impact on their intended audience. |
| **Ontological assumptions** | (1) Symbolic realism: Understanding of one reality in terms of actions, actors, artefacts, and social constructs.<br>(2) No denying of one or another research paradigm. | (1) I take a process perspective (a data value chain framework) to unravel roles, tasks, and actors' underlying assumptions.<br>(2) I combine research approaches and methods: application of abduction and Design Science Research. |
| **Epistemological assumptions** | (1) The human experience is knowable through actions.<br>(2) There is interdependence between the researcher and the study: the project of knowing and the background of the researcher (values, politics), influence the knowledge produced (which is one reality for one group of people). | (1) I have an inclination for methods that involve interactions and instantiations with the participants (Design Science Research)<br>(2) I have an inclination for research topics such as (citizen) participation, collaboration, social value, and impact, due to my background, personal history, and preferences. I use member-checks and popular science articles to evaluate my research output, without pretending they describe "the" reality. |
| **Role and shape of knowledge** | (1) Reach intelligibility in the flow of actions.<br>(2) Be useful for action (pragmatic): build models, guidelines, that provide insights for organising the world, the human experiences, and change a situation. | (1) The knowledge contribution aims at unravelling the roles, tasks, and output along an open data value chain, a representation of a flow of value-generating activities.<br>(2) The research output is delivered through the development of models, tools (typology), and methods (design principles), actionable by the intended audience. |
| **Role of the researcher** | Engaged in change, through the creation of data (experimentation) and theories. | Development of tools, frameworks and principles that are, when possible, developed, tested, and evaluated with their intended audience. |

Therefore, in accordance with the research paradigm and my research purpose, I opted for two complementary research approaches to answer my research questions:

- A **qualitative abductive research approach** to understand phenomena through the organisations of individuals' activities and their world's views, going back and forth between empirical evidence and previous literature. Specifically, I conducted a multiple cases study (Study 1) with an abductive approach.

- The **Design Science Research (DSR)** to develop usable artefacts relevant for practice and research, in an iterative research approach (Study 2 and 3).

I motivate these choices as I believed they are consistent with pragmatic constructivism and assessed as the more suitable to answer my questions in my empirical context (methodological implications), as far I as knew at the time of designing my methods (epistemological hypothesis).

Despite **Action Research** is mentioned as **a consistent alternative** research approach within the pragmatic constructivist paradigm (Avenier & Gavard-Perret, 2012), I did not follow this path due to the difficult access to the field (see Sections 3.3. and 3.4.) and my research purpose. Following a principle of "change through action", Action Research is highly context-dependant and implies some concrete "client" or partner involved to analyse a situation. Generalisable knowledge is produced by making sense of reality and solving local problems (Iivari & Venable, 2009). In my case, close collaboration with practitioners was difficult due to the low maturity of open data in Belgium (see Sections 3.2.4. and 3.3.), the covid-19 Pandemic and governmental measures ("stay at home"), and limited connections in the field due to my recent interest for the topic. Unlike Action Research, DSR also allows to solve a class of problems faced by a class of people or organisations in a way that is relevant and useful to fix specific problems for specific actors (further explained in Section 3.2.1.). I understand the misalignments as a class of problems. DSR does not assume any specific client or initial joint collaboration: it is focused on building innovative artefacts, creating new solutions, instead of making sense of the existing (Iivari & Venable, 2009). The researchers can conduct the development, tests, and evaluations as long as mechanisms to check the relevance for the audience are guaranteed all along the research process (e.g., setting goals, evaluating against goals' criteria; Hevner, March, Park, Ram, & Ram, 2004). Hence, I found DSR more appropriate and realistic in my context.

In the following subsections, I elaborate on the reasoning behind the two chosen research approaches: qualitative abductive research and Design Science Research, their development and use through my research journey (Section 3.2.), and their specific application and evaluation against other alternatives per study (Section 3.3.).

## 3.2. Overview of the research approach, research process, and empirical context

### 3.2.1. General research approach to theory development: abduction as a transversal approach

There are different approaches to theory development: deductive, inductive, and abductive. As presented in Figure 11, deduction involves the testing of the theoretical propositions by using a research strategy specifically designed to collect data for the purpose of its testing. That research aims for explanation and prediction and is typically associated with quantitative methods. The research starts with an extensive literature review to clarify a theory and draw hypotheses to test. Induction takes the opposite way: the research starts by analysing data already collected, from specific observations to generalisation and theories (Saunders & Lewis, 2018). It aims for interpretation and understanding, typically associated with qualitative methods (e.g., interpretivism, constructivism of Guba and Lincoln, Grounded Theory).

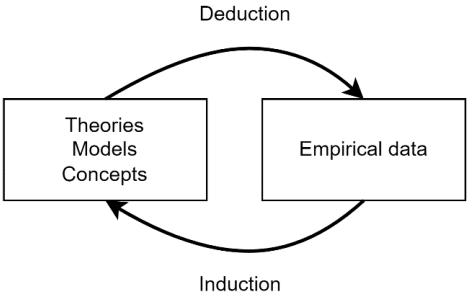


*Figure 11. The induction and deduction processes*

Finally, abduction moves back and forth, from theory to data and from data to theory. This approach aims at combining and leaving behind the long rivalry that opposes positivists and interpretivists, quantitative and qualitative methods. An abductive approach assumes that both can be combined to build valid knowledge. Morgan (2007) associates it to the pragmatic constructivism paradigm and does not limit it to the use of mixed methods in single studies. Qualitative researchers can use knowledge and hypotheses produced by quantitative researchers to explore a phenomenon.

I felt quite comfortable with abduction since the beginning of my research process and exploration of methods. I think it is more balanced than the constructivism as defined by Guba and Lincoln, which in my opinion, can lead the researcher to paths already explored in previous research, and in that sense, be inefficient. It is, nevertheless, grounded in empirical data and humans' activities, which matters in a pragmatic paradigm. As a result, I applied the principles of abduction in qualitative research (Study 1) and in Design Science Research (Study 2 and 3). I present the guiding principles of these research approaches in the following subsections.

### *Abduction in qualitative research*

An abductive analysis of qualitative data is a form of reasoning through which we perceive the phenomenon as related to others observations (Timmermans & Tavory, 2012). The phenomenon is seen as similar to other phenomena already experienced by the researcher or explained in other situations and previous research. Timmermans and Tavory (2012) explain that abduction is an inferential creative process of producing new hypotheses and theories based on surprising research evidence. Unlike the Grounded Theory, which argues that the researchers should begin with an inductive analysis of data, an abductive analysis recognises the role of the theoretical background of a researcher and relies, for a large part, on its scope and sophistication. The more the researcher is sensitized to theories and develops his theoretical repertoire throughout the research process, the more likely he will recognise the potential relevance of surprising pieces of evidence. The abductive analysis aims at generating novel theoretical insights that reframe empirical findings in contrast to existing theories. Theories and previous knowledge are involved at the start and during the research process. However, the abductive qualitative analysis still uses methodological tools from the Grounded Theory, such as detailed field notes, transcriptions, coding, and memo writing (Timmermans & Tavory, 2012).

Figure 12 illustrates an abductive research process based on Timmermans and Tavory (2012).

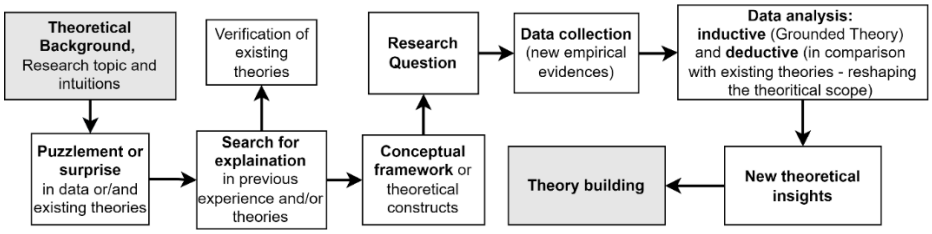


*Figure 12. An abductive research process (based on Timmerman and Tavory, 2012)*

I specifically applied abduction to my first study to understand the phenomenon of citizen participation in open data initiatives.

### *Design Science Research*

This approach is inspired by design. Design can be defined as both a process (a verb, a set of activities) and a product (a noun, an artefact resulting from the process). It can, amongst others, refer to routine development, a mindset (Design Thinking), or a research approach, Design Science Research (DSR). DSR is an approach of research that has its roots in engineering and computer sciences and was specially developed and theorised within the information system discipline. Information system is known to have a pragmatic approach to research, as the research output aims at solving business problems, and is known to have a multi-paradigmatic research community, as multiple means to conduct research and develop knowledge are accepted (Vaishnavi, Kuechler, & Petter, 2004). DSR can be anchored in a pragmatic paradigm (Goldkuhl, 2012).

DSR addresses **knowledge creation** through the building and the evaluation of **artefacts** designed to meet identified needs. Its **goal is *utility*** (Hevner et al., 2004). Based on the observed needs for practice and problems of interest to the research community, DSR can support the development, construction, and evaluation of artefacts that enable the transformation of situations by changing their conditions to better or desirable states (March & Smith, 1995; March & Storey, 2008). DSR integrates the prospective, prescriptive and normative aspects of knowledge (Goldkuhl, 2012). What makes the difference between routine design and DSR is the ***novelty of the artefact***: to contribute to knowledge, it should provide new solutions for known problems (improvement), extend known solutions (e.g., from other fields) to new problems (exaptation), or find new solutions for new problems (inventions) (Gregor & Hevner, 2013).

Specifically, the **artefacts** developed within DSR can be classified constructs (e.g., vocabulary, symbols), models (e.g., abstractions, representations), methods (e.g., algorithms, practices, governance strategies), and instantiations (e.g., implemented decision-support systems, prototypes) (March & Smith, 1995). In terms of **theory development**, Gregor and Hevner (2013) identify three levels of contribution. The more the knowledge is abstract and generalised, the more it contributes to design theories. The implemented artefacts (instantiations as situated implementations) are at the lowest level of knowledge contribution: they are specific, limited, and less mature. Nascent design theories, such as methods, models, design principles, or technological rules contribute to operational knowledge. Design theories are the most developed and mature knowledge contributions that can be made using DSR (Gregor & Hevner,

2013). Design theories are prescriptive knowledge: they say how to do something (Gregor & Hevner, 2013).

A DSR approach requires creativity and trial and error: the **research process** is not straight but iterative. Moreover, the research process is based on two types of knowledge that the researcher must explore and use to prove the novelty of his artefact:

- descriptive knowledge (what we know about the phenomenon and its environment),
- and prescriptive knowledge (how the phenomenon is or can be addressed, how artefacts are built and can be evaluated).

Accordingly, DSR relies on previous research but is not purely deductive or inductive (Iivari, 2015): the researcher can go back and forth in previous research between each iteration, in an abductive way. Deduction comes later, in the development and evaluation phases (Vaishnavi et al., 2004).

The utility principle of the knowledge created is applied to the environment and audience of the artefact, and the research community it contributes to. It means in the research process that high importance is given to the **evaluation and justification** of the knowledge created. This is done in respect with the scientific rigour of the process (existing knowledge and methodologies that were built upon and used, validity) and the relevance of the artefact (utility, quality, efficacy, and appropriateness of the solution to the problem domain, identified needs, utility and value outside the development environment) (Hevner et al., 2004; Gregor & Hevner, 2013). Several methods of evaluation can be used (see e.g., Hevner et al., 2004, p. 86).

Figure 13 presents the two main research cycles of a DSR approach as suggested by Hevner (2007).

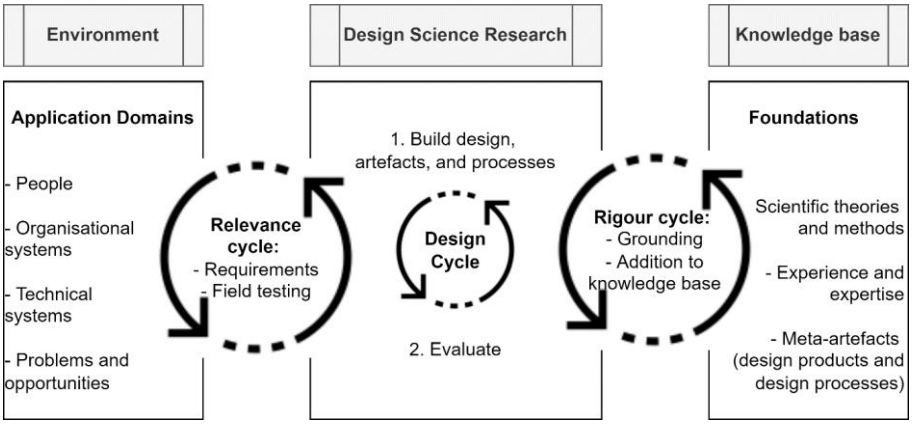


*Figure 13. Design Science Research Cycles (adapted from Hevner, 2007)*

To **implement DSR**, Peffers et al. (2007) recommend the following six iterative steps, as shown in Figure 14: (1) problem identification and motivation, (2) definition of the objectives for a solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication. Each iteration can go back and continue from any step of the process.

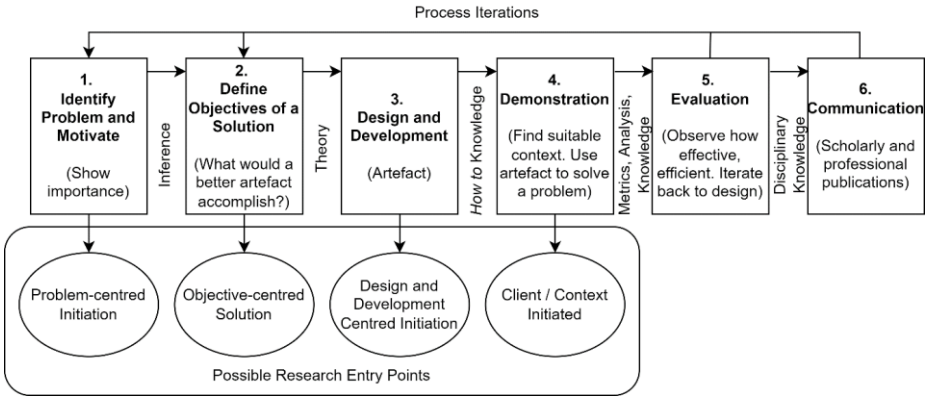


*Figure 14. Design Science Research Process Model (adapted from Peffers et al. 2007, p. 54)*

The next section situates when I used these approaches through my research journey. I highlight the conditions and the moments of puzzlement that initiated the three main papers, in reference to the abductive research process. In this puzzlement, the collaborative papers I was invited to join played an import role. I summarise the development of my ideas and iterations of papers, from posters to journals. I also explain the publication strategy and positioning choices that made me navigate between my initial field and the open data field.

### 3.2.2. My journey in open data research: knowledge development, role of the collaborations, and publication strategy

As mentioned in the Introduction, I first encountered open data through my readings about Smart Cities, the use of data in cities, and citizen participation with a background in management. Sharing my interest in open data to a colleague from the Computer Science faculty of Namur, I was invited for a **first collaboration** in a paper about open data reuse and re-users' impediments led by Jonathan Crusoe from Linköping University (Crusoe, Simonofski, Clarinval, & Gebka, 2019). We studied the impediments faced by the re-users, in our case, students at the University of Namur developing solutions for cities based on their available open data. For this project, I collected the data by interviews and analysed it, I co-wrote the method and result sections, discussion and conclusion, and reviewed the whole paper before submission. It was accepted and published at the IEEE Thirteen International Conference on Research Challenges in Information Science, held in May 2019 in Brussels. This work helped my research at the initial stage: I explored the subject, touched upon the technical side of open data, and got puzzled by some of the findings. It also opened me to the research community of my co-authors, Information Systems, where open data is studied at the intersection between information, management, and computer sciences.

I realised through the data collection and analysis the technical complexity of reusing open data, even for skilled users. Uneducated people in this field simply cannot use it. Another very interesting insight was the students reporting their frustrations and lack of ideas for valuable solutions using open data. They did not know about the cities' current challenges and priorities, citizens' needs, and market opportunities. Students who came with ideas projected their own needs on the "citizens", but then, they had to adapt them to the available data. Others brainstormed based on the available data, which means that their solution was not based on needs but data opportunities (Crusoe et al., 2019).

This finding drew my attention on two problems. (1) Given the complexity of reusing data, the re-users, here students, are not "any" lay citizens. Therefore, what is to be a "participant" in an open data-driven initiative? This reflection initiated Study 1. (2) The re-users might lack ideas and misunderstand the citizens' needs, in other words, the end-users' needs. How could we reconnect the re-users with the end-users' needs and involve them in the ideation stage? It initiated Study 3[5].

Given my background and the literature on innovation and user involvement, I assumed that the data skills should not be a limitation to participating in other ways and other stages of an open data solution development. I used this assumption and previous knowledge to focus on the citizens and their roles in open data initiatives in Study 1, and the methods to involve them and capture their needs in Study 3. I initially positioned these studies in innovation management, my field.

I started the data collection of **Study 1,** using a multiple cases study in an abductive qualitative research, in early 2019. I submitted an ongoing research paper at ISPIM (conference of the International Society for Professional Innovation Management) in June 2019. I targeted this conference for being a leading conference in the community of innovation management, close to my initial field and the research centre where I am affiliated. It was well received and got useful feedback. I finalised the first version of the paper for submission to in early 2020. The paper was rejected for being perceived as not fitting with the journal's scope and needing revision. I worked on a new version and submitted in early 2021 at a journal that focuses on research that aim at maximizing the benefits and minimize the negatives effects of technologies, with a specific focus on urban environments, Smart Cities, and a general audience of practitioners in this field. It is currently going through the first round of revision.

The first ideas of **Study 3** were submitted in 2019 as a poster at the RCIS conference, taking the opportunity because the paper of collaboration 1 was accepted as full research. To involve citizens and stimulate data reuse, I did not want to follow the models of the hackathons. I knew by previous research and by experience (I volunteered to be a jury member for the Hackathon of Futurocité in October 2019) that these popular events were attracting young developers facing the same difficulties we heard from the students in the collaboration 1 (Crusoe et al., 2019). Instead, with the co-authors, we organised workshops from October 2019 to test other methods to capture information needs in cities' everyday contexts. Because the intended research output was a method, we applied DSR. Encouraged by my co-authors, I sent an early version of the paper in January 2020 at a scientific workshop: the 17th Scandinavian Workshop on E-Government, SWEG 2020. This workshop gathers researchers from the eGovernment community and helps them to prepare their submission for the EGOV (Electronic Government) conference whose deadline is in March. The eGovernment research community is interested in topics such as open government and participation, electronic services, information and infrastructures, among others. Open data, in that community, is a growing topic (Scholl, 2014). This move took me away from the innovation management community but brought me closer to relevant feedback on issues surrounding open data, such as the public sector, citizen

[5] The reader would notice that the studies´ numbers do not follow the chronology of the research process. I switched them around in the thesis because Study 1 and 2 address a common misalignment and both contribute to the conceptualisation of roles.

participation, and information systems. The improved paper was submitted at the EGOV conference 2020.

During that time, we used the data collected at the first workshops to dig deeper into the concept of information needs and submitted this twin paper, my **second collaboration**, to EGOV (Crusoe, Gebka, & Ahlin, 2020). My contribution consisted of coordinating the whole research project, finding literature and co-writing the background, collecting and analysing the data, co-writing the method, and revising the whole paper before submission. We drew a tentative model showing the relations between the actors (publishers, re-users, end-users), the concepts at stake (primary needs, information needs, solutions and information sources), and the considerations to have in mind when trying to make the whole system aligned. This work contributed to my understanding of the information needs and the relations and requirements between the actors at an abstract level. It helped to refine the design principles of Study 3 between the iterations.

The paper of the second collaboration (Crusoe et al., 2020) was accepted as full research, while Study 3, that I was leading, was accepted as ongoing research. A collaboration with a partner, Futurocité, allowed us to run more iterations in a new data collection round in October 2020. The final version of the paper was submitted in February 2021 at the *Data Science Journal* and accepted in November 2021 after minor revisions. This outlet was chosen after discussion with my co-authors. I reviewed the scope of several journals focusing on data valorisation and we decided on this one for two reasons: it focuses on the availability and use of data, especially the principles, policies and practices for open data, and last but not least, it is available in open access. This especially interested us, as we were addressing the Design Principles to a community of practitioners.

During the evaluation of the design principles with two small Belgian municipalities, organised for Study 3 in January 2020, the research question of **Study 2** came. It struck me that the municipalities' employees and town councillors were confused about the scope of their role as publishers and public actors. I decided to investigate how the municipality's role shape the role of the citizens and develop a typology, using DSR. I collected data to submit the first ongoing research paper at ISPIM 2020 as, in continuation of Study 1, I saw the roles in the perspective of open data innovation. Like for my first submission at ISPIM, I had good discussions with researchers of the community and was suggested to use Role Theory, a set of theories popular in innovation process management. Using this theory was a very smart move. I completed the data collection in the autumn of 2020 and discussed the paper at SWEG (the 18th Scandinavian Workshop on E-Government, SWEG 2021) intending to publish it at EGOV the same year. I thought that, given the strong focus on the public actors and the implications of my findings, the eGovernment community was the best audience for this paper. It was accepted as a full-research paper at EGOV 2021, without revision. This last paper was also rewarded as *"The most innovative research contribution or case study. Awards the paper with the most out-of-the-box and forward-looking idea and concept",* thanks to the novel application of Role Theory on the field.

My research journey is represented as a timeline in Figure 15. The arrows represent the research process (data collection, analysis, writing paper) happening between each event (puzzlement and publication in outlets).

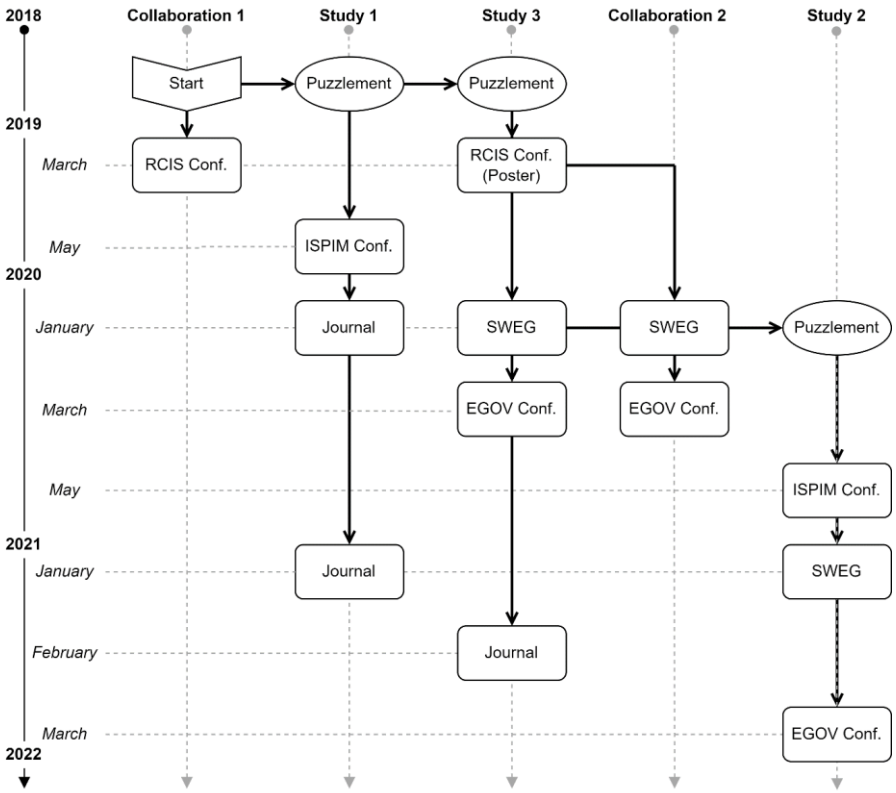


*Figure 15. The research process in a timeline*

Looking back at my research process and publication strategy, I would summarise it as follows. The role and the value chain were key concepts in my research from the beginning. In the third year, I conceptualised my thesis problem in terms of misalignments in the value chain. Influenced by my roots in innovation management, I look at open data through this lens and apply concepts and theories from my initial field to open data (i.e., value chain, Role Theory, Participatory Design methods). My approach wad double-edge when it comes to find the right publishing outlet. Either it makes it harder to find the right research community, as the paper can be perceived as speaking "too little", or not falling under the delimited scope of each of the intertwined fields (as experienced with Study 1 and 3). Or the paper can be surprisingly well-received and rewarded as innovative (as experienced with Study 2). At the end, I find my place within interdisciplinary research communities, such as the eGovernment community, and journals who assess and try to increase the value of new technologies for society and practice.

My knowledge development, iterative and abductive, is also reflected through the choice of theoretical lenses in my studies.

### 3.2.3. Plurality of theoretical lenses as a result of an abductive research process: justification and alternative unifying theories

Although the concept of role was central in my research from the beginning, I did not use a unique theoretical lens through the three main studies. Instead, I chose the theoretical lens I

thought was the most appropriate to each research question and study (further developed in Section 3.3.).

I started my research with a value chain perspective and tried to understand the problems along the process. I formulated my research problem in misalignments in the second half of my thesis journey and came across Role Theory lately, which eventually became significant in my discussion and conclusion. Once I have accumulated knowledge and better understood the topic, it is easy to think that I could have used Role Theory as a unifying theory from the beginning. However, research is an iterative and explorative process, and new alternatives, maybe best-fitting theories, can always come along the way.

For example, a theory I did not think of earlier but could be relevant to study the roles and their problems to reach a goal is **coordination theory.** Coordination theory is a set of theories concerned about the interdependencies between actors. Coordination is understood as *"the act of working together harmoniously"* achieved through the management of dependencies between actors (Malone & Crowston, 1990). The main components of coordination are the actors, the goals they want to achieve together, the required activities and the interdependencies (e.g., prerequisite, shared resources, or simultaneity) (Malone & Crowston, 1990). We can see parallels with the concepts I intuitively put together in my conceptual framework. Coordination Theory has been applied in several contexts at the task-level in computer interactions and organisational studies. One of the most famous theories and contribution in terms of coordination mechanisms is of Mintzberg (1980), and at the inter-organisational level, Demil & Lecocq (2006) identify the market, network, hierarchies, and bazaar as systems of coordination.

Coordination mechanisms are ways to manage the interdependencies, and it that way, could be an appropriate theoretical lens to solve certain misalignments, considering open data as an interdependency (the shared resource). Coordination theory has already been applied to open data to try to align offer and demand through infrastructure (Malone & Crowston, 1990), and "data collaboratives" (thematic open data ecosystems made of private and public actors; Susha, Janssen, & Verhulst, 2017).

To close the overview of the reasoning and circumstances that influenced the research design, the next section explains the choice for the cities as contexts, the selected national cases, and their specificities.

### 3.2.4. General empirical context: open data in cities and national cases

The three studies and two collaborative papers of this thesis have for context the cities, a choice that I justify for the following reasons.

#### Open data in cities

The main objectives (governance, participation, innovation) of open data are close to values and objectives making sense in cities. In democratic countries, all levels of governments are concerned by accountable governance, supported by transparency and participation. Especially at the local level, citizen participation is a historically valued mechanism in decision-making processes and urban planning. Citizen participation is part of the classic literature on cities (Krishna, Kummitha, & Crutzen, 2017). Cities are also incredible testbeds to try out the possibilities provided by ICTs as an attempt to renew and increase citizen participation in response to the democracy and trust crisis (see Section 2.1.2.). In parallel, two other leading

concepts combine the use of ICTs (including open data), and citizen participation in cities: the concept of open government (see Section 2.1.2.) and the concept of Smart City (see Introduction). The latter integrates the idea of improving the citizens' well-being with the use of ICTs and participation (Krishna et al., 2017). Open data has gained importance in the smart cities initiatives, to such a point that Ojo et al. (2015) claim that open data converge with smart cities. As a result, open data is becoming part of broader digital strategies in cities (Berends, Carrara, & Vollers, 2020). The municipalities are sitting on large amounts of data of social and economic value, which can feed the regional, national, and European portals (Berends et al., 2020). Cities have the potential to become leading publishers and are aware that open data can bring efficiencies and potentially more interaction with the citizens (Berends et al., 2020). Accordingly, this field is potentially rich in open data initiatives from where we can learn and build knowledge.

### National cases

The studies were carried out in three European countries. The origin of open data in the European context is succinctly presented in Section 2.1.2.). The rest of this section introduces specifically the national contexts of three countries: Belgium, Sweden, and France. They were selected for two reasons: (1) the access to the field and data (language, proximity, international collaboration opportunities), and (2) their complementary, since the three countries have different levels of maturity according to the European Maturity Report (van Hesteren, van Knippenberg, Weyzen, Huyer, & Cecconi, 2021). The report suggests four levels: beginners, followers, fast-trackers, and trend-setters. It measures open data maturity with metrics divided into four dimensions: open data policy, impact (among others, reuses), national portal, and data quality. I do not question the method used and the limits of the ranking: this discussion goes beyond the scope of my research. Instead, I used the ranking as a compass to navigate between the potential fields and chose purposefully interesting national cases. Despite the studies focusing on the local level, in municipalities, strategies and decisions made at the national level do have an influence on the local level. For example, the national governments have the authority to enforce a policy or a law applicable at the local level, and the national portals harvest the local ones, which is a good indicator of the local data quality and the reuses made possible by it. The national maturity provides thus insights into the status of open data in local initiatives. During the period of my thesis (2018-2022), France stayed stable in a trendsetter position, Sweden moved back and forth between follower (2018, 2019, and 2021) and fast-tracker (2020), while Belgium went down from fast-tracker (2018) to follower (2019-2020), and nowadays beginner (2021) due to its stagnation against the fast progression of other European countries.

#### *Belgium, the beginner*

In Belgium, open data is clearly encouraged by the federal government, following the incentive of the PSI Directive (European Parliament & European Council, 2013). Open data is one of the strategic priorities of the Belgian federal government since 2015 (Cabinet du Ministre de l'Agenda numérique, 2015). The law of 4 May 2016 on the reuse of public sector information (Conseil d'Etat & Chambre des représentants, 2016) was followed by the royal decree of 2 June 2019 (Conseil d'Etat, 2019), which specifies how to implement it. A federal strategy has been defined, accompanied by the establishment of a federal open data portal, which links the data of local authorities to the European data portal. A Task Force has been established to support the implementation of the strategy, support the administrations, and energize the open data community with events and calls for projects (e.g., the call for projects "Smart Mobility", 2019,

rewarding projects using and publishing open data, while solving public mobility issues). However, a challenge Belgium is facing is due to the political landscape: open data is also a competence of the regional governments, and each region (Flanders, Wallonia, and Brussels) is implementing it with its own way and pace. For example, Flanders, Wallonia, and Brussels have the equivalent of a law (Gouvernement de la Région de Bruxelles-Capitale, 2016; Ministère de la Communauté française, 2017; Vlaams Parlement, 2015), while only Brussels (Gouvernement de la Région de Bruxelles-Capitale, 2018) has an executive decree. It means that there are few practical guidelines for the Walloon and Flemish administrations regarding the implementation of the law, and no consequences if they do not.

Consequently, open data portals, considered as a means to facilitate open data release, are popping up at the federal, regional, and occasionally local (municipal) levels. In Flanders in 2018, an attempt to stimulate, coordinate and standardise open data release in municipalities was materialised by the creation of an Open Data Charter (Smart Flanders, 2018). It was signed by the 13 largest Flemish cities. In a survey I did in June 2021, only 8 out of 300 Flemish municipalities, all signatories of the charter, had published open data. Not all the signatories had published data and none of the non-signatories had. In the Brussels region, the legal framework specifies that municipalities should publish their data of subsidies and public procurement's inventories by the end of March 2021 for the last three years (Région de Bruxelles Capitale, 2021). I noticed a sudden increase in the data release on the regional platform at the end of the first trimester of 2021. Seven out of 19 municipalities published new datasets, which consists of only one to four datasets about subsidies for five of them. In Wallonia, the Digital public agency (Agence du Numérique), is still working on the structuration of the open data ecosystem, organising punctual thematic working groups and delegating the awareness, training, and incentive programs (e.g., hackathons) to other publicly funded agencies (especially FuturoCité). Only two big municipalities have a significant number of datasets on their own portals and the regional portal harvests isolated datasets from other municipalities and public institutions. According to my latest survey in June 2021, twelve out of 262 Walloon municipalities have published data. A few large cities, Gent, Brussels, Namur, and Liège, stand out with their own portal, a significant amount of data, and regular releases.

In civil society, the Open Knowledge Foundation is very active, especially in Brussels and Flanders. They organise a yearly practice and grass-root oriented conference (Open Belgium), meet-ups, events and projects to open more data and share knowledge. The IMEC, a Flemish research centre for nano and digital technologies drives a pool of research projects and researchers that, among others, focus on linked open data for decision-making in municipalities and collaborations with the Flemish government.

*Sweden, the follower*

The legislation on transparency has a long history in Sweden. The Freedom of the Press Act, which states that all the archives and official documents of governmental organisations should be made available to the citizens, was introduced in the constitution in 1766 (Safarov, 2019). The PSI directive is also influencing the release of data. It became law in Sweden in 2010 through the Reuse of Public Administration Documents Act (2010:566) and limits the conditions of reuse to foster a data market in Sweden.

The country is known for its high level of digitisation: it is second among the 28 European countries, according to the Digital Economy and Society Index (European Commission, 2020a). The index is based on indicators including broadband connectivity, digital skills, use of the

internet, digital public services, emerging technologies, the ICT sector and its R&D spending. The index reveals that Sweden comes second also when it comes to internet use and connectivity: most of the Swedes use the Internet daily for leisure (films, TV, music), online banking, and shopping. Sweden has also a strong human capital, since 72% of the population has at least basic digital skills and 46% has above basic digital skills. Digitisation is indeed important for the Swedish government. The government adopted a digitisation strategy in 2017 with the purpose to become the world's leader in making the most of the opportunities offered by digitisation and new technologies.

Interestingly, despite the public sector being digitally mature (European Commission, 2020a), the open data status in Sweden does not range among the leaders. The provision of data is uneven between sectors and municipalities, publishers have not fixed processes and standards to publish data, governmental activities promoting the reuse of data are few, and practitioners think that the number of public organisations working with open data is too low to establish structured public support and awareness (Safarov, 2019). Safarov (2019) and the European open data portal (European Commission, 2020b) point out the decentralized nature of the administration in Sweden as a challenge. Previously, the innovation agency, Vinnova, led the coordination of the effort to implement open data, amongst others by providing the first national data portal (oppnadata.se, renamed nowadays dataportal.se). That role was then delegated to the recently founded agency for digital government (2018), DIGG. It works to promote and make available open data from the public administrations and create a more efficient open data environment. Moreover, the government launched the Open Government Partnership Action Plan 2019–2021 (Sveriges regering - Infrastrukturdepartementet, 2019), a strategy dedicated to structuring and implementing open data nationwide with a comprehensive policy.

In July 2021, the national data portal provided 7415 datasets, from about 100 public organisations (agencies, museums, municipalities). Twenty-one municipalities are listed as publishers on that portal. Crusoe (2021) also identified 12 hackathons (one national), 11 data labs to foster innovation with data reuse, and 4 OGD events gathering both public and private sectors in 2018.

*France, the trend-setter*

The French government has voted a law in 2016 to become a "digital republic" (Assemblée Nationale & Sénat, 2016). The law is divided into three dimensions: data flow and dissemination of knowledge, protection of the citizens in a digital society, and accessibility to the digital solutions (digital divide, phone, wireless, and internet coverage), since then complemented by several decrees of application.

The first dimension is determinant for open data. It resulted, among others, in the creation of a public service for data (the idea that the public sector should make its data accessible), the notion of public interest data, and the will to facilitate the release and reuse of open government data considered as having an economic, social, health or environmental interest. The public data should be open by default by the municipalities above 3500 inhabitants and the public administrations above 50 employees. The purpose is to promote transparency, foster innovation, and promote an inclusive and open administration.

This ambitious legal framework facilitated the establishment of agencies, associations, and networks of data enthusiasts. Etalab is a ministerial department whose mission and structure

are made for coordinating the design and implementation of the national data strategy and the public service for data. They develop tools and services: coaching and advice for the administrations, national open data portal and other tools for APIs, promotion of data science and principles of open government, development of programs for public innovation (e.g., the concept of entrepreneurs of public interest, a kind of open innovation for administrations). Open Data France is an association that helps the administrations to open their data with working groups, methodological resources, open data principles, and coaching. It promotes also reuses and has developed an observatory of open data. The think-tank La FING has conducted research programs about the impact of open data. Datactivist is the private agency that advises, coaches, and trains the organisations with a leitmotiv: “release with purpose”. In civil society, the Open Knowledge Foundation is also present, and Team Open Data, a dynamic online community gathering more than 1000 publishers, re-users, and data enthusiasts. That vivid environment resulted in, according to the observatory of Open Data France, 598 local governments and public organisations having published open data in July 2021.

### 3.2.5. Studies’ overview

To conclude, Table 8 shows the result of the research approach choices, the coverage and articulation of the three studies around the research problems, stated in misalignments.

*Table 8. Research questions, research type and methods, studies' focus and empirical context to address the possible misalignments*

| Addressed misalignment(s) | Research questions | Research Type and *method* | Study's focus | Empirical context |
|---|---|---|---|---|
| **(1) Actors ↔ Roles** (actor's perception of other roles)<br><br>**(2) Actors ↔ Outcome** (municipalities' expected and realised outcome) | **Study 1:**<br>*(1) What are the innovation approaches taken by participatory data-driven initiatives in urban settings?*<br>*(2) What can be the citizens' roles and patterns per innovation approach?* | **Empirical study - Explanatory** (What - Why)<br><br>*Qualitative abductive research – multiple case studies* | Roles of the "citizens" as actors/participants in the value creation process (third parties) (what), according to the rationales of the initiatives' leaders (why). Type of value created (outcome) | Initiatives taking place in cities or addressed to citizens of Belgium and France |
| **(3) Roles ↔ Process** (Publisher's input >< users requirements)<br><br>**(1) Actors ↔ Roles** (actor's perception of other roles) | **Study 2:**<br>*What are the possible municipalities' roles within OGD release and reuse?*<br><br>*What expected users' roles are implied by the municipalities' roles?* | **Artefact building study - Descriptive** (What)<br><br>*Design Science Research* | Roles of the municipalities in open data, and related expected roles of the "users" | Municipalities of Belgium and Sweden |
| **(4) Roles ↔ Process** (publisher's input, re-user's output >< end-users' needs) | **Study 3:**<br>*What design principles could data providers and intermediaries follow to design methods to identify information needs of groups of users?* | **Artefact building study - Prescriptive** (How)<br><br>*Design Science Research* | Design principles to develop methods to re-align the End-users' needs with the Re-users' and Publishers' work. Roles as catalysts in the process. | Cities as a context of the scenarios, and citizens as roles of the participants |

In the next section, develop more in details the specific choices and circumstances of each study. I remind the studies' purpose and briefly present the challenges of the specific empirical contexts that constrained the methodological choices. I then explain the alternative theoretical lenses and methods I assessed, their strengths and weaknesses, and the reasons of my final choices. I also reflect on the data collection and analysis techniques used in the papers and the contribution of the studies to the main research problem.

## 3.3. Studies' methodological choices and strategies

### 3.3.1. Study 1

#### Research purpose and empirical context

The first study aimed to be explanatory. The purpose of an explanatory research is to understand how events occur and which one may influence a particular outcome (Hancock & Algozzine, 2006). I suspected that actors leading open data initiatives were driven by specific rationales (e.g., innovation, democratic values), and these rationales would define the way they see and involve the citizens. My study had the purpose to unfold the citizens' roles in data-driven initiatives taking place in cities, in relation to the leaders' rationales. It brings light on the causal link that exists between initiative leaders' rationales, citizens' roles and the type of value created by the participants.

#### Alternative methods and theory

At the early stage of my research project, having the purpose to understand the way citizens are involved in open data initiatives, I was interested in using Activity Theory. My initial idea was to use it in case studies to understand the roles of the citizens, including their motivations and barriers to participating in open data initiatives. Activity Theory is solution oriented. It suggests that activities are undertaken under certain conditions, are defined by an intention and driven by a purpose to reach an outcome, and are mediated by tools, instruments or artefacts. It has been applied in Human-Computer Interaction, project management, or education, with a problem-solving approach: it seeks to identify practices that can be improved or enhanced (McAvinia, 2016). Engeström (2000) has developed an activity system made of seven interacting nodes: the subject, the object (motivation for the activity), the outcome (transformation of the object), the rules, the community, the roles (division of labour) and the mediating instrument (artefact). Despite the lack of a precise method to apply it (McAvinia, 2016), the framework has the strength to direct the attention towards the contradictions between the nodes, especially between the shared and unshared objects and outcomes between the actors belonging to different systems and communities. It implies investigating the actor's perceptions of what is happening or taking a retrospective approach (what has happened). Activity Theory has been used in the field of open data (Ruijer et al., 2017).

However, I realised that the empirical context in Belgium was not appropriate. Open data is encouraged at the political level (see Section 3.2.4.) but the initiatives are still scarce and with low maturity. After selecting a first collection of initiatives, I saw that the participatory activities were seldom synchronous with my period of data collection, in-person (the participants being at home, using a personal device), and involving people that the initiative leaders knew personally. Most initiatives were side projects of organisations and hobbyists, or nascent start-ups of developers, whose main activity was not (yet) the running and maintenance of the initiative. It means that a data collection with long-term observation and interaction with the participants might have resulted in poor data. It is highly dependent on the project development phase and the motivation of its leaders. It made the use of techniques such as observation of participants and interviews about their motivations and points of view complicated. I also preferred, after a second thought, to build my contribution based on a conceptual framework and multiple cases as it can create knowledge that is more generalisable, which is in line with my practical research problem and purpose.

### Chosen research strategies, data collection and data analysis techniques

To answer the research questions, I chose to conduct a multiple cases study in two national cases: Belgium and France. The two fields offer more opportunities to find cases meeting the criteria of selection consistent with the research questions. In the meantime, France is better ranked in terms of open data maturity (see Section 3.2.3.).

The analysis aimed at understanding the innovation approaches and the roles given to the citizens in participatory initiatives. To build this knowledge, I applied an *abductive r*esearch approach and organised my *data collection in two stages*, based on *theoretical sampling*.

In a *theoretical sampling*, the choice of instances (cases) is based on the development of a "theory" that is grounded in empirical evidence. The confirmation or modification of the theory directs the selection of the next instances for the next phase of research, until the researcher has enough information and reaches saturation (Denscombe, 2010).

Therefore, at the beginning of the *first stage of the theoretical sampling*, I determined selection criteria relevant to the research question (See Chapter 4, Research methodology), with the aim to refine a conceptual framework (the bases of a "theory", a value chain). I ran the first round of selection going systematically through different sources of cases and picking out the cases that met at least two criteria. I proceeded this way because some cases required a deep analysis to validate the criteria, which interrupted the workflow in the initial selection phase, and because understanding the excluded cases helped me to make sense of the included ones and identify patterns. I targeted 50 initiatives to maximize the chance to reach saturation within the initial sampling. In *the second stage of the sampling*, I analysed deeper each case and I identified that 26 were meeting the three criteria.

Those 26 were through the *coding stage*: they were documented, using triangulation of data sources, coded and analysed. The coding was *abductive* because *mixing inductive and deductive methods*. I used an *inductive* (conventional) content analysis (Hsieh & Shannon, 2005) on the retrieved documents and interviews with initiatives leaders to identify the innovation approaches at stake. Patterns were revealed after about ten cases. Once grouped into theoretical categories, i.e., three innovation approaches, I purposively interviewed initiative leaders from each of the category (covering each sector and type of innovation approach identified) to triangulate the data source, enrich, and confirm the first findings, continuing with an inductive (conventional) content analysis (Hsieh & Shannon, 2005). Simultaneously, I developed *deductively* a conceptual framework based on Abella et al. (2017), Berends, Carrara, Engbers, and Vollers (2017), and citizens roles identified in previous literature. It helped to confirm the use of known citizens' roles, identify new citizens' roles, and develop an augmented open data value chain.

The method, relying mostly on documentary sources and reported information (interviews) has the advantage to broaden the access to cases with a limited time invested in each of them. Combined with the use of previous literature, it allows for building a bigger picture of the phenomenon. However, without observation and interaction with the participants, we miss their perspective and experience of open data. We can only hear the story of the respondents, the initiators of the participatory activities that might be embellished.

#### Contribution to the main research problem

The exploration of innovation approaches in participatory data-driven initiatives, in cities, has the objective to clarify, at least partially, the influence of the paradigms on the actors who make and send roles (the misalignment Actors – Roles) and the meaning they give to the "citizen" and "citizen participation". The use of a conceptual framework, an open data value chain, to identify and position the citizens' roles along the release and reuse process should allow us to see the output of their participation and the nature of the value created. The value chain shows the possible entries for participatory activities and the possible value resulting from them. This contributes to cover the misalignment Actor-Outcome, given that *a fortiori*, not all participatory activities can be understood with a civic meaning and result on social value creation, as municipality's broad objectives could let it believe.

### 3.3.2. Study 2

#### Research purpose and empirical context

In the second study, I keep on exploring the role issues but focus on the municipalities. The purpose of the study was to understand the possible roles of the municipalities and the expectations they generate on the opposite roles, the "user" of their output (input for actors in the rest of the process). Given that, municipalities can be both publisher and user, and are, first of all, a public actor in the city, it opens to a broad range of outputs and expectations.

I chose the national cases of Belgium and Sweden for three reasons: both countries have a different level of maturity (see section 3.2.4.), have integrated open data into their law and national digital strategies, and I had an active collaboration with Swedish researchers. They could help with relevant friendly reviews along the research project.

#### Alternative methods and theory

For that study, I first considered the **Actor-Network Theory** (Latour, 2005) as a theoretical framework. I thought of conducting a multiple case study based on interviews of different actors inside the municipalities (political leader, open data manager, operational staff) and their perspective on the user/citizen role in open data. The Actor-Network Theory (ANT) is relevant to studying the actors, roles, and their relationships in a network. It has the specificity to consider any agents, human or non-human (e.g., objects, documents, technologies), as influential actors. This theoretical framework rejects the techno-deterministic views, the idea that human actors fully control non-human actors, and that non-human actors are neutral (Silvis & M. Alexander, 2014) which fits with my assumptions. However, the focus is on the power relations and power flows, which I consider more appropriate in an interpretivist or transformative paradigm than a constructivist pragmatic. Methodologically, it implies identifying the actors, tracing their interactions, constructing a model of network to identify cohesions, strengths, and influences (N. Carroll, Richardson, & Whelan, 2012). Once again, my empirical context turned out to be not appropriate for such a method. The first prospection for respondents and cases gave a very low rate of response. I targeted medium-sized to big European cities for their maturity with open data. Seventeen cities were contacted for only two positive answers. After conducting three in-depth interviews with the open data managers and operational staff of the city of Gent in Belgium and Goteborg in Sweden, I also understood that they had little contact with other actors such as the re-users (which corroborates the previous literature and the possible misalignments) and the political leaders. It strongly compromised

the case study as initially designed and an ANT approach. Moreover, after consideration, studying only the large leading cities and their initiatives would bear the risk to create knowledge that is hardly transferable to the small municipalities, with limited resources. The last significant constraints that determined my method choice, in spring 2020, were the Covid-pandemic and unpreceded governmental restrictions that locked everyone down at home for a few weeks that became months.

### Chosen research strategies, data collection and data analysis techniques

Refocusing on my objective of producing useful knowledge, in a pragmatic perspective, I decided to use instead a Design Science Research approach to develop an artifact developed and tested on a broader set of cases. I chose Role Theory as theoretical lens. The advantages I saw in Role Theory, which is not one theory but a group of research streams studying roles, are several. Each research stream in Role Theory has a specific approach to roles (Nyström et al., 2014), for instance, the role as a social structure or the role as a position to control resources. Used in combination, they can bring interesting role-related issues to light, taking into consideration actors with different rationales (e.g., municipalities and developers). Roles can be defined based on the activities performed by the actors (Nyström et al., 2014; Selig & Baltes, 2018) and to understand the activities, we are not limited to data collection techniques such as interviews (reported activities). They can be observable through written sources (e.g., activities and tools made available for other actors). For instance, Herrmann, Jahnke, and Loser (2004) used written dialogues in a web platform to find roles in web-based community systems. Role Theory has been used in innovation and management research to understand how innovation processes are organised, what roles exist among them and what are their output (e.g., Dedehayir, Mäkinen, & Ortt, 2016; Heikkinen, Mainela, Still, & Tähtinen, 2007; Öberg, 2010). Role Theory also brings the concept of role expectations and the idea that roles can be complementary and interdependent to reach a common outcome (Heikkinen et al., 2007). This perspective on roles is said transactional, and describes roles in terms of focal position and counter position (Markham, Ward, Aiman-Smith, & Kingon, 2010) which is touching the research gap I identified for the study. Finally, to my knowledge and at the time of writing the study, Role Theory had never been applied in the context of open data despite the high potential I saw in it.

I developed a typology (artefact) that should help its intended audience to see the existing municipalities' approaches to open data, the type of users they call for and the role-related issues that can interfere with the realisation of expected benefits. I used the method of Nickerson, Varshney and Muntermann (2013), anchored in DSR, and I followed the steps of Peffers et al. (2007). I used data from previous literature, conceptual and empirical papers, to access more cases and potentially identify more roles than any other method, especially during the Covid-19 lockdown. To identify relevant papers, I used two tools: the browser Google Scholar and the database Science Direct. Google Scholar, because it is known to index a large scope of references not available in controlled databases, including scientific reports, working papers, white papers, presentations, thesis, and magazines (Hancock & Algozzine, 2006). Despite these sources are not always usable as scientific references, they can help the researcher to understand concepts and perspectives, which support the critical understanding of a body of literature (Boell & Cecez-Kecmanovic, 2014). The access to a large type of documents can also ease the snowballing and citation tracking methods suggested by Boell & Cecez-Kecmanovic (2014), which is also facilitated by the features of Google Scholar (Hancock & Algozzine, 2006). However, Google Scholar can miss recent publications under embargo, the

reason why I combined it with the controlled database Science Direct. I chose this one specifically has it includes the Journal of *Government Information Quarterly*, recognised by the research community as one of the main source of government literature (Virkar & Viale Pereira, 2018), and publishing a large amount of papers on open (government) data. The papers' selection and exclusion criteria are detailed in the method section of Study 2.

Then, I consolidated the typology by testing it on the municipalities of two national cases: Belgium and Sweden. I limited the data collection to written sources, websites, portals, and all types of reported offline activities that constitute the output and create a channel for interaction between the municipality and the user. On the other hand, it allowed me to include in the study every municipality of Belgium and Sweden publishing any form of open data, and not only the medium-sized or leading big cities. The main weakness of the method is that it can miss ongoing but unpublished open data projects, future projects. I did not plan to conduct interviews due to the high number of cases and the persistent difficulty to find contact details and responsible persons for open data in municipalities.

#### Contribution to the main research problem

This study contributes to the misalignment Actors – Roles, as it seeks to clarify what activities are under the scope of a counterpart role. The aim is to find role-related issues, as a way to improve the alignment of the actors to each other and towards a common outcome. Therefore, the study also touches upon the misalignment Actor-Outcome: one of the dimensions of the typology is the output delivered to the user, which can open the discussion of its consistency and relevancy with the initial objectives and expected outcomes of the publishers.

### 3.3.3. Study 3

#### Research purpose and empirical context

The purpose of the third study was to understand how to capture the needs of end-users and help the publisher and the re-user to better align their work to these needs. The most popular methods to involve people with data are addressed to the re-users. I argued that end-users do not need data, but information. The concept of information needs is especially developed in information sciences with the seminal work of Wilson (1981), library and documentation sciences (e.g., Devadason & Lingam, 1997), and information behaviour (e.g., Nicholas, 2010) among others.

#### Chosen research strategies, data collection and data analysis techniques

Like Study 2, Study 3 is anchored in DSR. In our first attempt to understand how to capture information needs relatable with open data, and with methods having a limited impact on the participants, we used complete observation (Baker, Lynda, 2016) in combination with interviews in a crowded information-rich environment: a bus hub (relatable with public transport data). However, we realised that it was not possible to associate an information need with a behaviour (e.g., look at a billboard or smartphone) without the participants' explanation or outside of a controlled environment (e.g., lab experiment). The semi-directed interview, using questions based on previous experiences and contextualisation was more effective and revealed the need to guide the respondent in this specific intellectual process. Therefore, we decided to investigate the scenarios in a collective workshop, with DSR as research approach, as the purpose of the study was to develop an artefact (a method). I followed the steps of Peffers et al. (2007) to develop a method. I was challenged by reviewers for the value of a

method, judged as too context specific. I elevated the learnings of the instantiations to a more abstract level of knowledge, Design Principles. I ran three iteration loops, first with end-users (workshops with a group of researchers and a group of students, as inhabitants of a city), and then with publishers (civil servants from local and regional administrations in digital workshops). The design principles include the involvement of end-users and allow the design principles' implementers some flexibility given their context. For instance, the last workshops involved the publishers in a user-oriented reflection, accompanied with tools and ideas to involve their end-users, in their environment, after the workshop. We used a popular science article to evaluate the artefact with the audience of practitioners and a scientific workshop (the 17th Scandinavian Workshop on E-Government, 2020) and a conference (EGov conference, 2020) to discuss the contribution with the community of researchers before the final submission.

#### Contribution to the main research problem

This last study addresses the misalignments at the Role – Process level. The produced artefact is a set of design principles. They are methodological recommendations for the publishers, re-users, and possibly end-users, to capture the information needs of groups of end-users. Information needs are complex and difficult to capture by nature, as they are personal, intangible, temporary. Once we know, the information need disappears, and information sources can be everywhere. To deal with this complexity, the concept of role is once again central. I used it to circumscribe the end-users in a given context and social position, needed to guide the intellectual process of identifying specific information needs. The papers use previous knowledge from Service Design and Participatory Design and apply it to the open data domain, where I perceived their potential value despite being seldom seen and used in previous open data research and practitioners. The research contribution is what Gregor & Hevner (2013) call "*exaptation*": we compiled and extended known solutions to new problems or new fields. Since the main audience of this paper is the practitioners, we have also made the paper available in open access.

## 3.4. Methodological limitations

The researcher's paradigm influences the methodological choices and, therefore, implies methodological limitations. Referring back to the second and third hypotheses of the constructivist paradigm (Section 3.1.2.), Avenier and Gavard-Perret (2012) argue that the researchers should be aware that the knowledge they develop depends on their knowledge project, background, personal history, and that the intention to know influences the way they experience it (what they study and how). In terms of methodological limitations, it means that methods are selected based on their appropriateness to the research question, but also my skills, previous knowledge, and current methodological repertoire.

I am aware that other theoretical lenses could have been chosen (e.g., the coordination theory mentioned in Section 3.2.3.) and would have led to other findings. The choice for specific tools and methods influenced the data accessed and collected. In Study 2, for instance, I can have missed papers not indexed in the browser Google Scholar and the database Science Direct that I chose to use, and papers not using the keywords that I used in my selection criteria.

Nevertheless, the main limitations I have experienced through my research were the access to the practitioners. It determined the methodological choices. In Study 1, there was seldom a name on the initiatives and channels to get in touch with the leaders, as such, it was based on

documents and selected interviews. In Study 2, a few municipalities had open data integrated in their process, and therefore, an organisational structure that identifies responsible employees, and make them reachable by the citizens. Moreover, there was not always a desire to facilitate interaction with the re-users. I was able to see how few municipalities were giving contact details for the re-users to ask questions, what I could have used to reach out practitioners. Consequently, the framework (Study 1) and typology (Study 2) have not yet been submitted to practitioners to discuss, test their viability and actionability in practice.

The access to the data was also made complicated by the exceptional circumstances. In March-April 2020, the whole country, Belgium, was put under lockdown for a few weeks due to the covid pandemic. The governmental recommendations for homeworking affected drastically people's life, work, habits, and the opportunities to organise collective activities. This situation was one of the reasons to choose to use secondary data from the literature in Study 2, despite we know published documents does not show all the intentions and projects of the municipalities. In Study 3, the best evaluation of the design principles would have been to let the audience try to implement them. I did not have the opportunity to find partners and run or observe practitioners running the method in presence due to the persistent "stay-at-home" recommendations in Belgium and Sweden.

# Chapter 4. Citizen Participation with Open Data: Unpacking Innovation Approaches and Citizens' Roles in Urban Settings

## Related publications

**Ongoing research paper** published in a conference

Gebka, E., & Castiaux, A. (2019). Data-driven initiatives in Smart cities: citizen participation and value creation. In *ISPIM Conference Proceedings* (pp. 1–9).

**Full-research paper**

Currently under revision in a journal (submitted the 19th February 2021, first round).

## Abstract

Participation in innovation is progressively conceptualized as a valuable practice to enrich the innovation output. Open data, as an innovation means in urban settings, is believed to enhance citizen participation and social value creation. Yet we have little knowledge about the scope of the citizens' roles in data-driven initiatives and their relation with innovation approaches. This paper explores that gap by analysing initiatives that aim at generating social value in urban settings. We observed three innovation approaches – open data innovation, service innovation, and civic innovation – and extended an open data value chain to include the citizen's roles, who contribute in the data layer, project layer and society layer.

**Keywords:** open data, citizen participation, urban innovation, social value

## 4.1. Introduction

Data-driven initiatives are projects that reuse data to provide new information solutions, services, or products or to improve the existing ones. One source of data can be open data, data that can be freely used, re-used, and redistributed by anyone for any purpose, subject only to the requirement to attribute and share-alike (Open Knowledge Foundation, 2015).

In open data literature, citizen participation in innovation and governance is an abundantly used concept and claimed objective. Open data is seen as a means for value creation, through innovation and citizen participation as key generative mechanisms (Jetzek, Avital, & Bjorn-Andersen, 2013; Pereira et al., 2017). The Open Knowledge Foundation, a civil society leader in the open data movement, argues that the concept of "openness" implies "universal" participation: anybody should have access to data and information to create and share knowledge without restriction, for self-empowerment, improved efficiency of products, services or governments, and democratic control (Open Knowledge Foundation, 2015). Beside that movement, one of the main sources of open data is the public sector (Berends et al., 2017). While releasing open government data (OGD), the publishers expect that it will create value by improving transparency, participatory governance, innovation, and that it will enable citizen participation and engagement in these objectives (Attard et al., 2015; Charalabidis et al., 2018a). In the public sector, both citizen participation and open data stem from democratic values, good governance and a need for legitimacy of the public actions and authority.

Open data, especially in urban settings, is, therefore, anchored in the assumption that participation is a significant value generator mechanism. Indeed, data is basically released for the use of somebody else. However, the way the actors approach innovation and open data, in other words, their rationale, can imply different forms of participation and citizens' roles. In the literature, the "citizen" is used as a catch-all term, where are mixed up the roles of direct data reusers and indirect end-users of data-driven solutions (Safarov et al., 2017). The objectives, forms and results of the citizen participation are unclear, especially when it comes to data-driven innovations in urban settings. In that context, participation can carry multiple meanings and be associated with a source of social value creation.

In sum, we have little knowledge about the scope of the citizens' roles in data-driven initiatives, and the relation between the performed participatory activities and the innovation

approaches. This question is particularly relevant for the initiatives aiming at creating social value in urban settings. In research and practice, there is a need for clarification regarding the concept of citizen participation in open data-driven initiatives, from a process and output point of view, and in relation to the rationale of the actors.

The purpose of this research is to unfold the citizens' roles in data-driven initiatives that take place in urban settings and aim at creating social value. We want to understand the rationales and implications of participation and open data reuse in different approaches to innovation, by answering the following questions:

- *What are the innovation approaches taken by participatory data-driven initiatives in urban settings?*
- *What can be the citizens' roles and patterns per innovation approach?*

Our contribution is twofold: (1) we propose three innovation approaches of participatory data-driven initiatives in urban settings, and (2) fine-tune one open data value chain model (Berends et al., 2017a) to discuss the relationship between the innovation approach, the open data reuse, the participatory activities performed, and the resulting value.

## 4.2. Background

In this section, we introduce the diversity of meanings and forms of innovation, citizen participation, and open data, with a focus on urban settings. This conceptual clarification establishes the knowledge gap addressed by this study and the reasoning used further in the method.

### *Innovation in urban settings: open and participatory*

In complex ecosystems, like cities, wicked problems involve multiple stakeholders. Innovating for the public interest is not anymore the duty of the public actor only, cities are becoming fields for public, collaborative, social, and open innovation. Behind these similar concepts stands a variety of rationales.

*Public innovation* relates to innovation in public organisations, processes, services and governance (De Vries et al., 2016). The central activity of public managers is the creation of *public value* and serving society as a whole (Moore, 1995). *Public Innovation* enhances public value creation and is therefore realised through the public actor's actions. However, citizen participation increases to develop new public services, in what is called *Collaborative Innovation* (e.g. Bommert, 2010; Sørensen & Torfing, 2011).

When the public actor has not the resources or knowledge to innovate and solve issues of public interest, it can also call for *Open Innovation*. The paradigm, coming from the private sector and developed by Chesbrough (2003), supports the idea that inbound or outbound knowledge flows can benefit the organisation. By sharing internal knowledge with external innovators or importing external knowledge inside the internal processes, an organisation can maximize innovation effectiveness. In the same vein, the thesis of Von Hippel (1986) is that the user is an expert of his environment and needs and therefore is the most aware of the solution required to fulfil his needs. A lead user is an expert and skilled user capable of innovating for himself, whose talents and ideas can help organisations to innovate. For instance, in Living Labs, user-centred and participatory methods are used to co-create technological innovations with

the users and stimulate bottom-up innovation (Dell'Era & Landoni, 2014), which method is popular in cities (e.g., Björgvinsson et al., 2010).

Innovation can also contribute to a greater good without the impetus of the public actor. To drive changes in social practices and solve social problems is *social innovation's* main purpose (Cajaiba-Santana, 2014). *Social value* is the value created when resources, inputs, processes or policies are combined to generate improvements in the lives of individuals or society as a whole (Emerson et al., 2000). Participation of the stakeholders is a core component.

This diversity of participatory innovation paradigms leads also to a conceptual ambiguity concerning citizen participation, especially in urban settings. The citizen can in fact relate to several roles and participation purposes. In this paper, we use the definition of Anttiroiko (2016, p.8): the citizen is (1) a representative of a consumer, customer or user group, or (2) a political actor, community member or a resident. Citizen participation is therefore a voluntary involvement in (1) a facilitated innovation process, or (2) a democratically organised policy-making, planning or governance process.

The widespread enthusiasm for participation in innovation, especially in cities, feeds a growing interest in open data. The next subsections elaborate on the open data as an innovation means, its use in urban contexts, and the barriers and opportunities it raises for participation.

### *Open Data as an innovation means*

Open data has become a popular innovation resource in the public sector and can be seen as a critical resource in the innovation process (Magalhaes & Roseira, 2017), especially public and open innovation. The underlying idea that the benefits of open data should be realised using the wisdom of crowds and collective intelligence (Janssen et al., 2012) in a paradigm of urban open innovation (Chan, 2013). This involves at least two roles: the publisher, who produces and share data, and the re-user, who makes secondary use of the data (not to be confused with the user or end-user of services).

There are various degrees of openness and potential open data sources. The Open Data Institute depicts it in a continuum: the *Data Spectrum*[6]. It ranges data from *closed* (the personal and private data), to *shared* (data open only within named, group-based access), and finally to *open* (from public data, to open data, accessible for anyone, anywhere). Open data sources can include public bodies with Open Government Data (OGD), but also Open Business Data (OBD) or Citizen-Generated Data (CGD). The latter is data produced by citizens, which can be obtained by using, for instance, Internet of Things (IoT) devices (Charalabidis et al., 2018a).

In conclusion, the origin of the data shall determine the degree of openness, accessibility and usability of the data. It has two impacts in the context of our research: (1) in urban settings, the data sources, publishers and re-users can be many; (2) the degree of openness, accessibility and usability can become a barrier for participation and innovation. We elaborate on these two considerations in the next subsections.

### *The urban settings, a place where flourish open data publishers and re-users*

[6] The Data Spectrum, Open Data Institute: https://theodi.org/about-the-odi/the-data-spectrum/

The Public Sector Information (PSI) Directive (2003) in Europe has been a strong incentive for governments, territories, and municipalities to publish OGD. Public actors share nowadays a substantial amount of open data, that cover themes such as business, education, health, manufacturing, and public safety (Okamoto, 2017). Jetzek, Avital and Bjorn-Andersen (2014) suggest that the act of opening data and treating them as a shared resource has the ability to contribute to the simultaneous generation of economic and social value, which is a strong argument for the public sector to invest in open data platforms. For example, civic apps based on open data can address a huge range of public issues (Desouza & Bhagwatwar, 2012). According to a European report (Berends et al., 2017), the most reused and combined open datasets by the private sector come from the data categories of regions and cities, transport, environment, and population and society. In some cases, the cities, publishers, are also the re-users and develop data-driven solutions in the form of applications, improvement of public services, efficiency and decision-making in public actions, citizen participation in urban planning, or greener environment (Hemerly, 2013; Ojo et al., 2015; Hemmersam, Martin, Westvang, Aspen, & Morrison, 2015; B. Wilson & Chakraborty, 2019).

The other main data re-users are the private sector and civil society. Data re-users can belong to different social groups: citizens, developers, students, researchers, journalists, members of a civic tech community, civil society organisations, city managers, non-governmental organisations, large cooperations, entrepreneurs, or start-ups (Lassinantti et al., 2018).

*Barriers and opportunities for participation with open data*

The principles of "universal" participation, open innovation, and citizen participation enhanced by open data enthusiasts face up a hard reality: a high level of technical and analytical skills is needed to deal with such datasets (Zuiderwijk et al., 2012). This recalls one of the main limits of the Lead User Theory (Von Hippel, 1986; Franke, von Hippel, & Schreier, 2006): only a limited of skilled and expert users are capable of innovating. Citizens are said beneficiaries and re-users of O(G)D (Lassinantti et al., 2018). However, the role and the actual participation of the citizens in open data initiatives are questioned (Attard *et al.*, 2015; Canares, 2016).

Using datasets is not obvious and data re-users are suffering important barriers and impediments in the reuse process. Governmental publishers tend to stay in a supply-logic as OGD platforms are evaluated according to whether they are complying with the law or not, and not based on the usefulness of the information provided (Attard et al., 2015). Consequently, there is a discrepancy between the offer (the data provided) and the demand (the datasets desired by the data re-users to create value and meet their purposes) (Gonzalez-Zapata & Heeks, 2015).

Moreover, data re-users face several socio-technical barriers and challenges (Janssen et al., 2012;Gascó-Hernández et al., 2018; Crusoe et al., 2019). Data literacy and high technical skills are needed in a data-driven economy (Hemerly, 2013). In addition, the poor quality and a lack of metadata are the first technical impediments that can slow down the reuse process, increase the need for analytical skills and result in a time-to-value ratio prohibitively long for the reuser (Sadiq & Indulska, 2017). Consequently, the reuse can become challenging if the data delivered is neither usable nor perceived as useful by the reusers. The assumption that citizens will develop innovative solutions just because the data is released is challenged by researchers (Charalabidis et al., 2018a). Intermediaries are needed to help communities to reuse open data (Yoon, Copeland, & McNally, 2018).

Nevertheless, we can foresee openings for participation in the innovation processes of data-driven initiatives. As mentioned before, the citizen can be an open data provider (Citizen-Generated Data), the data that individuals consciously collect and share, and are available for use in the public domain, for example geographical, environmental or mobility data. Data-driven business models also struggle with open data quality and availability and have to combine internal and external data sources to tackle the issue (Hartmann, Zaki, Feldmann, & Neely, 2014). Crowdsourcing and sensors help in improving data density, quality, trustworthiness, and real-time data procurement (Mirri, Prandi, Salomoni, Callegati, & Campi, 2014). Finally, the citizen can likely add value to the data-driven initiative. User-centric approaches in the development of applications are becoming a common way of working in the digital sector. User participation is known to be valuable in new product development (Cooper, 2018). In conclusion, the dichotomy of citizen's roles stated by Safarov et al. (2017), re-user or end-user, can be reasonably questioned.

## 4.3. Research methodology

This qualitative research is a multiple case study. The qualitative approach was the most appropriate to understand the *why* and *how* of the phenomena under investigation. The unit of analysis is the data-driven initiative taking place in urban settings and aiming at creating social value. The units of observation are the rationale of the initiatives (why), and the activities performed by the citizens (how). The research has an explanatory nature given that open data is a recent phenomenon and there is no dominant model or theory in the open data literature (Zuiderwijk, Helbig, et al., 2014).

We selected cases based on criteria, then conducted the data collection and analysis in two parallel loops of inductive and deductive analysis, as presented in Figure 16.

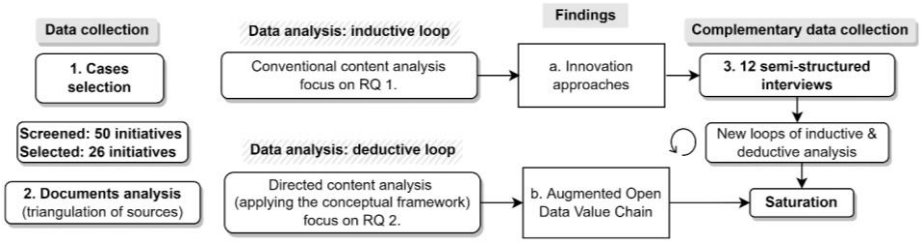


*Figure 16. Research approach in relation to the research questions and findings*

The selection of initiatives was carried out by screening the description and documentation of data-driven initiatives reported on 10 Belgian open data portals, 28 French open data portals at the municipal, regional level and national levels, the laureates of one Belgian national call for projects related to open data reuse for mobility, one open data community website, open data events' reports and programs, and one French open data association. Open data portals and dedicated events were good pools of initiatives, since advertising reuses is a typical strategy of public publishers and data enthusiasts to display the best practices. Both countries (Belgium and France) were chosen for their cultural proximity and field accessibility, and because the French national digital strategy decrees the open data release as a legal obligation (Conseil d'Etat & Chambre des représentants, 2016), which has positioned the country as a European trendsetter (van Knippenberg, 2020).

The data-driven initiatives are expressions of innovation practices, to reach an outcome with the use of data and citizen participation. Therefore, we used three criteria to select the initiatives. First, the context and expected outcome of the projects: the initiatives should aim at delivering a product, service, solution, or action for the urban territory or the citizens as residents which at least one expected outcome is to generate social value as defined by Emerson et al. (2000). Second, the data: the initiatives should make use of open data (public data, open data, OGD, CGD, or BOD). Third, citizen participation: the initiatives should include a form of voluntary and conscious citizen participation in the process of value creation. For each relevant initiative, we triangulated the data by collecting empirical material from different sources: the initiatives' sources (web pages, projects documentation and press releases issued by the initiatives), and third sources (websites and news reporting the project, and descriptions in the app stores when available). A list of all the selected initiatives (anonymized with code names), with the number of documents collected and analysed per type of sources is available in Appendix I. A. The documents were stored and coded in the software NVivo. We supplemented the document analysis with twelve purposely selected semi-structured interviews (see the interview guide in Appendix I. D. and the list of respondents in Appendix I. E.) to confirm our preliminary findings and our understanding of the initiative's rationales. The interviewed initiatives represented the three innovation approaches identified, in the public sector, private sector, and civil society.

To analyse the data, we used an abductive approach mixing directed and conventional content analysis (Hsieh & Shannon, 2005). The inductive coding process was used to answer our first research question, which initial code scheme is available in Appendix I. B. We used the directed approach to analyse the citizen participation in data-driven initiatives at a process level, with the help of a conceptual framework (Figure 17), adapted from Abella et al. (2017) and Berends et al. (2017). It adds the outcome dimension of Abella et al. (2017) into the data value chain of Berends et al. (2017) that focused on the actors, roles and tasks performed to add value. The conceptual framework was used to identify the activities and roles performed by the citizens in the participatory initiatives across the data value chain's phases and discuss the resulting value.

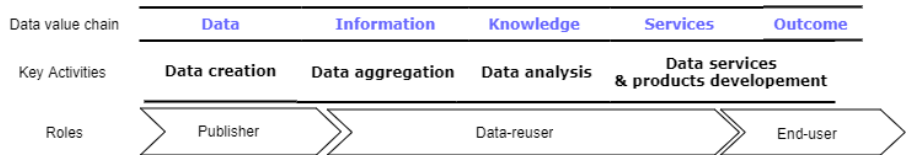


*Figure 17. Open data value creation framework, adapted from Arbella, et al (2017) and Berends, et al (2017)*

The role and set of activities served as initial codes (see Appendix I. C.). New categories of participatory activities and roles came out of the analysis and were added to the original open data value chain.

In total, 50 initiatives were screened, wherein 26 were selected for deep analysis. After deeply analysing about 10 initiatives, we started to see patterns. We continued to look for new initiatives to reach a good coverage of each approach. In the end, we analysed and coded 12 interviews and 163 documents from direct and indirect sources, including the documents of the rejected initiatives that contributed to our understanding. The interviews confirmed the

preliminary categories of innovation approaches, and added nuances and richness to our analysis. They revealed that initiatives can share values and ideas of different approaches, and make use of participatory activities not reported in the written sources. The last half of the analysed cases did not provide evidence of other innovation approaches or new participation forms. We concluded that we reached saturation.

## 4.4. Findings

After introducing the context of the cases, we present the findings in two parts: (1) the rationales behind the use of data and participation in the initiatives taking place in urban settings, revealing three innovation approaches, and (2) an augmented data value chain as a synthesising analysis of the participation activities and citizen roles.

### *Context*

In Belgium, open data is a strategic priority for the federal government. In 2015, a task force was set up to implement the PSI Directive, energize a community of re-users, and provide tools and incentives to help the public actors to release more OGD. For example, a recent incentive was the Smart Mobility public call for projects (2019), which funded initiatives making use of and generating new open data. Local initiatives started earlier at the municipality level, for instance in Gent, where the first open data portal was launched in 2009 and the first hackathon run in 2011. However, data-driven initiatives in urban settings are still in their infancy and at the European level, Belgium is ranked 24$^{th}$, in the "followers" category (van Knippenberg, 2020).

In France, releasing OGD is compulsory by law since 2016. The republic has established a ministerial department dedicated to the design and implementation of the OGD strategy at all levels of government. It has resulted in a dynamic open data ecosystem, including, for instance, an association of local governments that helps its members to release OGD and promote reuses, and a vivid online community and forum that get together nearly 1000 data enthusiasts of all backgrounds. France is ranked third by the European maturity report, as a trendsetter (van Knippenberg, 2020).

### *Participatory data-driven initiatives taking place in urban settings: three approaches to innovation*

*Rationales behind data reuse and citizen participation*

The screening and analysis of initiatives revealed significant differences in the rationales of the initiatives underlying the use of open data and citizen participation.

In our context, both open data reuse and participation can be based on two rationales. The three innovation approaches are the result of the combination of these logics, as illustrated in Figure 18.

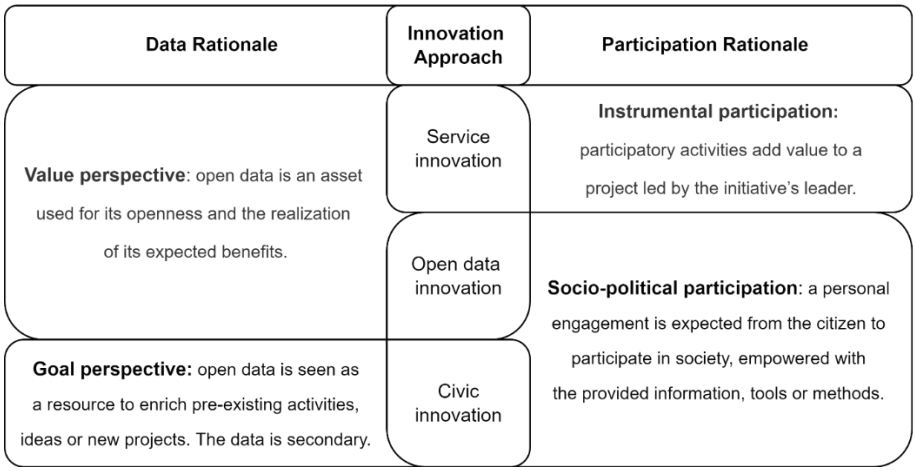


*Figure 18. Data and participation rationales in the innovation approaches*

Initiatives can present characteristics of two approaches, but one was always dominant (see Appendix I. A.). In the next sections, we elaborate on the approaches.

*The open data innovation approach: Data is open, and openness is a cause*

Innovation, in that approach, relies on the openness (accessibility, reusability) of data and aims at realizing the expected benefits of open data for society. For the initiative's leaders, access to open data is a right and a pool of opportunities to create value, share knowledge, and solve issues for the benefit of most. Table 9 presents the main features of the approach with empirical examples. The codes (e.g., 2/CS) refer to the cases reported in Appendix 1, and to the sector of the initiatives (CS: Civil Society, P: Public, B: Business/private sector).

*Table 9. The open data innovation approach: features*

| *Features* | *Examples* |
|---|---|
| ***Initiative leaders:*** *community-driven (civil society)* | |
| ***Participation rationale: socio-political***<br>*- The citizens should have control over their personal data and access to public data to benefit from its value: make better decisions, monitor the public action, participate in the management of public issues, or act as a counter-power.*<br>*- To succeed, a priority is to stimulate the development of skills, data literacy, and participation in the definition of data reuse, for instance.*<br>*- Participation is encouraged in every stage of the data value chain.* | ***Participation forms:***<br>*Enable the citizens to collect data and develop visualisations tools to join the public debate (2/CS), ease data reuse for more citizens and enable the development of innovative/alternative mobility solutions (4/CS), collectively identify data and ideate about potential reuses for a territory (15/CS),* |
| ***Data reuse rationale: value perspective***<br>*- Open data is an asset and the creation and reuse of open data is the purpose.*<br>*- Initiatives' leaders want to explore data, show its intrinsic value to the world, and claim more openness in a logic of open data advocacy.* | ***Services and products (output):***<br>*Sensors kits combined with visualisation dashboards (2/CS), open-source tools for the community of re-users (4/CS), methodology for citizens and public managers to identify data and solutions for more efficient public management (15/CS).* |

An example of initiatives in that approach is Civic Air Sensors 1 (2/CS). The initiatives' leaders believe open data has an intrinsic value that the society should benefit from. They work collectively on solutions and want to empower the citizens.

> *« There is the idea that we are going to work on projects open to anybody, and that benefit to anybody. We work with open source, open data" (2/CS).*

For the members, openness is a way to contest power, claim the right to create and develop independent solutions, access to transparent information.

> *"In fact, it is like an act of rebellion. They [the participants] know that we can do the things ourselves and we will prove and show the reality [the air pollution]. It is interesting because we are not moved by any political party. We are neutral. We make the data speaks ». (2/CS)*

To reach that purpose, the project Civic Air Sensors 1 (2/CS), for example, ran two participatory workshops to measure air quality. By the end of the first workshop, people were able to extract the data provided by the sensors, analyse it and visualise it. The second workshop aimed at teaching people how to build their own sensor so they would only need to plug it in at home to be operational. The initiative's purpose is also to collect facts to ask the government to take measures to improve air quality in Brussels. Participation can cover all the steps of the data value chain, i.e., creation, aggregation, analysis, development of visualisation, which tends to attract people who are digitally skilled, have an interest in understanding data. The realisation of the social outcome, for instance, the civic actions to take based on the information visualized, can be challenging for these initiatives. The initiative's leader of 2/CS reported that "make the

data speaks" to understand air pollution is still a challenge for both the participants and himself. They miss knowledge and network to reach their purpose: they are not experts in air pollution, not recognized as an environmental NGO, and not connected with the relevant experts.

*Service innovation approach: Open data is a resource for innovative services*

In service innovation, data and digitalization is an opportunity to create new services or digitalize the existing ones, and in the end, generate value for the users and its community. This approach is by far the most prolific amongst the initiatives observed (16 out of 26). Features and examples are provided in Table 10.

*Table 10. Service innovation approach: features*

| ***Features*** | ***Examples*** |
|---|---|
| ***Initiative leaders:*** *start-ups, social entrepreneurs, and developers, followed by cities and civil society organisations* | |
| ***Participation rationale: instrumental***<br>*- Participation improves and increases the value of the output: the citizens are involved in data provision and service development.*<br>*- The output is believed to improve society (social outcome) by encouraging citizens to adopt civic behaviours to tackle public or societal issues* | ***Participation forms:***<br>*Collect original data (CGD) that feed the solution (46/CS), or supplementary data that improve the delivery of a service (45/P), test a solution under development (1/B), support the local economy and encourage shopping in local stores (1/B), ease the recycling of waste and promote the respect of the environment (38/B, 49/C), encourage the use of bicycles (11/CS).* |
| ***Data reuse rationale: goal perspective***<br>*- Reuse data for the development of new ideas and solutions, as open data is a freely available resource.*<br>*- Initiatives combine the data sources (on-demand, under licenses) to develop attractive services.* | ***Services and products (output):***<br>*Digital services that aggregate information to localize local services and points of interest (1/B, 28/B, 40/P, 43/P), infrastructures (11/CS, 49/CS), local food producers (20/B, 31/B). Route planners that integrated different public transports (3/B, 5/B), web apps that ease the navigation of disabled citizens (39/P, 41/B).* |

In that approach, the purpose of the initiative is not the creation or the reuse of open data, but the improvement of the information, service delivery, and maintenance of the public infrastructures, for a better experience for the inhabitants of the city (goal perspective).

For example, Local Information App (1/B) is an application that aggregates local information in one app and develops micro-services for the citizens. They sell their solution to cities as an innovative public information service.

> *"At the beginning of our reflection comes always "what does the citizen need?", "what is a local information?". We have integrated a lot of information related to the smart cities, mobility, the open data from the public transport, and the touristic data from the region (…)."*

The choice for open data is pragmatic and opportunistic. For example, Handicap-friendly Map 2 (41/B) is an application that helps disabled persons to navigate in the city and report obstacles.

*"I won an open data competition in 2011 with this app. (…) Often, it is through an open data initiative that the cities contact us. To be honest, we do not care whether the data is open or not, but we created Handicap-friendly Map 2 as a part of an open data initiative of a municipality and today we are branded as the app that can come on territory once there is an open data portal."*

The use of open data can be reconsidered when too difficult or lacking availability, reusability and perceived utility. One respondent (36/B) admitted:

*"We do not need to work with open data (…). We chose Google maps because it is a tool everybody knows".*

Five initiatives communicated to be based on open data, which brings them close to the open data innovation approach. Nevertheless, that communication was not driven by a logic of open data advocacy, but by the branding of the innovativeness of a public initiative or recognizing themselves in some of the values held by the open data movement. For example, public initiatives promoted that the solution was based on open data as a part of their innovation strategy (40/P), or a good practice of their open data strategy as a city (45/P). For 46/CS, the initiative was the result of the association's purpose, but reflected their identification in the open data values:

*"Our signature is really open data. We find it really powerful, the open data produced by the citizens are more updated, and more often, than the public data. And because our purpose is that the most people possible find the old pedestrian paths, it is logical to share our findings with everybody in open data".*

5/B said releasing open data and APIs produced by the project to "help" other developers. Going for open data is also a way for the actors to keep their autonomy and independence from the big players, owners of data and APIs (e.g., Google), and support their peers in the time-consuming task of web scrapping and APIs development (1/B). Open data is a strategical and political choice, which come on the side of the initiative's purpose.

Participation in the process is mainly instrumental and used to serve the purpose of the initiative's leaders. For example, by reporting holes on the street (providing complementary data) (45/P), the municipality can improve its service and maintenance of public infrastructures, and by identifying local producers in an app that localizes local products (31/B), the developer improves the value of the app. Participation in tests to give feedback is a common way of working in the development of digital solutions (1/B, 3/B, 5/B, 11/CS, 20/B, 38/B).

Participation as a citizen in public or societal issues is also implied in the objectives of seven initiatives, as an indirect outcome. The use of bicycles promoted by 9/B is also an argument to choose a responsible and environmental-friendly means of transport, just as the promotion of local producers is for food consumption (20/B, 31/B). 20/B said:

*"The idea is to compile tons of information close to the user in order to enable them to have a clear overview of what happens close by and change their behaviours."*

However, the realisation of this social outcome is not often properly evaluated. For example, five respondents (1/B, 5/B, 7/CS, 20/B, 47/CS) mentioned that they measured their success in the number of users or sensors instead of impact. Participation of the users in the desired civic behaviour or change is not either facilitated by the initiatives, beyond the claim of the objectives.

*Civic innovation approach: Open data is an information and tool for empowerment and change*

In that approach, data and digitalization offer new ways of being a citizen, understood as an inhabitant with political action and concerns for its community. Information provided by open data reinvents the civic engagement and the initiative's output offers new opportunities for participation in governance and collaboration in solving public issues (e.g., urban planning, traffic congestion, environmental problems). For the features and examples of this approach of innovation, see Table 11.

*Table 11. Civic innovation approach: features*

| *Features* | *Examples* |
|---|---|
| ***Initiative leaders:*** *environmental activists, associations for mobility and environment, companies, and social entrepreneurs.* | |
| ***Citizen Participation: socio-political***<br>*- Citizens are believed to be empowered with accurate data and information.*<br>*- Citizens participate in the provision of data and should engage in social issues.*<br>*- Tools for interacting with the stakeholders and following up actions can be included in the output.* | ***Participation forms:***<br>*Report and track environmental issues (44/CS), participate in the governance of urban mobility planning (8/CS), take action as an activist (e.g., environment, politics) (47/CS).* |
| ***Data use rationale: goal perspective***<br>*- Data is a powerful source of information to raise citizens' awareness, demonstrate objective facts and increase the pressure on the governments or municipalities on specific issues.*<br>*- The use of CGD is privileged for its relevance, quality, and transparency.* | ***Services and products (output):***<br>*Personal sensors (7CS) or public sensors (8/CS) combined with visualisation dashboards, reporting tools and interactive maps (44/CS), participation app to interact with stakeholders (23/B).* |

These initiatives' leaders, driven by civic values, consider the data as a means to reach their goal. For instance, Environmental Monitoring app (44/CS) is an initiative developed by a federation of environmental associations. They use CGD to centralize, report, and communicate about environmental issues and positive initiatives.

> *"This tool allows us to have an information medium to communicate and raise awareness about nature violations, when we give training or organise public events, based on trustful information. The citizens are the information source, and after the reporting, if they feel like, they can join a member association to deal with the information themselves, and follow the case. It is a way to recruit volunteers for us."*

Only in this approach, we observed initiatives that provided the tools, places, or functionalities in their output to make the expected social outcome really happen. For example, Traffic

Counter (8/CS) is an initiative of a mobility association using crowdsourcing to count the numbers of motor vehicles, bicycles and pedestrians passing by the streets. The project included also enabling the citizens to collaborate and participate in the design of mobility policies with the municipality.

> *"They will have that counter at their window, they can see their result on the website, and then what? We will study the data together. And the technical organisation will guide them in suggesting solutions, and see directly if we cut the street, over there, what will be the impact on the traffic. (…) And then at the end, we will write all this and our organisation will guide them in real suggestions like traffic technical suggestions they can bring to the municipality."*

However, the citizens who participate in collecting primary data are not necessarily the ones who will engage more time and energy in civic actions (44/CS, 47/CS). With sensors, the participation is active and conscious, but requires little or no time. Taking action based on the information can require the initiatives' leaders to mobilize other groups of citizens.

### *The Citizens' roles in an augmented open data value chain*

The analysis of the rationale of innovation approaches allowed fine-tuning the open data value chain. We identified new roles: primary data provider, secondary data provider, idea generator, and active citizen. We ranged them into three layers of participation: the data layer (participation in the back-end of the project and infrastructure), the product or service development layer (participation in the front-end and value proposition of the innovation), and the society layer (participation in society with actions, for the expected social outcome). Figure 19 presents the augmented open data value chain and the possible citizens' roles, described thereafter.

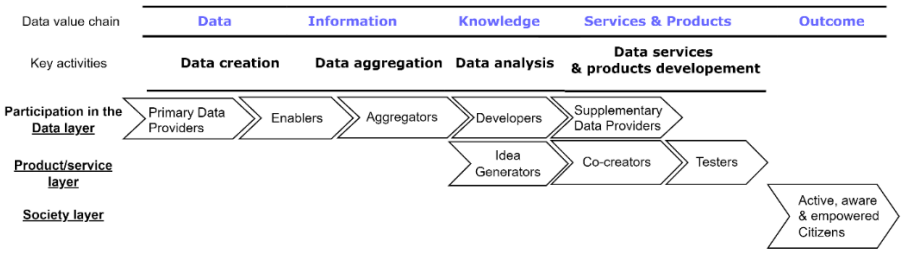


*Figure 19. Citizens roles across the Data Value Chain and three layers*

### *Participation in the Data layer*

- *Primary Data providers:* the citizens are involved in the primary data creation and collection with tools like sensors, dedicated platforms or simply paper and pens. The data is compiled by the initiative's leaders for analysis and transformation in the desired format for reuse.
- *Enablers*: the citizens bring their skills and knowledge to help the initiative leader in the development of open APIs and software to enable data reuse.

- *Aggregators:* the citizens bring their skills and knowledge to help the initiative leader in the analysis, combination and aggregation of data.
- *Developers:* the citizens bring their skills and knowledge to help the initiative leader in the development of data visualisations and applications, to deliver value to the end-users.
- *Supplementary data providers:* the citizens contribute to the creation and/or collection of supplementary data to improve the value of the new or an existing service or product, by adding, validating, or improving data and pieces of information via a mobile application or web platform. It improves or enables the delivery of the service or product.

*Participation in the Product/service layer*

- *Idea Generators*: the citizens are invited to brainstorm about the potential reuse of the data, at the beginning of the development phase
- *Service/product co-creators:* the citizens are invited in the early development of the solution or product to participate in user research, focus groups, and give input on the interface developed (front-end) or the value proposition. Later, user contests and user labs can be used to co-create extra micro-services and functionalities based on the existing solutions.
- *Testers:* the citizens are consulted to test and give feedback on the developed solutions via feedback forms or user labs.

*Participation in the Society layer*

- *Active, aware, and empowered citizens:* thanks to the information or solution provided, citizens become aware of a situation with objective facts, and the initiative's leaders expect them to adapt their behaviours (i.e. mobility, consumption, waste management), and speak their voice (i.e. collaboration with the municipality for policy design) or even become activists (i.e. in environmental issues).

Not all the roles were observed in every initiative. Table 12 presents the occurrences by initiative and reveals patterns by approach. The grey-ticked boxes show the main approach and orchestrated participation forms. White-ticked boxes appear when the initiatives showed an inclination for a secondary approach, and expressed the expectation of civic engagement, but not were proven or facilitated by the initiative.

The open data innovation approach is the only one where participation activities were orchestrated at every stage, from the primary data provision to the outcome. All initiatives belonged to civil society and focused on empowering citizens. In the service innovation approach, initiatives focused on service development and improvement, which is operationalized in participation that mainly brings relevant data and feedback. In the civic innovation approach, the citizen is empowered by digital solutions based on accurate information that should encourage engagement in solving issues, or interactions with the stakeholders enabled by the digital application (the output). The participation is focused on the provision of data and the action in society.

*Table 12. Occurrences of the citizens' roles by initiatives and approaches*

| Initiative name (code) | Open data innovation | Service innovation | Civic Innovation | Primary Data provider | Data supplier | Data aggregator | Developer | Supplementary data provider | Idea generator | Service/product co-creator | Tester | Active, and empowered citizen |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Civic Air Sensors 1 (2CS) | X | | X | X | X | X | X | | | | | X |
| Open Mobility Dev Tools (4CS) | X | | X | | X | X | X | | | | | X |
| Civic Data-driven Governance (15CS) | X | | X | X | | | | | X | | | X |
| Street Damages Reporting app 1 (45P) | X | X | X | | | | | X | | | | X |
| Local information App (1B) | | X | X | | | | | X | | X | X | X |
| Public Transport Route Planner (3B) | | X | | | | | | | | | X | |
| Multi Transport Planner (5B) | | X | | | | | | | | X | X | |
| Pupils on Bike app (9B) | | X | X | | | | | X | | | | X |
| Bicycle Parking app (11CS) | X | X | | | | | | | | | X | |
| Local Consumption app 1 (20B) | | X | X | | | | | X | | X | | X |
| Garage Locator (28B) | | X | | | | | | X | | | | |
| Local Consumption app 2 (31B) | | X | X | | | | | X | | | | X |
| Recycling Together app (38B) | | X | X | | | | | X | | | X | X |
| Handicap-friendly Map 1 (39P) | | X | | | | | | X | | | | |
| Handicap-friendly Map 2 (41B) | X | X | | | | | | X | | | | |
| Small Paths app (46CS) | X | X | | X | | | | | | | | |
| Recycling Centers Locator (49CS) | | X | X | | | | | X | | | | X |
| Touristic Map (40P) | X | X | | | | | | X | | | | |
| Local Services Map (43P) | X | X | | | | | | X | | X | | |
| Talking with the City app (48B) | | | X | | | | | X | | | | X |
| Bike commuters app (7CS) | | X | X | | | | | X | | X | X | X |
| Participatory Urban Planning app 1 (23B) | | | X | | | | | X | | | | X |
| Participatory urban planning app 2 (50B) | | | X | | | | | | | | | X |
| Traffic Counter (8CS) | | | X | X | | | | X | | | X | X |
| Civic Air Sensors 2 (47CS) | | | X | X | | | | | | | | X |
| Environmental Monitoring app (44CS) | | | X | X | | | | | | | | X |

## 4.5. Discussion and conclusions

The purpose of our study was to identify the innovation approaches taken by participatory data-driven initiatives in urban settings and unravel the citizen participation forms within them. The findings section addressed it by developing three innovation approaches, and a new open data value chain aligning the citizen's roles. In this section, we will discuss further the relative importance of open data and participation in the data-driven initiatives, as expressions of innovation practices, the patterns in participation forms by innovation approach, and finally the nature of the value they add. We conclude with the limitations and implications of our study.

A first notable observation was that despite the open data narrative often implying citizen participation and the creation of social value (Attard et al., 2015), in practice, the initiatives combining these features are more the exception than the norm. Communicating about the use of open data is a political or strategic choice, which means that the initiatives that do not see the point in telling open data is reused are more difficult to find. Instrumental participation is becoming a common way of working to develop innovative digital solutions, while participation in society, as an outcome, is more often claimed than proven, as previously stated by Worthy (2015). This maintains a usual confusion in open data between the expected benefits (objectives) and the actual results (Zuiderwijk et al., 2019).

Each innovation approaches privileged certain participatory activities, which limits and opportunities deserved to be discussed. In the open data innovation approach, despite the will to engage and empower citizens at every stage of the project, the leader of 2/CS admitted that the participants are a very specific group of people, with specific motivations (e.g., the open principles, skills development). These initiatives can be compared with open software communities (von Hippel, 2001): the innovation is developed within a group of like-minded and skilled lead-users, capable of reusing data and innovating to meet their needs. As a result, the solutions do not always reach and meet the needs of a bigger group of users. Experts in data, the members can lack topic knowledge and external recognition, which can lead to missing part of their social outcome. The project of 15/CS intended to be more inclusive and developed a methodology to involve larger groups of citizens in the identification of valuable datasets and ideation of reuses. It is the only initiative that let the citizen be an idea generator, maybe one of the most influential roles in the data value chain since it can determine the direction of the development. Giving a voice to the citizens at this stage could be a way to avoid the gap of the lead-user innovation, towards a democratic approach of data-driven innovation.

In the service innovation approach, open data and participation are used as a resource to build new solutions. Whereas a noticeable open data advocacy narrative in six initiatives, the realisation of the open data benefits was never presented as the main purpose. Participation is instrumental, anchored in user-oriented innovation, using for example users tests and feedback (Cooper, 2018), and co-creation workshops (Sanders & Stappers, 2008). The use of open data and the social outcome, however, can reflect a strategy more than a value-based choice, to fit, for example, with a public call for projects or demonstrate the innovativeness of a public organisation. It can raise questions about the genuine intention and feasibility to realise the expected social outcome.

In the civic innovation approach, relevant and transparent data is a key input, which encourages the initiative leaders to engage citizens in data provision (CGD). Only initiatives in this category provided tools to enable participation in society and realise their expected social outcome,

although instrumental participation (collect primary data) and socio-political participation need to be addressed to distinct groups of citizens.

When ranging the participation forms in the open data value chain, we can foresee that they contribute to different natures of value (see Figure 20). Lassinantti (2019) and Abella et al. (2017) had previously argued that different types of values are created and captured by different groups of actors during the data reuse process. We build upon that argument by refining the citizens' roles and the value they contribute to creating.

The roles at the early stages of the data reuse process contribute to an "intermediate" value, the data value. It can be defined as the potential of data (for its intrinsic qualities) to enable practices of value creation by someone for anyone (either oneself or the others) in some given contexts (Cabitza, Locoro, & Batini, 2018). When contributing to the development of the output, a solution for an end-user, the actors create user value: the value of the specific finished product properties and how those properties fit into the end-user behaviours, daily habits (Boztepe, 2007). In the end, the use of the output can generate social value, i.e. improvements in the lives of individuals or society as a whole (Emerson et al., 2000). This can happen, in some initiatives, thanks to the *engagement* of citizens. Ehrlich (2000, p.6) describes *civic engagement* as *"working to make a difference in the civic life of our communities and developing the combination of knowledge, skills, values and motivation to make that difference. It means promoting the quality of life in a community, through both political and non-political processes"*. Civic innovation, in fact, reinvents civic engagement with open data, a concept much more appropriate than the overused and catch-all term "citizen participation" to realise social value with and for the citizens. However, there is room for improvement to move participation from the technical and project layer to the society, as previously argued by Barber & MacLellan (2019).

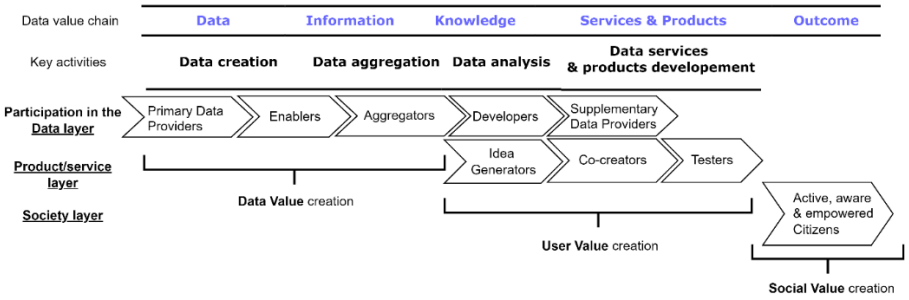


*Figure 20. Types of value created resulting from the citizen participation*

To increase the potential social value of data-driven innovation, a promising path could be to seize collaboratively the social value of open data and therefore reinforce the idea generator role. Cabitza, Locoro and Batini, (2018, pp 3-4) claim that *"the social value of open data can be considered the potential of open data to bring benefit at a societal level (group, community or the whole society)"*. The social value of open data is the extent to which citizens' behaviour, affected by or based on data, have an impact on the social group, or the extent to which the citizens perceive the expected impact.

For research, implications are that to study the value of data-driven initiatives, researchers should investigate the rationales of the projects leaders and not take for granted the

proclaimed outcomes. It could advance our understanding of the realisation of open data expected benefits. However, a process and role perspective can be insufficient to understand the dynamics of value creation and capture between the stakeholders. For practice, the extended open data value chain can help the initiative's leaders to deploy participation activities based on the value they are looking for. The initiative's sponsors should challenge the initiative leaders on measuring their impact by designing indicators relevant to their proclaimed outcome in order to avoid broken promises.

Our findings are limited by their contexts and access to data. Ongoing initiatives could have been missed since they do not always communicate about the use of open data or participatory activities. Participation activities were not directly observed because they were asynchronous with the data collection. The application of the analysis to new contexts could strengthen the theoretical contribution of the three innovation approaches, started in this paper. Moreover, it could be interesting to study the factors (i.e., policies, stakeholders, and available data) that enable the development of innovation approaches, in order to know how to encourage specific approaches according to public priorities. The citizens' roles could be further developed in terms of tools, type of motivations, and profiles required to successfully engage with the "citizens". Research could also explore methods and mediation tools to enable citizens to define the social value of open data, improve the citizens' influence on the value proposition, and move a step towards the realisation of the benefits of open data.

## Disclosure Statement

No potential conflict of interest was reported by the authors.

# Chapter 5. A Typology of Municipalities' Roles and Expected User's Roles in Open Government Data Release and Reuse

## Related publications

**Ongoing research paper** published in conference

Gebka, E., & Castiaux, A. (2020). Open Government Data innovation: A typology of Government and Citizen roles. In *ISPIM Conference Proceedings* (pp. 1–11).

**Full research paper** published in conference

Gebka, E., & Castiaux, A. (2021). A Typology of Municipalities' Roles and Expected User's Roles in Open Government Data Release and Reuse. In H. J. Scholl, J. R. Gil-Garcia, M. Janssen, E. Kalampokis, I. Lindgren, & M. P. Rodríguez Bolívar (Eds.), *Electronic Government. EGOV 2021. Lecture Notes in Computer Science, vol 12850* (pp. 137–152). Springer, Cham. https://doi.org/10.1007/978-3-030-84789-0_10

## Abstract

The purpose of this paper is to identify the roles municipalities take when engaging in Open Government Data (OGD) and the expectations of user's roles they imply. According to the output delivered, the user can relate to data or data-based solutions. OGD is data released by public organisations to enhance government transparency, innovation, and participation. The realisation of those benefits involves different roles, from providing data, developing solutions, to using them for a certain purpose. However, the definition of the municipalities' and users' roles in that context is unclear, which can impact the realisation of the OGD benefits. This study uses Role Theory's concepts as an analytical lens, following the Design Science Research approach to create a typology. We conducted a hermeneutic literature review, identified, and analysed 52 papers, to build a typology of the municipalities' roles based on the goals, tasks, output delivered, and the expected users' roles they generate. It results in seven classes of roles coming in pairs. We tested the typology on empirical cases: the 28 Belgian and 158 Swedish municipalities engaged in OGD. Five role pairs were encountered in the empirical cases, and two occurred only in previous literature. The typology can help municipalities to understand how their role choice calls for a certain type of users that cannot be generalized as a "citizen". Role Theory opens new perspectives of research to understand their interdependence and raises fundamental role-related questions that should be given the same importance as technical and technological challenges.

**Keywords:** Open Government Data, Municipalities, User, Roles, Typology.

## 5.1. Introduction

In 2013, the European Commission adopted the *Public Sector Information* (PSI) Directive (Directive 2013/37/EU), which encourages public organisations to share their information and data for reuse. The idea is that Open Government Data (OGD) is funded by public money and can generate social and economic value (Attard et al., 2015), therefore it should be made accessible to all. OGD is data released by a public organisation, the publisher, for secondary use, by a user, without any restriction or limitation in use. For a public organisation, the benefits of publishing OGD can be better transparency and accountability, innovation and improved efficiency, and/or an increased engagement and participation in governance (Charalabidis et al., 2018a). The municipalities play an important role because they own plenty of data (e.g., on transport, pollution, geographic data) (Berends et al., 2020) that can supply the regional and national portals and are amongst the most reused data sets (Berends et al., 2017).

However, publishing OGD is new for most municipalities and creates ambiguities in terms of role and scope of action, compared to their traditional role. To serve the common interest,


H. J. Scholl et al. (Eds.): EGOV 2021, LNCS 12850, pp. 137–152, 2021.
https://doi.org/10.1007/978-3-030-84789-0_10

several institutional paradigms co-exist, have different views about what is at stake, and bring different answers on how to achieve efficiency, accountability, and equity (Stoker, 2006). They create a variety of citizens' roles and modes of interaction (Vigoda, 2002). In the same vein, what is included in the actors' roles to realise the OGD benefits is variable. For example, the public organisation can be limited to the publisher role (Carrara et al., 2015) or considered as a data user (Mergel et al., 2018; Pereira et al., 2017), a duality that is under-researched (Mesquita et al., 2020). The citizen, which involvement in OGD is lacking empirical evidence (Safarov et al., 2017), is broadly assimilated to a group of data users, or indirect beneficiaries (end-users) depending on intermediaries, developers or companies, to benefit from OGD (Abella et al., 2019; Safarov et al., 2017).

There are different approaches of role -classifications in OGD literature: a process approach based on data value chain (e.g., Attard et al., 2015; Carrara et al., 2015), the data provision (e.g., Sieber & Johnson, 2015), the reuse process (e.g., Abella et al., 2019; Lassinantti, 2019) or the data ecosystems (e.g., (Dawes, Vidiasova, & Parkhimovich, 2016). They help to model roles to reach data value creation in an ideal world, as things happened in a continuous process and perfect interactions. However, they provide little insight into the influence of other roles (municipality *and* publisher) and projected expectations towards others. The purpose of this research is to investigate the possible roles of the municipalities, considering that they can be publisher and user, and deliver a certain type of output, data or solutions, for others to use. They generate expectations towards the user. To this end, we develop a typology of roles through the lens of Role Theory's concepts, following a Design Science Research approach. The research questions guiding the study are:

1. *What are the possible municipalities' roles within OGD release and reuse?*
    - *What expected users' roles are implied by the municipalities' roles?*

The typology should help to differentiate the municipalities' approaches to OGD, roles, and type of users they call for, and raise role-related issues that can impede the realisation of benefits. The paper is structured as followed: the background introduces role ambiguities and the Role Theory's concepts used to develop the typology. Then, we explain the research approach, present the findings, discussion and conclusion.

## 5.2. Background

### OGD Roles, Outputs, and Ambiguities

To generate benefits, the data needs to be made available. Then, it needs to be accessed by users, handled and repurposed to give it a new use and a broader value (e.g., insights, visualisations, or information solutions) (Lassinantti, 2019). Therefore, in OGD, the most comprehensive and acknowledged roles are (1) the publisher, the actor who publishes data, and (2) the user, the actor who makes a secondary use of it. Publishing and reusing data is not easy and can require new roles and intermediaries between the publisher and the user. For example, a publisher might need the help of a portal provider to structure the data released on the web. A user might need the help of an enabler who provides tools or visualisations to facilitate data reuse in context (Berends et al., 2017). The user can also have the expertise to develop solutions for others, make decisions, participate in governance processes, or benefit from new digitalized services. A complexity of OGD lies consequently in the role coordination around an output, to reach a higher purpose. The output delivered by the municipality can vary

from raw data in an excel spreadsheet, datasets on a portal, visualisations and charts, or a complete information solution. To use the output, the user can need analytical skills or just be able to use a computer or smartphone. The expected user is moreover coloured with citizen's roles, to participate in the discussion, exchange of ideas and decision-making process with the government (Harrison et al., 2012), enabled by OGD. Users can collaborate with decision-makers to create solutions based on OGD that will be implemented in the city (Noveck, 2009) or consume the developed services.

The extent to the municipalities engage with OGD, therefore, has an impact on the user expected tasks and activities, to realise the higher purpose. When the municipalities' role and expectations are not in line with the user capabilities, resources, and motivations, there is a risk that the expected benefits will not be realised. To understand the relation of the actors through their role, Role Theory provides a relevant lens.

### Role Theory

The concept of roles is widely used in the area of social sciences to explain human and organisations' behaviour patterns. It assumes that people have social positions and hold expectations for their own behaviours and those of other persons, according to their role (Biddle, 1986). The concept of role can be extended to the concept of actors, understood as persons, entities, or organisations. "Role Theory" is a catchphrase grouping different research streams that study roles, with different perspectives and terminologies. In our context, three approaches can complement each other to understand roles in OGD.

The *functional approach* has focused on the behaviours of individuals occupying a social position within a stable social system. This perspective suggests that individuals within their social systems are taught norms and are expected to conform to those norms and sanction individuals who do not (Biddle, 1986). In the traditional conception of roles of the citizens and the government, as voters vs. elected representatives, or public managers vs. users, the boundaries of the roles are shared, normative. Clear expectations prescribe and explain behaviours. In OGD, the municipality might expect the public to reuse OGD as a tool to monitor public action because it is the duty of the citizens.

The *symbolic interactionist approach* assumes that roles are not consequences of one's position in a social structure, but that an actor can change its position as roles are context-specific (Anderson, Havila, Andersen, & Halinen, 1998; Ashforth, 2000). Roles are created through interactions with others, they are emergent and negotiable (Biddle, 1986). Network and innovation studies have focused on a processual aspect of roles that describes what actors intend to do. It implies that actors' roles can also be used for granting access to important resources. The roles are products of actors' interpretations of situations (Anderson et al., 1998). In OGD, it explains how the publishers rely on the users to innovate, providing creativity and skills that it has not internally. They bring essential resources to realise the expected benefits. The first two approaches of Role Theory can be considered normative.

Finally, *the task-based approach*, as suggested by Nyström et al. (2014) for the study of open innovation networks (in Living Labs), looks at individuals using an ideal role to achieve a certain goal. The actor's role is created through action: the goal and the related tasks determine the role, which resources are allocated, and which actors are teamed up with. The roles emerge in the innovation process and, as the process is open, roles are not predefined. The same actor can have different roles. This approach is particularly relevant for our study, as the OGD reuse

is an open innovation process by principle. It gives the appropriate flexibility to create role categories, necessary in OGD, where the roles are emerging and varying with the local contexts.

In sum, the **roles** are functions, tasks, and behaviours expected of parties in particular positions and contexts (Herrmann et al., 2004). A **role-set** of an actor is related to the expected acting out of a role: required duties, activities, standards, objectives and responsibilities (Heikkinen et al., 2007); a role-set emphasises the interdependence between the actors within a certain structure: the actor who sends the role through expectations, and the focal actor, who receives the role (R. L. Kahn, Wolfe, Quinn, Snoek, & Rosenthal, 1964). The **role expectations** can be explicit (e.g., job description) or not (e.g., informal notions, agreements) (Herrmann et al., 2004).

Herrmann et al. (2004) argue that the repetition of social interaction patterns until they can be anticipated, based on patterns of expectations, gradually develops new roles. They add that the development of roles is accompanied by role-mechanisms, i.e. interaction patterns for role-taking and role-making. The role-taking is acting with respect to the expectations, which can be potentially enforced by sanctions being imposed on the actor (Herrmann et al., 2004). The role-making is how a person lives a role, and how she transforms the expectations into concrete behaviour (Herrmann et al., 2004). Role expectation is, therefore, a key concept in Role Theory. In the functional approach, the focus is on the given expectations (role-taking). In the interactionist approach, roles are emergent and negotiable, consequently, the actors interpret, organise, modify the expectations (role-making). In the task-based approach, the roles are defined by the network's goal and needs depending on the situation, there are constant occurrences of role-making and role-taking (Nyström et al., 2014).

The study of role-sets and expectations can help to understand dysfunctions in the role-taking or making and its impact on the process. For example, role ambiguity is the lack of clarity of role expectations and the degree of uncertainty regarding the outcomes of one's role performance (Miles, 1977). Role overload occurs when a person is faced with too many expectations (Biddle, 1986). Role malintegration occurs when interdependent roles do not fit well together (Biddle, 1986). In this study, we use concepts of Role Theory to develop a typology of roles and discuss the role-related issues.

## 5.3. Research approach

The research approach to develop the typology is based on Design Research Science (DSR). This paradigm of research aims at developing solutions (artefacts) meeting defined goals, that contribute to the scientific knowledge base (rigour) and provide utility in the environment (relevance). To reach that purpose, a research project should, in as many iteration loops as needed, follow 6 steps: *identify and motivate the problem domain (1), define the objectives of the solution (2), develop (3), demonstrate (4), evaluate (5)*, and finally *communicate (6)* the results to the audience (Peffers et al., 2007).

*The problem (1)* we identified is the lack of clarity in the role of the municipality and citizens identified in the literature. Accordingly, *the motivation and objective (2)* of this study are to design a typology of municipalities' roles when dealing with OGD, enlightening the expectations they project on the users. Such a typology can be used as a tool to diagnose the OGD approach of municipalities and help them to raise the critical questions of new role integration that OGD implies for all the actors.

*To develop the typology (3)*, we used the method suggested by Nickerson et al. (Nickerson et al., 2013) anchored in DSR. A typology is a system of conceptually derived grouping. The method of Nickerson et al. (Nickerson et al., 2013) starts with the determination of the meta-characteristic, the most comprehensive characteristic that will serve as the basis for the choice of dimensions and their characteristics in the typology. In this research, it is the interdependence between the role of the municipality releasing or reusing OGD and the user of the provided output. Each characteristic should be a logical consequence of the meta-characteristic, in our case, relevant concepts from Role Theory: components of the role-set (e.g., tasks, responsibilities), output, and role expectation. The typology development, made in iterations, combined a conceptual method (deductive: conceptualizing the dimensions of the taxonomy without examining actual objects) and an empirical method (inductive: identifying a subset of objects that we want to classify).

The typology development was based on previous literature as it helped us to access more cases and potentially identify more roles than an empirical method. To review the literature, we used the hermeneutic method of Boell et al. (Boell & Cecez-Kecmanovic, 2014), consistent with the DSR approach. It allows a progressive and critical understanding of a body of literature, through two intertwined circles of research that can be repeated several times: the searching and acquisition circle, and the analysis and interpretation circle. The software NVivo was used to store, code, analyse, and sort the selected papers. We conducted three cycles. We used the databases Google Scholar and Science Direct and, in the first cycle, the keywords "role", +"Government", +"Citizen", + "Open Government Data", then "Task", "Actors". We selected empirical papers wherein the words "role" and "open (government) data" appeared in keywords, abstract or body text, with the cities as a context. It resulted in the first iteration of the typology, with roles based on the level of engagement of the municipalities and the user's roles as citizens. For the second cycle, we extended the research with citation tracking (backward and forward literature search) and keywords of concepts discovered in the papers to find new papers. We came to the second iteration of the typology, but noticed overlaps between roles and dimensions due to conceptual ambiguities in the chosen characteristics and reuse of role classes of previous research. We detached ourselves from previous role classes and re-focused on the key concepts of Role Theory for the third iteration. We analysed one more time the goals, tasks, outputs and expectations to let emerge classes of municipalities' roles, and sort the empirical literature. In the third and last cycle of literature search, we focused on acquiring empirical and conceptual papers for the less covered roles (e.g., OGD + "commercial reuse", OGD + "citizen participation") and refined the typology in its fourth iteration. Conceptual papers helped us to strengthen the logic of the typology, i.e. its dimensions in accordance with the research purpose, and to understand the school of thoughts of the empirical studies. We coded and used 40 empirical papers and 12 conceptual papers to develop the typology. The roles are more often indirectly presented than explicitly researched in the current OGD literature. They were not found with the names and the combinations presented in the typology, as they result from the analysis and understanding of the researchers.

*To consolidate the typology and demonstrate (4)* its relevance in practice, we collected primary empirical data for two national cases, Belgium and Sweden, wherein we analysed the municipalities engaged with OGD. Two cases allow better generalization. Those countries were

chosen for their different level of maturity in OGD, according to the European Maturity Report[7], Belgium and Sweden being respectively follower and fast-tracker. Both countries have translated the European PSI Directive into their laws and encourage their municipalities to publish data. Belgium counts 581 municipalities, and Sweden 290. By comparing national data portals and lists of publishers, we identified 28 municipalities in Belgium and 158 in Sweden publishing or reusing OGD at the time of the data collection (January-February 2021). We applied the typology to each of them and used a directed content analysis (Hsieh & Shannon, 2005). We analysed the websites, portals, and all type of reported off-line activities that constitute the output and create a channel for interaction between the municipality and the user. We identified and tabulated their stated goals, expected users, activities and tasks through the output delivered to sort them per role classes. At the end of the analysis, *we evaluated (5)* the typology based on the ending conditions of Nickerson et al. (Nickerson et al., 2013) and concluded that they were met. The publication of the findings in that paper is part of the *communication (6)*.

## 5.4. Analysis and findings

This section presents the typology of the municipalities' roles and expected user's roles (Table 13). The first column lists the municipalities' role, characterized by a set of goals, tasks, and output. In pair come one or several expected users' roles that interact with the municipalities in unique ways. Key references and the number of occurrences are specified. The sum of occurrences is higher than the total number of municipalities analysed, as the combination of roles is possible, which is further presented in the findings. Each pair of role class is subsequently elaborated with empirical examples.

[7] https://www.europeandataportal.eu/en/dashboard/2020

*Table 13. Typology of the Municipalities' Roles and Expected User's Roles*

| *Municipalities' Role* | *Role-set (G: Goal, T: Examples of Tasks and Activities)* | *Example of Output Delivered to the User* | *Expected Users' Roles* | *Key References and Number of Occurrences in Belgium (BE) and Sweden (SE)* |
|---|---|---|---|---|
| **The Compliant Data Provider** | G: Compliance with the law and public values (e.g., transparency).<br>T: Answer requests; send data on demand, without specifically publishing it. Open "on-demand". | Unstructured data, formatted for internal use, delivered on-demand by email, or imported web pages showing a selection of data, with low engagement to improve it. | **The Data Hunter:** The users are experts in data reuse and know what data they need to satisfy their goals (find information, innovate). | (Agrawal, Kettinger, & Zhang, 2014)<br>BE : 0 -- SE : 127 |
| **The Partner** | G: Support the development of new services and public value.<br>T: Participate in collaborative processes (for innovation, governance). | Client briefing, guidance and feedback, expertise for the under-development solutions, funds. | **The Partner** in governance or innovation processes, as a project led by third organisations or researchers. | (Ruijer et al., 2017) (Conradie, Mulder, & Choenni, 2012) (Toots, McBride, Kalvet, Krimmer, et al., 2017)<br>BE: 0 -- SE: 0 |
| **The Stand-Alone Publisher** | G: Openness, transparency, economic growth, innovation, participation (multiple and broad goals).<br>T: Publish data on a website or portal, with a supply-driven and often scattered approach. | OGD portal or website, with data as the main content. Do not always provide contact forms. | **The Rare Bird:** the users are the expert and can conduct all type of activities required to reuse data (searching, finding, cleansing, enriching, combining, visualising, developing solutions). The provided data is believed to be enough. | (Ham et al., 2019) (Attard et al., 2016a) (Attard et al., 2016b)<br>BE: 23 -- SE: 24 |

| | | | | |
|---|---|---|---|---|
| **The Dedicated Publisher** | G: Make the data appealing to reach the above-mentioned multiple goals.<br>T: Make the data easy to use, accessible, provide extra tools and resources, publicize the data released and reuses (apps), release new data on a regular basis. | OGD portal with extra content and functionalities to provide support and feedback (tools, tutorial, documentation, API's, technical standards, selection of best cases of reuses, feedback form). | **The Data Analyst and Developer:** develop new applications and solutions based on data, exchange ideas with the open community members, gives feedback to the publisher, analyse data for monitoring the public action. | (Attard et al., 2015) (Abella et al., 2017) (Khayyat & Bannister, 2017)<br>BE: 4 -- SE: 6 |
| **The Enabler** | G: Make the data reused to outsource innovation, service development, or to solve identified problems.<br>T: Raise the awareness and capabilities of the ecosystem, identify public issues, provide means and call for actions. | Beside the portal, organises hackathons, training programs, workshops, ideation platforms, innovation contests, guidelines, policies, places of collaboration and exchanges (ideas, resources). | **The Ideator, Innovator, Co-producer, Co-implementer:** share community needs, provide ideas, prototypes of solutions, applications, technical know-how, creativity, solve public challenges. | (Foulonneau, Martin, & Turki, 2014) (Sangiambut & Sieber, 2017) (van Loenen, 2018)<br>BE: 4 -- SE: 2 |
| **The Solution Provider** | G: Provide public e-services and digital tools based on data for the citizens.<br>T: (Co-)fund solutions, develop tools. | Dashboards, policy evaluation tools, improved public services. | **The Smart Citizen:** use enhanced public services, make informed decisions, participate in governance processes. | (Mergel et al., 2018) (Barns, 2018) (Matheus, Janssen, & Maheshwari, 2018)<br>BE: 8 -- SE: 2 |
| **The Orchestrator** | G: Coordinate means and strategies together to reach a vision and purpose (smart city, data ecosystem).<br>T: Develop policies, strategies, tools. | Living labs, policies, change management strategies, global approach of data production, management and reuse, pilot projects. | **The Data Producer:** generate data that is reused by the municipality.<br>**Innovator, Smart Citizen** | (Bakici, Almirall, & Wareham, 2013) (Gupta, Panagiotopoulos, & Bowen, 2020)<br>BE: 0 -- SE: 0 |

**The compliant data provider** goes for the simplest way of providing data by responding to external data demand or pressure. For example, in Sweden, municipalities can freely upload their data on a platform managed by an association (Kolada), to allow the citizens to compare their performance. Nine municipalities refer to that website under a page labelled "Open data" or "PSI data", 118 just imported a script from Kolada's website that displays a selection of datasets, and only 13 of those provided contact details. For very broad goals (e.g., "*promote participation, democracy and growth*"), the municipalities provide what is strictly necessary and create huge expectations on the users. They are true **data hunters**: to find data, they have to be ready to explore websites and dig into unstructured datasets. These roles are not very documented in research, since they bring little knowledge about OGD, but they are the case of most municipalities that have not intention to invest time and resources in OGD.

Municipalities and citizens can both be **partners** and collaborate in projects led by third parties for new service development (user-centred approaches) and governance (decision-making, policy-making, monitoring). As the goal of the data reuse is clear, the process led and facilitated by a third party, municipalities and users can be called for expertise other than data analytics: field knowledge, community needs, creativity, voices in public debate. The expectations are directed to a role in society (user of public service, citizen), instead of technical expertise. However, the collaborative processes identified in our literature review were research-led projects and no cases were reported in the empirical material. These roles are temporary and reactive to external impulsion.

The **stand-alone publisher**, unlike the compliant data provider, shows intentions to join the OGD movement. Still, it has either not stated or very general goals (e.g., *"foster innovation and the development of applications", "promote the participation of all"*). The output can consist of only one to five datasets on a regional or national portal, a catalogue on the website, or owned external portals. The publisher follows a supply-logic: what is in house, cleanable and openable, or thought as a priority by the municipality, is released (Ham et al., 2019). The publisher expects the rest of the reuse to be undertaken by the user, a **rare bird**. Despite the little resources provided and lack of channels to interact with the publisher, he would navigate between portals, find and reuse data, develop new solutions, and participate in governance processes and public debate, empowered by the information he would have extracted himself (Attard et al., 2016a, 2016b).

The **dedicated publisher** has understood data only gain value being reused and tries to make it appealing. A common strategy to enhance reuse is to publish as much data as possible on an owned portal, which accessibility and user-friendliness depend on the functionalities offered by the portal provider. Besides extra information, technical documentation, portals can also include tools to visualise data and improve the data reusability, discoverability (Abella et al., 2017), even extent the technical development of data (e.g. linked open data, 1 municipality, BE). The municipality is aware of the difficulty of reusing OGD, but the output (portals, tools), can still be complex for a lay user. Part of the support is addressed to the **developers and data analysts**, who still are expected to develop services and solutions for the community.

As an **enabler,** the municipality moves beyond the publishing activities, seeing the need for a more interventionist approach. It enables the actors of an ecosystem to realise the benefits of OGD. It has still no control over the developed services but shows more leadership because its goal is to create public value, stimulate innovation to solve public issues and meet the citizen's needs. To achieve that, the municipality undertakes an enabling role that can be oriented

towards the capabilities of the actors, the functioning of the ecosystem, or the motivation to solve specific issues. This role was well documented in previous research, as it was an ideal to aim. In practice, only six municipalities took that role, and organised workshops and hackathons. They serve as places to meet, raise public needs, exchange ideas and develop prototypes based on data. The municipality expects the citizens to be an **innovator, ideator, co-creator** of new services of public interest. Interestingly, that role can be limited in time, as it relates to a specific project. Three municipalities organised a single hackathon or similar, in two cases, funded by European projects, two ran yearly hackathons but stopped due to the pandemic, and one stopped due to the lack of sustainable results but is considering new ways to energize the user community.

When the municipality reuses its data, it becomes a **solution provider**. It can improve its processes and digitalized its public services (Mergel et al., 2018). Innovation and reuse are internalized. The user is a **smart citizen** who gets tools like dashboards and visualisation of key facts to monitor the performance of the city (Barns, 2018; Mergel et al., 2018) (9 municipalities), make better decisions, and trust the government (Matheus et al., 2018). Municipalities try to reduce the information asymmetry and encourage "citizen participation" in the decision-making process, but in the empirical cases, dashboards do not come along with off-line governance processes. It is more an open window on key figures and performance. Information can also be developed into applications (2 municipalities, e.g., an app for parking spots, cemeteries). Open data is integrated to the municipalities' core activities, and the key outputs are tools for transparency and digitalized information services.

In the role of **orchestrator**, the municipality put means together to reach a certain purpose (under its control), the transformation of the city into a smart city (Bakici et al., 2013), a smart data ecosystem (Gupta et al., 2020) or a platform model of data-driven public services (Cordella & Paletti, 2019). The municipality takes the lead to fulfil its goal and strategy. The OGD is a piece of a larger program, which can enrol any of the municipality and citizens' roles previously cited. The difference is that the goal of the municipality is clear and its maturity in reusing data enables it to make strategic choices instead of experimentations with OGD. These roles were not observed in the empirical material.

We observed municipalities **combining roles**. One municipality (SE) was a stand-alone publisher (basic data catalogue) and temporary enabler (European project addressed to the citizens), three were stand-alone publishers and solution providers (app of visualisation of the key figures). Two added to the latter combo the enabler role (organisation of single or yearly hackathon). Finally, the dedicated publisher role was combined once with temporary enabler (unique hackathon), once with solution provider (app of cemeteries, visualisation of key figures), and twice with solution provider and permanent enabler (yearly activities with the users).

## 5.5. Discussion

The purpose of this study was to create a typology of roles for the municipalities within the spectrum of data release and reuse. Previous research on OGD roles took a technical perspective, dividing them between a succession of tasks and operations to reuse data and create value (e.g., Abella et al., 2019; Attard et al., 2015; Carrara et al., 2015; Dawes et al., 2016; Lassinantti, 2019). Sieber & Johnson (2015) take a government-citizen perspective, limiting the scope to data provision. We were missing a more comprehensive definition of roles, considering that in our context the data provider is also and first a municipality. They

have to define the limits of their new OGD roles, which creates expectations towards the user. In this section, we further discuss the nuances between the identified roles and role-related issues. Then, we highlight discrepancies between the given importance to the roles identified in the literature, and their occurrences in practice.

Through the lens of the Role Theory's concepts, we can highlight three ways municipalities approach the OGD roles, as shown in Figure 21, the partner and the compliant data provider are in reality in a focal role (R. L. Kahn et al., 1964). The role sender can be a citizen asking for data, a supra-government or an institution that push the municipality in the role of data provider or partner. The stand-alone, dedicated publisher, and enabler, believe in the benefits of data and embrace a new but *distinct* role, imposed by the new activities coming along with data release. The stand-alone is doing what Sieber & Johnson (2015) call "data-over-the-wall". The dedicated, with better tools, expects the user to provide new public services themselves ("Do-it-Yourself Government"; Linders, 2012). The enabler makes its resources and knowledge available to the public, provides support to foster greater public value, embedded in an ecosystem view, but without active involvement in the development of solutions ("Government as a platform"; Linders, 2012). However, the solution provider and orchestrator integrate data reuse in their operations to deliver digitalized services or improve public management.

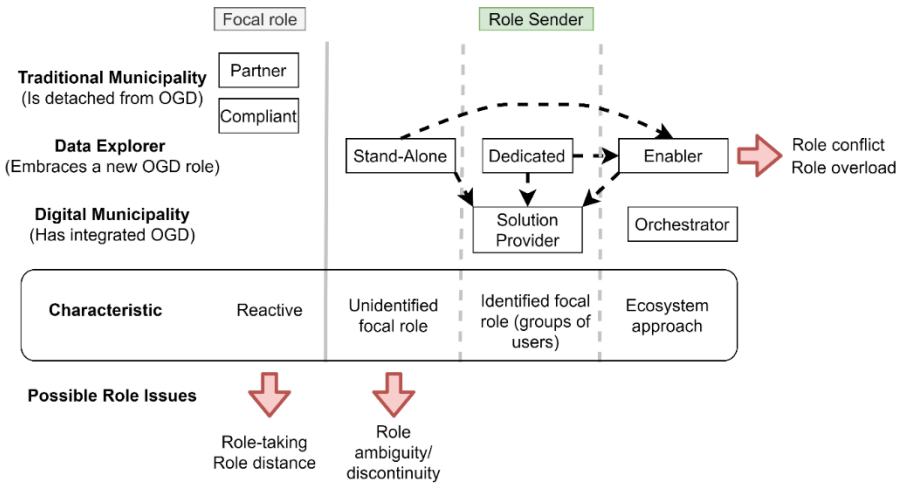


*Figure 21. The municipalities' roles combinations, characteristics, and issues*

More complex municipalities' perspectives on OGD show through the combination of roles. The stand-alone/service provider goes for the quick-wins: basic portals for the experts, easy to read visualisations for the other and transparency. The stand-alone /enabler/solution provider adds user interactions, which is interesting when the enabler role is recurrent: basic portals are balanced with regular hackathons. The stand-alone/temporary enabler is an experimenter, and the dedicated publisher/temporary enabler chooses to invest in a rich portal more than in user interaction. The dedicated publisher/solution provider is a logical combination that comes with experience and time: the more they publish data, the more they see opportunities for new

services. The dedicated publisher/permanent enabler/service provider is probably the most engaged in OGD, intending to become a "data-driven" city.

Role Theory also allows identifying issues and ambiguities that can impair the realisation of OGD benefits, as presented in Figure 19. In a focal role, the municipality is in a position of role-taking that can be potentially enforced by sanctions (Herrmann et al., 2004) (laws regarding OGD), which can increase a role distance (Herrmann et al., 2004) and the absence of interest in providing data. The stand-alone publisher has an unidentified focal role ("anyone", role ambiguity), without interaction with the user. Consequently, the output might meet the requirements of no one: there can be a role discontinuity (Biddle, 1986) between the publisher and the user in terms of skills, objectives, and data available. The stand-alone, dedicated publisher, and enabler can raise a role conflict (Miles, 1977) for the municipality. If the data reuse is entirely delegated to independent users, can the public interest be guaranteed? Who is responsible for? The provided data (limited in quality, interoperability) combined with the high expectations (create applications for the citizens) can also generate role overload (Biddle, 1986), a role wherein few users recognize themselves. In the context of OGD, where roles develop and are performed through the provision and the use of a complex output, a functional approach, which can fit with the organisational culture of traditional public organisations, can impede the OGD reuse and realisation of its benefits.

In terms of occurrences, five of the seven roles were taken by the municipalities. The most represented role is the compliant data provider, in Sweden. Importation of Kolada's datasets on their website is considered sufficient to comply with their obligation of delivering data for transparency. It is reasonably arguable, however, that this operation results in the claimed objectives ("*promote participation, democracy and growth*"), as none of these pages reported initiatives making use of the data.

The most documented role, in our literature review, is the enabler (13 papers identified). According to research, the municipalities can, among others, provide training to build data-literacy (Foulonneau et al., 2014), hackathons, workshops, understand the needs of the users (van Loenen, 2018), develop policies that focus on the availability of resources and good governance (Zuiderwijk et al., 2016), encourage participation and balance the benefits of all type of users (companies, citizens, social organisations) (Abella et al., 2017). In practice, the instances of enabler's roles are focused on the organisation of hackathons, which remained a unique or abandoned experiment for four cities. The lack of resources can be a reason, as two projects relied on external funding, but also the question of the role attribution: is that the role of the municipality?

Finally, the roles of partners and orchestrators did not occur. The partner, as understood in the literature, is, in fact, a reactive role: the municipality joins an experimental process that intends to create value with data, led by other organisations. It depends therefore on external impulsion. The orchestrator is a role that requires a certain maturity with data and digitalization, together with ambition, vision, and resources. Capitals can more easily gather these conditions, such as London (Gupta et al., 2020). Brussels and Stockholm appear to be dedicated publishers. They invest more in appealing and well-provided portals than in the integration of the data in their operations with a coordinated vision.

The "citizen" turns out to be an elusive role that does not help to grasp the nuances between the expectations generated by the different municipalities' roles. This lack of clarity is

substantial in the empirical material. The citizen's role is not stated ("*Open data is available for anyone*") or implied (transparency for the citizen to monitor) by the compliant data provider. For the stand-alone publisher, the citizen is the data user or user of future applications. The dedicated publisher addresses clearly its output to experts, while the enabler can see the citizens as idea providers. Solution provider and orchestrator, on the other hand, have a clear objective and output, and therefore defined user groups of democratic processes or digitalized services.

Implications for practice are that an interactionist and task-based approach would enable the actors to shape their role in relation to each other and the resources available, through time and experience. "Open" should not mean "abandon" of and disconnection with the users. Until then, if the municipality can combine the role of solution provider with other roles, it can increased the perceived value of OGD for both municipalities and lay users, and avoid a role conflict that could cease the municipality's engagement in OGD. For research, Role Theory offers a new perspective to understand the difficulties and barriers faced by the actors in OGD. It raised fundamental role-related questions that should be given the same importance as technical and technological challenges.

## 5.6. Conclusion

Municipalities are encouraged to publish OGD, which is for most, a new role. There is no definition or limits regarding what a municipality is supposed to do and deliver to the user. However, the way they frame their role creates expectations towards the user. The interdependence between roles and generated expectations is not researched in previous literature, as roles are defined from technical and process perspectives. The expectations are ironed out or ignored, although it is a central concept in Role Theory. This study is the first to use Role Theory's concepts in the OGD field. With that theoretical lens, we develop a typology of municipalities' roles, coming in pair with expected users' roles. The way each pair interacts is unique. Seven pairs were identified and applied to the municipalities of two national cases, Belgium and Sweden. One of the main findings is that municipalities can detach themselves from the OGD roles, create new ones focussing on the provision of data and support for the user, or integrate OGD as part of their main operations. The aim of the typology is not to suggest an ideal-type role or path to follow, but to allow identifying municipal approaches to OGD and highlighting the possible role-related issues that could impede the realisation of OGD benefits. The main contribution lies in the originality of the theoretical lens used, which opens a new perspective to understand the difficulties and barriers faced by the actors in OGD.

The study has limitations. It has focused on the municipalities and their users. It is however evident that they do not work in closed environments and other factors and actors influence the way they perceive and deliver output. The analysis of the data, information, tools and activities provided by the municipalities, pictures a situation at a time. It does not reveal uncommunicated intentions or future projects that could affect the role classification of the cases.

Future research could use the typology to conduct case studies and explore the factors or conditions that encourage the municipalities to take certain roles. With a time perspective, future research could explore what experiences and learnings make them evolve between roles, combine them, or stop and leave OGD. The typology could also be compared with the user perception of the municipality role, and explore how role-making interplays with role-taking for both users and municipalities.

Acknowledgements. The study was performed with financial support by the FEDER and the Walloon Region (H2020), as part of the project “Wal-e-cities”.

# Chapter 6. Methods to Capture User Information Needs: Design Principles for Open Data Intermediaries and Data Providers

## Related publications

**Poster** published in conference

Gebka, E., Clarinval, A., Simonofski, A., & Crusoe, J. (2019). Generating Value with Open Government Data: Beyond the Programmer. In *2019 13th International Conference on Research Challenges in Information Science (RCIS)* (pp. 1–2). IEEE. https://doi.org/10.1109/RCIS.2019.8877054

**Ongoing research paper** published in conference

Gebka, E., Crusoe, J., & Ahlin, K. (2020). Open Data Reuse and Information Needs Satisfaction: A Method to Bridge the Gap. In S. Virkar, M. Janssen, I. Lindgren, U. Melin, F. Mureddu, P. Parycek, … H. J. Scholl (Eds.), *Proceedings of Ongoing Research, Practitioners, Workshops, Posters, and Projects of the International Conference EGOV-CeDEM-ePart 2020* (pp. 41–50). Linköping. Retrieved from http://ceur-ws.org/Vol-2797/paper3.pdf

**Full research paper** published in Journal (open access)

Gebka, E., Crusoe, J., & Ahlin, K. (2021). Methods to Capture User Information Needs: Design Principles for Open Data Intermediaries and Data Providers. *Data Science Journal*, *20*(1), 37. https://doi.org/10.5334/dsj-2021-037

**Popular Science Article** (open access)

Gebka, E., Crusoe, J., & Ahlin, K. (2021b). To Capture the Information Needs of End-Users of Open Data: Guidelines to design methods. Retrieved from https://www.diva-portal.org/smash/get/diva2:1621345/DATASET01

## Abstract

Data providers share open government data (OGD) to be transformed by intermediaries into products and services (solutions). OGD is believed to lead to many benefits. However, OGD is not reaching its expected level of reuse, which can come from a lack of awareness, interest to fulfil the user's needs, or novel perspectives to understand them. This paper presents a set of design principles to develop tailored mixed methods that capture activity-based information needs of users which could be satisfied by building information products based on OGD. The produced insights can help data providers and intermediaries to realign ideas of solutions to the user's information needs. The set of nine design principles are developed using design science research and are based on previous research and empirical testing. They have been implemented with two groups of users and three groups of data providers as participants of, respectively, face-to-face and digital workshops. The design principles and the produced insights were evaluated with practitioners. Implications for practice are that starting with the users' information needs can open a broader range of solutions and potential paths of OGD reuses, while following the design principles can help the practitioners cope with the fuzziness of the information needs and ideation process. For research, we propose a novel method that goes beyond the exclusive data provider-intermediary interaction to study new paths to improve the realisation of OGD benefits.

Keywords: Open Government Data, Information Needs, Design Principles, User Involvement, Innovation Method

## 6.1. Introduction

Public organisations, in the role of data providers, are releasing open government data (OGD) (Davies, 2010). OGD focuses on public sector information associated with agencies at federal, state, and local levels that collect and record data in order to monitor or inform social arrangements relating to human populations, economics, and public services (Open Knowledge Foundation, 2015). The providers produce and provide the data to others without any restrictions on its use or redistribution (Janssen et al., 2012). The providers hope that OGD will lead to better transparency, citizen engagement, and innovation (Charalabidis et al., 2018b). On the other side, intermediaries develop information products or services (solutions) on the data for others (Davies, 2010; Janssen & Zuiderwijk, 2014; van Schalkwyk et al., 2016). Users who are seeking information can use these solutions to satisfy their information needs, which can be experienced when they attempt to satisfy a primary need and encounter a gap in their knowledge (Belkin & Vickery, 1985; T. D. Wilson, 1981). Users can draw facts from data provided by data providers (Davies, 2010). However, the rawness of the data can make it difficult to use it for any meaningful purpose relating to the life events or decisions of users (Weerakkody et al., 2017). Intermediaries play a vital role in assisting users in interpreting and acting on information drawn from the data (Hunnius & Krieger, 2014), adding value to the data

First publication in *Data Science Journal*, *20*(1), p. 37 (2021), available in https://doi.org/10.5334/dsj-2021-037 

(Janssen & Zuiderwijk, 2014), and increasing its accessibility and utility (van Schalkwyk et al., 2016). However, OGD is not reaching its expected level of reuse (Zuiderwijk, Susha, Charalabidis, Parycek, & Janssen, 2015), which can come from a lack of awareness or interest (Hellberg & Hedström, 2015; Huang, Lai, & Zhou, 2017). Data providers attempt to tackle the low awareness and engage intermediaries to develop innovations with OGD via innovation contests, like Hackathons (Johnson & Robinson, 2014). However, satisfying users' needs is the developers' fourth motivation to participate in such events, after the fun and enjoyment, intellectual challenges, and status and reputation (Juell-Skielse et al., 2014). Therefore, it is possible that the methods used to stimulate the use of OGD do not meet the information needs of the crowd and carve a gap between intermediaries' information solutions and real users' information needs.

To solve the above problem, the purpose of this paper is to develop design principles that can help data providers, intermediaries, and users design methods to capture information needs of users that can be satisfied by building information products based on OGD. A design principle is a statement that prescribes what and how to build an artefact in order to achieve a predefined design goal (Chandra, Seidel, & Gregor, 2015). Providers can use the information needs to identify sought-after data, intermediaries can use the information needs as an input to the development of information solutions to satisfy these needs, while users can follow the design principles to help them express information needs of their community. The design principles are a first step towards bridging the gap between a group of users, their information needs, and information solutions in a specific everyday context of information use. The information needs expressed by implementing the design principles could cover three functions of information use for the users: fact-finding (answering specific questions), staying aware (keeping up to date on a topic) and briefing (refreshing one's knowledge and understanding in more depth) (Erdelez, 2005; Nicholas & Herman, 2010). This research is guided by the following research question:

- What design principles could data providers and intermediaries follow to design methods to identify information needs of groups of users?

In this paper, the presentation of the research process and results follow the recommendations of Gregor & Hevner (2013). The paper is structured as follows. Section 2, the background, introduces the concept of information needs, approaches and methods to capture information needs, and the design principles as prescriptive knowledge for designing methods. It has contributed to the design and development of our design principles. Section 3 develops the research process, anchored in Design Science Research. Section 4 presents the final design principles, with examples taken from the last instantiations, and the evaluation. Then, Section 5 discusses the results and implications, which is followed by a conclusion.

## 6.2. Background

### Information needs and use

People gain information when they make sense of data, where data can be defined as a representation of objective facts or unprocessed information (Hey, 2004). Open datasets can offer many different reuses, but first aim to produce information (European Open Data Portal, 2019; Hey, 2004). When there is an extra processed layer added to information, it can become a digital information solution (e.g., dashboards and journalists' blogs (Davies, 2010)), an improved service (e.g., waste collection optimized by the real-time data of sensors placed in

the trash bins), or an aggregated information service (e.g., an optimized route planner based on different data sources and public transport data) (Berends et al., 2017). These solutions can be used to satisfy information needs, which T. D. Wilson (1981) considers to be a secondary need that can emerge when people try to satisfy primary needs. When people recognize a gap in their knowledge, information needs arise (Belkin & Vickery, 1985) that are personal and contextual. Information needs are determined by the roles an individual fills in social life, the environment within the role is performed, and depend on his level of knowledge on the matter (T. D. Wilson, 1981).

Information can help people in their work, with problem-solving, or pursue hobbies (Nicholas & Herman, 2010). Information is used to find facts, stay aware, research, obtain background understanding, stimulate thoughts, and recreation (Nicholas & Herman, 2010). Its value does not decrease as it is used. Rather, the value can increase as one piece of information is added to the other (N. Moore, 1998). The search for information follows these stages: (1) needing (the first inkling that some information might be needed), (2) starting (attempting to understand their need and preparing for a search), (3) working (to identify the information), (4) deciding (on the value of any results), and (5) closing (and wrapping up the work) (Westbrook, 1993). Any of these stages can be the final one, skipped, iterated between, and lead to a new effort to seek information. People can encounter information when they are searching for information related to another topic or bump into it as they are working (Erdelez, 2005).

Information needs are difficult to understand for the person experiencing them and the people who study them. They can be perceived as a hologram that the person walks around and through and may have difficulty putting it into words and seen as an anomaly (Westbrook, 1993). If the situation is unclear, confusing, or difficult to comprehend, the person may be unable to formulate questions that can help them identify information for their need (Daft & Lengel, 1986). This situation can be accompanied with a feeling of uncertainty and apprehension (Kuhlthau, 1991). The person might need some initial gathering of information to make a general, internal expression of their need (Westbrook, 1993). They can discuss their need with friends or relate it to previous experiences and knowledge (Kuhlthau, 1991). New pieces of information can cause the frame of reference to change (Westbrook, 1993), which can change their understanding of the need or the need itself. On the other hand, people might lack awareness of their need (dormant or unrecognized), know about it, but not work to satisfy it (unexpressed) (Erdelez, 2005; Nicholas & Herman, 2010), or misunderstand their expressed needs (misguided) (Shenton, 2007).

From an external perspective, the actions of a person attempting to satisfy their information need can appear chaotic. They can jump between past, current, and future information needs and different information-seeking processes, which can lead to cross-pollination. These behaviors can make the person seem to behave irrational, change focus, and not follow search directions provided by experts (Erdelez, 2005). At the same time, encountered information (1) can help a person satisfy another information need in the present, (2) could have helped them or other persons in the past, and (3) can help them satisfy a future information need (Erdelez, 2005). Consequently, other people, experiences, and situations are all sources of information and knowledge, which means information can be found in unexpected ways and places (Erdelez, 2005).

People could want, demand, or need information (Nicholas & Herman, 2010). Information wants are similar to a dream scenario and are often based on a bit of wishful thinking. In a perfect world, it would be the same as their information need. Lack of time, resources, motivation, and knowledge can mean people are tempted to obtain information that they in fact do not need. A want can lead to a demand for information, which is formulated as a request for an item of information based on what they think they want. However, there can be a mismatch between their understanding and the information, which causes them to demand information they do not need (Nicholas & Herman, 2010). The above qualities of information needs add to the difficulty of identifying and satisfying information needs of users for data providers and intermediaries.

### Related works: approaches and methods to capture user information needs

Previous research has explored how to capture needs and information needs of users, within different schools of thoughts. It is often framed as a form of users involvement that has become a common practice in, for example, Information System, Human Computer Interaction, Service Design, Innovation, New Product Development, and Marketing. A better understanding of the needs can increase the user's satisfaction and acceptance of the designed solutions, by improving the solutions all along the process (Kujala, 2003). The active involvement of users is a growing innovation strategy to develop internet-based applications since it can help intermediaries to understand user requirements, access to useful information about their context of uses, habits, preferences, get new ideas, and define the scope of a project (Komninos, Pallot, & Schaffers, 2013). Therefore, understanding the users' needs can serve multiple purposes and the used methods vary according to the objectives of the intermediaries. But, information needs are specific needs, requiring specific methods.

The methods that focus on the identification of information needs originate in Information System and Information Science literature. They have developed strategies to overcome the difficulties for the users to express information needs. One such method is to involve the users in collaborative workshops (Barbosa Tavares, Hepworth, & De Souza Costa, 2011; De Tuya, Cook, Sutherland, & Luna-Reyes, 2017; Hogan et al., 2017; Staron, Meding, & Baniasad, 2019). Barbosa Tavares et al. (2011) used techniques including scene setting, brainstorming, cards for people to express ideas, individual and group work, and discussions as they are efficient to help people identify their information needs. In fact, they combine different strategies to enable the participants to express their needs and constraint them to follow a structured mental framework to unravel the unexpressed or misguided ones. Given that information needs are secondary needs and are identified when people experience a knowledge gap (Belkin & Vickery, 1985), the users need to be guided in an information seek workflow (Hörold, Mayas, & Krömker, 2012; Timmerman, Boer, Hisschemöller, & Mulder, 2001). Scenarios (J. M. Carroll, 1999), or vignettes (short scenarios) (Urquhart, 2001), and the critical incident technique (Urquhart, 2001) can be used to create a situation where participants experience knowledge gaps. They can help the user to project himself in a clear context, role, and tasks. In term of group dynamics in workshops that involve several stakeholders, to enhance the active participation of everyone, Staron et al. (2019) used the brainwriting and similarly, Hogan et al. (2017) used the idea writing technique of Warfield (1994).

Information needs are complex to capture as they are temporary, personal, contextual, and sometimes misguided or unexpressed. To tackle such a complexity, Devadason & Lingam (1997) have developed a set of methods, which are extensive and time-consuming by nature.

They argue that no single method can identify information needs. They use their methods to study users (role, tasks, behaviours by direct involvement and indirect user studies, observation), their environment (available channels, tools, organisation by case study, desk research) and information use (which need, information form, frequency of need, by interviews, surveys). Methods to capture information needs can have different focuses (Westbrook, 1993): (1) understanding the information needs and the problem (knowledge gap and its reasons (e.g., Staron et al., 2019), (2) understanding the information self (what the user is looking for performing specific tasks, its form (e.g., Hess, Diebold, & Seyff (2017) and (3) designing the interface between the two (the solution, information system, its functionalities, that will deliver the information (e.g., Paulus, Meesters, & van de Walle, 2018).

Evidently, methods developed in innovation and service design for the capture of users' needs can be applied, to some extent, to capture information needs (e.g., Al-Shboul & A.Abrizah, 2014). Information needs are difficult to express in words (Shenton, 2007), just as the so-called latent needs of users (Sanders, 2002; Sanders & Stappers, 2008). As in service design, the fulfilment of information needs is embedded in the development of solutions. In service design, co-design and empathetic design, two approaches to participatory design, are particularly relevant to the capture of unexpressed or misguided information needs. They involve users to imagine or envision a future practice or product, and to seek inspiration together with the designers (Steen et al., 2007). In participatory design (Abras et al., 2004), users are partners with the intermediaries, their involvement is considered as central in the process. Therefore, a special effort is put to the development of tools and techniques to enable the expression of the participants, and the communication between the users, experts in their usage experience, and the intermediaries, experts in their own field and technologies (Sanders, 2002; Sanders & Stappers, 2008). The generative tools guide participants in small steps to construct and express deeper levels of knowledge about their experiences, in order to get to know tacit knowledge and latent needs, often difficult to express in words (Visser et al., 2005). The activities are intended to engage users in the group cognition, leading to the emergence of new ideas, concepts and solutions (Pallot et al., 2011; Sanders, 2002; Sanders & Stappers, 2008).

Previous OGD research that study methods to capture information needs are scarce, and have focused on how OGD can empower citizens (Barbosa Tavares et al., 2011), how to engage intermediaries to create value (De Tuya et al., 2017; Hogan et al., 2017), or how to develop an OGD solution that solves a specific issue (Paulus et al., 2018). The users' information needs are to elicit to reuse OGD, but the researchers' aim was not the development of methods or guidelines that tackle this challenge. In conclusion, several schools of thought have explored and developed methods to capture needs difficult to express in words (information needs and latent needs), but to our knowledge, no method has been developed to capture the information needs of users for OGD. This paper aims at developing design principles as a first step towards filling this gap.

### Design principles in design science research

Design Science Research (DSR) is an approach to research which output, the artefact, aims to contribute to the knowledge base (rigor and scientific legitimacy) and be useful in its environment (relevance) (Dresch, Lacerda, & Antunes, 2015). In other words, DSR develops artefacts that should achieve a goal and solve a problem, which are presented as the result of the research. The artefact is designed and developed based on problems, motivations, and objectives identified by researchers in practice and previous research. This artefact is then

demonstrated and evaluated in practical contexts to observe how well it supports a solution to the problem (Peffers et al., 2007). Artefacts can be of four types or classes: classified constructs, models, methods, or instantiations (March & Smith, 1995). Given the practical orientation of the artefact, the contribution to knowledge can be either more limited and specific (e.g., one instantiation), or more abstract (e.g., design theories) (Gregor & Hevner, 2013). Design principles belong to the latter and are considered as prescriptive knowledge. A design principle is a statement that prescribes what artefact and how to build it in order to achieve a predefined objective (Chandra et al., 2015). Its purpose is to capture knowledge about the creation of other instances of artefacts belonging to the same class (Sein, Henfridsson, Purao, Rossi, & Lindgren, 2011). It moves the knowledge contribution from the specific case to the generalization. Design principles should be evaluated based on their reusability and utility in their environment (Iivari, Rotvit Perlt Hansen, & Haj-Bolouri, 2020).

An effectively formulated design principle contains three kinds of information: (1) information about the actions made possible through the use of an artefact, (2) information about the material properties making that action possible, and (3) the boundary conditions under which the design will work (Chandra et al., 2015). Gregor, Kruse, & Seidel (2020) add three components that should make a design principle understandable and useful in real-world design contexts: the specification of the actors (implementer vs. user of the artefact), the aim of the user (its problem to be solved) and the rationale (why a mechanism is recommended).

In the end, the conceptual schema of a design principle by Gregor et al. (2020) is made of seven components: the aim, the implementer, the context, the users, the mechanisms, the enactor, and the rationale. *The aim* is what is to be achieved. *The implementer* is the actor that instantiates abstract specifications in a concrete design context. *The users* are those whose aims are to be achieved. *The mechanisms* are human actions, the use of artefact, or the combination of both, that lead or allow users to achieve an aim. The mechanisms may require the help of *enactors*, who perform actions as part of the mechanisms that are used to accomplish the aim. *The rationale* is a justification for believing that the mechanisms will lead to achieving the aim. Gregor et al. (2020) suggest the following formulation for design principles: *"For the implementer, to reach the aim of the user, in a context, use the mechanisms, because of the rationale"*.

Despite design principles having their roots in the Information System literature to develop socio-technical artefacts (e.g., tools, computer systems, interfaces), they have also been used to prescribe co-creation methods (e.g., Kohler, Fueller, Matzler, Stieger, & Füller, 2011).

## 6.3. Research process

The development of the design principles followed the design science research (DSR) methodology and the steps suggested by Peffers et al. (2007): (1) problem identification and motivation, (2) definition of the objectives for a solution, (3) design and development, (4) demonstration, (5) evaluation, and (6) communication. We iterated between the steps, and the result is a set of design principles.

The design principles have been communicated to practitioners in a popular science article (description of the design principles with examples of implementation based on the instantiations, addressed to practitioners without prior knowledge and accessible at http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-181126).

### Problem identification and motivation

DSR is oriented towards solving specific problems and requires a pre-evaluation of the problem to be solved (Dresch et al., 2015; Goldkuhl & Sjöström, 2018; Peffers et al., 2007).

The problem identification for this study is that the data providers and intermediaries' low awareness of the users' information needs, adding to the problems of which OGD to publish and what solutions to develop. Previous attempts in the area, such as hackathons, are rarely driven by the information needs of the users, more built on existing datasets, technology, and the experiences of the intermediaries (Carr & Lassiter, 2017; Johnson & Robinson, 2014; Juell-Skielse et al., 2014). In reality, this situation means that the user's information needs are most likely not met and that the use of OGD continues to be low. We understand that the three actors are different and separated by cultural values, possibly geography, and goals, but there is also a focus on data and technology-driven solutions over actual needs. There is a misalignment between the three parts: providers' data, intermediaries' solutions, and users' needs (Belkin & Vickery, 1985; Davies, 2010; T. D. Wilson, 1981).

At the same time, we know that information needs emerge when people try to satisfy primary needs and experience a knowledge gap (T. D. Wilson, 1981), but they can be dormant, unrecognized, or unexpressed (Erdelez, 2005; Nicholas & Herman, 2010). They are personal, contextual (T. D. Wilson, 1981), and can be difficult to put into words (Westbrook, 1993). Situations that are unclear, confusing, or difficult to comprehend, can make it impossible for the users to express their information needs (Westbrook, 1993). Moreover, wants, demands, and needs for information are not the same and only the last is valuable (Nicholas & Herman, 2010).

### Definition of objectives

The objectives of the artefact, a set of design principles, are: (1) to enable data providers, intermediaries, and users to develop and implement an efficient method that allows a defined group of users to express their information needs and (2) to enable them to reuse the output of the method to inform data providers and inspire intermediaries. The output should provide insights to the intermediaries to develop desirable information solutions for users and support data providers to identify datasets in relation to information needs. The design principles should be actionable methodological guidelines that help the actors to realign the OGD release and reuse on the information needs of users, in a given context, for relevant issues and user groups. They should contribute to the knowledge base by addressing the lack of impact of OGD with an original approach: the integration of users information needs in methodological design principles.

The next steps, design and development (3), demonstration (4), and evaluation (5) were developed in three successive iteration cycles, as summarised in Figure 22.

| Iteration cycles | | | |
|---|---|---|---|
| | **3. Design** | **4. Demonstration** | **5. Evaluation** |
| **Iteration 1** | **Face-to-face** Workshop<br>1:30 hour<br><br>**Purpose:** test of vignettes and collective approach | **Instantiation 1**<br>**One workshop**<br>**Participants role:** Users<br>**Profile**: Researchers<br>**Number**: 9 | **Instantiation level**<br>Method: debriefing with the participants<br><br>**Artifact level (V.1.)**<br>Method: observation and field notes |
| **Iteration 2** | **Face-to-face** Workshop<br>1:30 hour<br>**Purpose:** improvement of the intellectual process (question scheme); Definition of categories of information needs | **Instantiation 2**<br>**One workshop**<br>**Participants role:** Users<br>**Profile:** Students<br>**Number:** 11 | **Instantiation level**<br>Method: debriefing with the participants<br><br>**Artifact level (V.2.)**<br>Method: interview with 4 data providers and 3 intermediaries |
| **Iteration 3** | **Virtual** workshop<br>2:00 hours<br>**Purpose:** test of scalability (new participant roles, settings, and activities based on the Design Principles), and preparatory activities | **Instantiation 3**<br>**Three workshops**<br>**Participants role:** Data providers<br>**Profiles:** Civil servants<br>**Number:** 8 (a) - 15 (b) - 7 (c) | **Instantiation level**<br>Method: questionnaire to the participants<br>**Artifact level (V.3.)**<br>Method: interview or questionnaire with 3 data providers |

*Figure 22. The Three Iteration Cycles*

An instantiation is an implementation of the mechanisms, or methodological principles, developed based on our current knowledge. An artefact is the result: the abstraction of the instantiation (design principles addressed to a specific audience). Each iteration cycle was based on lessons learned from previous research, evaluation of the previous instantiations and artefact, and the problem domains of OGD and information systems.

### Design and development

The first iteration initiated with previous research. Collaborative workshops are more efficient than interviews to enable people to explore and identify their information needs as members of group or community (e.g., Barbosa Tavares et al., 2011). Therefore, we started by developing and implementing a preliminary version of the design principles (artefact V.1.) into a face-to-face workshop with users. The workshop was based on the idea of critical incident technique where participants, as citizens, were asked to re-examine a brief, but memorable information seeking episode in a given situation, called vignettes (e.g., socialize or find food in a new city)(Urquhart, 2001). The participants had to describe their possible actions in those circumstances, identify related information needs, existing solutions and their issues, following the intellectual process of Hörold et al. (2012). In the iteration 2, we refined the question scheme and wording for a better understanding of the participants and decided to define categories of settings or situations of information seeking behaviours to avoid dispersion of the participants that resulted in a limited reusability of the insights for a specific intermediary. In the last iteration (3), we improved the relevance and scalability of the design principles that

were challenged in the evaluation of the artefact (V.2.). We instantiated the design principles in three digital workshops, with participants in a new role, future data providers, and adapted the activities as a new interpretation of the design principles. The participants were also invited to execute short sensitizing activities designed to raise their awareness of the topic and concepts (Visser et al., 2005).

### Demonstrations

We ran two demonstrations in face-to-face workshops with users, and three demonstrations in online workshops with data providers. For the first face-to-face workshops in iteration 1, we invited nine Belgian researchers as users and citizens of the city they work in (see Appendix II. A.), and for the second, in iteration 2, eleven Belgian students as they belonged to a homogeneous group of users and citizens and perceived to share similar needs regarding information related to the city they study in (see Appendix II. C.). All participants were volunteers and responded to an open call for participation. Both workshops lasted 1:30 hours.

We ran three demonstrations with future data providers in iteration 3, as initial session of a training program about OGD. The three digital workshops lasted 2:00 hours each, and involved respectively eight, fifteen, and seven civil servants of Belgian municipalities, association of municipalities, or regional administration (see Appendix II. F.). The online format was a constraint by the governmental measures to control the COVID pandemic, but considered as an opportunity to test the scalability of the design principles in different settings.

### Evaluation

In DSR, evaluation can be performed in many different ways, as it depends on the type of artefact and problem (Hevner et al., 2004; Peffers et al., 2007). Nevertheless, this step should ensure that the artefact is relevant and useful (Gregor & Hevner, 2013). We evaluated on two levels: the instantiation (implementation of the design principles as workshops) and the artefact (set of design principles, as abstraction of the instantiations). On the instantiation level, we evaluated whether the mechanisms reached the predefined objectives (Chandra et al., 2015; Peffers et al., 2007). For that purpose, we evaluated the smooth execution of the principles with observation (Hevner et al., 2004), analysed the quality of the output produced by the participants, collected participants' feedback (Peffers et al., 2007) (see Appendices II. B., II. D, and II. G.), and finally, provided informed arguments based on previous knowledge (Hevner et al., 2004). On the artefact level, we evaluated the relevance and usefulness of the design principles for the audience. We reported the design principles as a popular science article and submitted it to potential users of the artefact. The first version of the design principles, a method under development, was evaluated in three semi-structured interviews with four representatives of future providers (two small municipalities) and three members of a team of experienced intermediaries (a Belgian digital start-up that developed an OGD based application for the group of users whose information needs were investigated in iteration 1 and 2) (see Appendix II. E.). As the design principles were reaching their final form in the artefact V.3., as presented in this paper and the current popular science article, they were evaluated based on the design principles evaluation framework of Iivari et al. (2020). The popular science article, together with a questionnaire (online form) was sent to six experienced data providers, as a way to complement the feedback got from less experienced data providers (respondents of the first evaluation and participants of the iteration 3). They were offered to either fill the form at their own pace or suggest us a time at their convenience for an interview,

based on the given questionnaire. Two respondents accepted the interview (the OGD managers of a large municipality in Belgium and a public traffic agency in Sweden), and one filled the questionnaire (a regional coordinator of OGD in Sweden) (see Appendix II. H.).

## 6.4. Results

This section presents the final design principles, a demonstration of the third iteration using these principles, and, finally, the evaluation of the principles.

### The design principles

We followed Gregor et al. (2020) for the structure of the final set of design principles (DP), which all share the same aim, actors and context, as presented in the first box. The following boxes present nine mechanisms and their rationales.

**Set of Design Principles: Aim, Actors, and Context**
For the data providers, intermediaries, or users (implementers) to capture the users' information needs (user and aim) in an everyday, information-rich context, we recommend:

**Design Principles 1: General principle**
**Mechanism:** To employ a mix of data collection methods, activities, and tools to better grasp the complexity of information needs, involving users, but with little impact on them.
**Rationale:** Information needs are personal, context-related, and role-based: the involvement of the users is essential to visualize these needs. But, they are also complex and time-consuming to identify (Devadason & Lingam, 1997; Nicholas & Herman, 2010). We sought a balance between complexity and time invested (efficiency) in the implementation of the method and the interactions with all the stakeholders (especially, the users).

**Design Principles 2: Main data collection method (1/2)**
**Mechanism:** To employ a collective intelligence-based workshop as a key data collection method, using creative tools and activities (i.e., tools that encourage the ideation with the use of text, pictures, to present ideas) and idea-writing techniques (i.e., written structured brainstorming, that uses variation of work dynamics, personal and in small groups, with a final presentation of the ideas in a plenary session), with the guidance of a facilitator.
**Rationale:** Information needs are difficult to express spontaneously and in words (e.g., people have dormant information needs, unexpressed needs, misguided needs (Shenton, 2007)). Collective workshop has been proven to be an efficient technique to identify people's information needs (Barbosa Tavares et al., 2011; De Tuya et al., 2017; Hogan et al., 2017; Staron et al., 2019). Creative tools, e.g., generative tools, can help participants to express latent needs, thoughts and feelings (Sanders, 2002; Sanders & Stappers, 2008), and idea writing techniques ensure input from a diverse range of people and temperaments (Staron et al., 2019; Warfield, 1994).

Design Principles 3: Main data collection method (2/2)
Mechanism: To use preparatory material and activities for the participants that aim at developing their understanding of the information needs concept and issues of the context.
Rationale: An appropriate preparation can raise the awareness and understanding of the participants about the abstract concepts and questions used in the workshop, stimulate their memory about the plots given in the vignettes (critical incident)(Urquhart, 2001), improve the quality of their production during the session (Visser et al., 2005), and, finally, increase the social interactions, their enjoyment and sense of efficacy, as they are known as important motivational factors in innovation communities (Antikainen, Mäkipää, & Ahonen, 2010; Roberts, Hughes, & Kertbo, 2014).

Design Principles 4: Adaptation to the aim actors and context (1/3)
Mechanism: To employ short scenarios (vignettes) to focus the participants on clearly identified roles and contexts, and the plots (critical incident technique) to trigger a knowledge gap (a brief, but memorable information seeking episode, based on their own experience but prompted by the problematic situation). The participants-short scenario fit implies a thoughtful recruitment.
Rationale: Information needs are context-related, personal, and appear when people realise a lack of knowledge (Al-Shboul & A.Abrizah, 2014; Belkin & Vickery, 1985; Nicholas & Herman, 2010), which can be prompted by the critical incident technique (Urquhart, 2001).

Design Principles 5: Adaptation to the aim, actors, and context (2/3)
Mechanism: To suggest predefined categories of information needs, in regard with the roles and context clearly circumscribed.
Rationale: Information needs can be broad and poorly pre-selected information needs can lead to irrelevant information solutions (Warner, Murray, & Palmour, 1973).

Design Principles 6: Adaptation to the aim, actors, and context (3/3)
Mechanism: To suit the vignettes with a relevant issue for the three actors (data providers, intermediaries, users), and link them to the OGD already or likely available.
Rationale: The interest of the data provider is needed to release OGD. The motivation and knowledge of the intermediary is needed to invest time, skills, and resources in developing solutions. A fit with the need of the users is needed to increase the relevance, acceptance and use of the solution. This also helps to identify actual needs.

Design Principles 7: Intellectual process
Mechanism: To use a sequential question scheme that guides the participants in the intellectual process of information needs exploration. It starts with the identification of primary needs in a given context and role, and the actions or tasks required to fulfil that need. It continues with identification of the information needs behind the tasks (by listing the needed and missed information and the weaknesses of existing solutions). It wraps up with ideas for new solutions.
Rationale: Information needs are secondary needs and can be identified by tracking first the tasks, then information, and finally categories of data (Hörold et al., 2012). People often need time to reflect and process information needs to express them properly. The use of activity, existing solutions, and future solutions help to give a frame to the need and make it more expressible.

Design Principles 8: Purpose of the result (1/2)
**Mechanism:** To ensure that the method results in the identification of information needs of groups of users.
**Rationale:** The development of information solutions should satisfy the needs for more than one person to be sustainable and economically viable (Nicholas & Herman, 2010; Westbrook, 1993; T. D. Wilson, 1994). It is important to study the result for similarities and quirks, but also for needs, wants, and demands (Nicholas & Herman, 2010). Results that are fragmented often indicate that the participants lacked experience with the information needed, as such it is important to ensure that the participants have a good understanding of the target group and discuss with them, or represent them.

Design Principles 9: Purpose of the result (2/2)
**Mechanism:** To ensure that the material produced by the method can be used by data providers for dataset prioritization and by intermediaries for inspiration to OGD innovation, considering the alternative information solutions that exist.
**Rationale:** Data providers tend to follow a supply logic (Susha et al., 2015), intermediaries tend to develop solutions with little impact (Carr & Lassiter, 2017; Hjalmarsson, Johannesson, Jüll-Skielse, & Rudmark, 2014), and the information needs are embedded in an environment (stakeholders, channels, habits in information use, other partial solutions available) that is important to evaluate when designing new information solutions (Devadason & Lingam, 1997; Hess et al., 2017; Staron et al., 2019; Timmerman et al., 2001).

### Demonstration of the design principles

For clarity purposes, Table 14 shows the way we combined the design principles in a mixed method, conducted with data providers in webinars (last instantiations). We summarise the objectives, activities, deliverables, and related design principles (DP + n°). The first DP was applied by developing a mixed-method divided into three phases: (1) the preparation, (2) the collaborative workshop, and (3) complementary analysis and relevance check to complement the insights. The following DP were spread across the parts.

Both face-to-face and digital workshops followed the same structure and objectives, involved the users at a moment, and resulted in the identification of information needs. However, the practical activities and tools were scaled to the participants. Given the focus of this paper, on the design principles as abstractions of the instantiations, the readers who want more extensive examples of tools and presentations of the instantiations can find them in the popular science article (accessible at http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-181126).

*Table 14. Overview of a mixed-method (DP 1): a virtual workshop with data providers*

| Phase | Objectives | Activities | Deliverables |
|---|---|---|---|
| *Phase 1. Preparation* | 1. Define a context and issue (DP 5) based on previous knowledge (experience, strategic priorities, literature)<br>2. Define the profile, roles (DP 8), and categories of situations creating information needs (DP 5) of the users based on previous knowledge<br>3. Preparing the participants for the workshop (awareness of the topic and concepts) (DP 3, 6) | 1–2. Kit of preparatory activities with clear instructions for the participants to define a problem (trigger questions), the context (mapping exercise), and user group (persona)<br>3. Previous to the workshop, a short presentation of the concepts (information needs, sources, solutions) given by the implementer | 1. A precise context, problem and group of users with categories of situations creating information needs<br>2. A kit of preparatory activities for the participants |
| *Phase 2. Collaborative workshop (DP 2)* | 1. Enable the participants to identify progressively user' information needs, existing solutions and weaknesses, and possible imagined solutions (DP 7) in a given role and context (vignettes) (DP 4) | 1. Problem situation brought by the participants converted into vignettes, used as a start for a written brainstorming. Use of the participants' field experience to limit categories of situations of information needs given the context. Brainstorming tool: question scheme written on a canvas, following the track: primary needs in the situation, related user tasks, related information needs, existing solutions and their issues. | 1. New insights:<br>– Information needs (spontaneous and latent)<br>– Overview of the competing solutions<br>– Imagined solutions |
| *Phase 3. Analysis + Relevance check* | 1. Confirm the ideas and intuitions (imagined solutions) with users and knowledgeable experts (DP 8)<br>2. Connect the collected information needs with released or releasable OGD (DP 9)<br>3. Inspire possible developments and OGD intermediaries (DP 9) | 1. Examples of activities and tasks given to the participants to compare the brainstormed information needs with their user group and colleagues (question scheme in a workshop, interview, or focus group format)<br>2. Comparison of the categories of data with data owned in their administration<br>3. Presentation of the projects to developers (intermediaries) in a post-event | 1. A project<br>2. The identification of information needs and related OGD<br>3. The identification of future development avenues |

### Evaluation of the design principles

The evaluation focused on whether the objectives were achieved by the design principles, and the perceived relevance and usefulness of the artefact by the audience. The main insights of both are presented subsequently and are based on empirical material and previous literature (informed argument, Hevner et al. (2004))

#### Achievement of the objectives: value of the design principles.

**Objective (1)** was *to enable data providers, intermediaries, and users to develop and implement an efficient method that allows a defined group of users to express their information needs.* To achieve that objective and overcome the intellectual challenge of information needs identification, the iterations and evaluations revealed the importance of (1) the group, (2) the preparation of the participants, and (3) the guidance in the intellectual process.

**The group** has proven its worth for two reasons. First, in the search for efficiency. An information seeking behaviour is only one manifestation of a complex situation, influenced by individual's roles, tasks, the environment, existing solutions. Exhaustive but time-consuming design principles (Devadason & Lingam, 1997) are likely to be rich in results but irrelevant to their audience. Hence, as a trade-off, we developed the three-phase method **(DP 1)**, with a collaborative workshop as the main method **(DP 2)**, since Barbosa Tavares et al. (2011); De Tuya et al. (2017); Hogan et al. (2017); Staron et al. (2019) have proven that collective intelligence improves the individual understanding of one's information needs. The group support helped the participants to understand the concepts, questions, and generate more ideas. Second, the group contributed to the participants' good experience. The data providers found a community of practice where to share field experience and ideas. The users had a good time. Roberts et al. (2014) found out that in value co-creation with community of users, the value of participation, for the participants, is in the process, not the output, and in the social interactions of like-minded people. The fun, enjoyment, and reward for participation are important motivational factors in innovation communities (Antikainen et al., 2010). This type of motivation was noticeable in the user workshop. In comparison with the face-to-face workshops, the online workshops brought difficulties in the work dynamics which can impact the efficiency and participants' experience. More time was required to break the ice and engage the participants, simple participation rules were not always followed, and the first technical difficulties (access to the shared documents, broken links, dysfunctional microphone) can break the participants' motivation and engagement down very quickly and for long, due to the enhanced difficulty to use the tools and the distance between the participants and implementer.

A need for **a good preparation (DP 3)** was also perceived from the first iteration of the users workshops. The participants tended to mix up the question (information need) with the answer (sources and solutions). For the users, we developed a sensitization kit inspired by Visser et al. (2005), and data providers were involved in narrowing down the problem, users profiles, roles **(DP 5 and 8)**, and mapping their environment (Devadason & Lingam, 1997). As argued by Mulder & Stappers (2009), both time for reflection and a tool of expression are needed to help the participants become *"experts of their experience"*, which increases the breadth and depth of the results of participatory activities. The activities are designed to be an invitation to reflection, without conditioning the participants, and raise their awareness about habits and environment. The preparatory activities were tested on the data providers in the instantiation

3. Yet, they were first received by the participants as an extra and unnecessary workload, until they felt confused at the workshop with concepts they are not used to think about, and acknowledged their value in the post evaluation. In theory, the sensitization kits and similar are promising (Visser et al., 2005). In practice, the implementer has limited control over the engagement of the participants in off-site activities. To move a necessary preparation out of the facilitated workshop, to reduce the impact of the method on the participants' agenda, still comes with the risk of reducing the participants' awareness, their feeling of self-efficacy, and the quality of the insights produced. Self-efficacy is an important motivational factor for participants to join a workshop: they have to feel that they can contribute (Roberts et al., 2014). But, they might not foresee the value of such activities and a lack of awareness can influence their motivation and production (Roberts et al., 2014).

The guidance in the intellectual process (DP 7) (Hörold et al., 2012) was ensured, first, by the presence of a skilled facilitator (method implementer), highly important to keep the focus of the participants and challenge their preconceptions, otherwise they can lose the track of their role and produce irrelevant output. Second, it was enabled by the use of the critical incident technique (Urquhart, 2001) that helped the participants to remember life episodes, and the wording of the questions was good enough to make the tools (brainstorming cards or canvas) self-sufficient. It increased substantially the quantity and reusability of the insights produced by the participants. We gained insights into how participants can be helped to express the hologram of information need and demand for an information solution (Nicholas & Herman, 2010; Westbrook, 1993).

Objective (2) was *to enable the data providers, intermediaries, and users to reuse the output of the method to inform data providers and inspire intermediaries. The output should provide insights to the intermediaries to develop desirable information solutions for users and support data providers to identify datasets in relation to information needs.* The guidance in the intellectual process allowed us to perceive (1) degrees of importance in the information needs expressed but also (2) a difficulty for the provider to come closer to the users.

We noticed degrees of importance. The participants start with, what we call, their top-of-mind information needs, for a role, in a context, given the existing solutions ("What information do I need for X task, that I already get from Y solution"). Then through the identification of issues with existing solutions, they can express a deeper level of needs, the latent and not spontaneously expressed needs (frustration and missing features help them to find out the must-have, what information they value most). The creative part, their imagined solutions, allowed them to express their wishes, the "nice-to-have" (what they would value and is not yet satisfied with present solutions, the difference between information needs and wants (Nicholas & Herman, 2010)). It can help the intermediary to identify potential paths of developments.

However, we also noticed the difficulty of getting the data providers and users closer, which is later reaffirmed by the evaluation of the design principles. Complementary methods were implemented by the researchers to check the relevance of the output of the user workshops (DP 8, 9). The same was suggested to the future data providers, participants of the last instantiation and users of their insights, as they were working on personal projects. They were suggested to meet users, use the participatory mechanisms and communication channels of their municipality to challenge their ideas with the target user groups. None of them did it. We identified several reasons for that reluctance: lack of time of the participants, lack of perceived

relevance and unfamiliar with user engagement in the development of solutions, fear that citizen participation commits the municipality to deliver the solution, other projects and topics of citizen participation in the strategic plan, rigidity of the administration and its processes, exceptional circumstances (COVID and confinement). In conclusion, for the implementation of the DP 8 and 9, the participants needed more arguments and guidance (ideas of methods and tools relevant to their experience and context) to do it on their own, which opens for future research.

Perceived usefulness and relevance by the audience: data providers and intermediaries.

Following the evaluation framework of Iivari et al. (2020), the design principles were perceived by the experienced providers (OGD managers) as accessible (understandable with the provided examples and tools), important as they address a real problem in their practice (lack of reuse and understanding of users' needs), novel in their approach and useful. However, they perceived their actionability differently. The first respondent challenged it, as it is usually difficult to get in contact with the right kind of users and stakeholders. The lack of budget, resources, and time to organise the implementation of the design principles and process their insights places the exploration of user's information needs as a secondary task, which maintains the supply logic that current and previous research try to tackle (Susha et al., 2015). The second respondent felt ready to try the design principles and was only concerned about the difficulty to recruit the right participants, the strategies to adopt to motivate people for such activities (this corroborates the DP 4), and the adaptability of the questions to their level of expertise (field knowledge, technical, or user expertise). The last respondent solved it by thinking about trying them out with internal intermediaries: colleagues of his administration or other public services that are part of the data-sharing ecosystem. A group of professional intermediaries share a clear context, role and tasks, and is known and accessible by the data provider, which eases the preparation phase. This hesitation to implement the design principles was not noticeable when we presented the design principles in an early development stage to the start-up, as they were used to UX methods and user engagement in the development of their digital services.

The users' insights were, according to the intermediaries (the experienced start-up), bringing up interesting perspectives. Questioning a problem by the information needs prevents the intermediary to fall into confirmation bias by validating their own ideas at an early stage of development instead of exploring user's problems. Therefore, the implementation of the design principles revealed unexpected users' information needs they never thought about before, which can generate new ideas. However, the risk is that the users stay too close to what they know and their individual needs (user-centred methods are known to generate incremental improvements (Verganti, 2008)), wish unrealistic solutions, or go beyond the core business of the intermediary. During the ideation of solutions in the workshops, both future data providers and users were sometimes going in every directions, coming up with solutions that would solve everything at once. They lacked realism in data availability (users wanted solutions based on privately owned and generated data, and data providers dreamt about Waze-like real-time traffic data produced and shared by the citizens). The data providers in the digital workshops were lacking knowledge regarding the data availability in their organisation, the technical possibilities, and skipped the involvement of the users, focusing on their organisational priorities and projections. Therefore, for the data providers, participants of the digital workshop, the insights were useful to frame a project to address to intermediaries and to challenge their understanding of OGD. The impact of these insights on the data release

remains limited, likely due to the context of the participants (civil servants with limited power of initiative).

## 6.5. Discussion

The design principles presented in this paper originally combine previous research and approaches to capture information needs. The limited research about information needs within OGD research focuses on identifying needs in a specific context by researchers (e.g., Barbosa Tavares et al., 2011; De Tuya et al., 2017; Hogan et al., 2017; Paulus et al., 2018). This approach limits the generalizability, as information needs are determined by social roles, environments, previous knowledge, and experiences (T. D. Wilson, 1981). Our research overcomes this limitation by developing design principles that can be used by other researchers and practitioners to capture information needs. Similar to previous research, our design principles are based on collaborative workshops (e.g., Barbosa Tavares et al., 2011; Hogan et al., 2017; Paulus et al., 2018; Staron et al., 2019). This approach fuelled the creativity of the participants (Pallot et al., 2011; Sanders, 2002; Sanders & Stappers, 2008) and helped to overcome problems relating to expressing information needs (Westbrook, 1993) since the participants could discuss their need with peers (Kuhlthau, 1991). On the other hand, the design principles originally combine approaches to capture information needs with approaches from service design. In the information needs approaches, Barbosa Tavares et al. (2011) explain that it is important to unfold the information needs of participants following a structured mental framework (Hörold et al., 2012; e.g., Timmerman et al., 2001). Urquhart (2001) explains that critical incident technique helps participants remember information needs previously encountered. Westbrook (1993) explains that the approaches can focus on understanding (1) the need and the problem, (2) the information, and (3) the solution. In service design, Steen et al. (2007) explain that participants can imagine or envision a future product, which requires tools to help them express themselves (Sanders, 2002; Sanders & Stappers, 2008). The tools guide them in small steps to construct and express their experiences, tacit knowledge, and latent needs that are often difficult to express in words (Visser et al., 2005). Consequently, our design principles cover the three categories of Westbrook (1993) and their future possibilities. The combination of collaborative workshop, structured unfolding, and creative tools allows participants to identify present information needs, but also future and possible avenues of innovation. This extension is original for OGD research, as it takes a step away from the supply logic of OGD (Susha et al., 2015) towards an enrichment logic where OGD can supplement existing information solutions or find routes to surpass them. At the same time, it moves the responsibility of identifying information needs from researchers to practitioners where it can make the most value and impact.

### Limitations and future work

The conceptual schema of Gregor et al. (2020) to write understandable and useful design principles helped us to frame our thoughts. Nevertheless, for the practitioners, the design principles became understandable with examples and illustrations. Stand-alone design principles might not be sufficient to be practically applicable, despite the coverage of the suggested conceptual schema.

The set of design principles is not either a recipe for success, but should be seen as a method that can sparks new ideas and practices in the intermediaries and data providers' jobs. It should be continued with methods that value the involvement of the three actors: users, intermediaries, and data providers. Each actor has a limited perception of each other's needs,

objectives, and technical possibilities, that have to be mediated. Verganti (2008) suggests the role of the interpreter, who can help to connect user needs with the priorities and capabilities of the firms in the technological innovation process. The intermediary could be an interpreter to close the gap between a data provider's fuzzy objectives and the user's needs, just as the data provider can be interpreter between the intermediary's market oriented solutions and the public needs, to move towards an OGD reuse that benefits to a broader group of users.

The objective of producing a maximum of insights with a low impact on the users implies *limitations*. The design principles could be developed to capture the information needs with the same scientific rigor as Devadason & Lingam (1997). However, we propose a dense set of methodological recommendations to approach information needs in the context of OGD and to get insightful results at short term. Yet, it requires considerable time and reflection for the implementer to recruit people, design, and run the required activities, and process the output, which could make the implementer to be selective or expeditious and lower the quality of the insights. *Future work* could enrich and develop design principles for the preparation phase, especially regarding the questions of user's profile and recruitment based on their knowledge and role in context, and user involvement from a data provider's perspective, to maximize the collective intelligence and the scalability of the insights. Finally, the design principles could be implemented in different contexts (e.g., administrations, companies) to develop further new principles regarding the integration of the information needs in the data provider's and intermediary's work and processes.

### Implications

For practice, starting with the users' information needs can open a broader range of solutions and potential paths of OGD reuses. For the data providers, it can instil a better understanding of the OGD potential. The insights can also inspire the intermediaries with new reuses of diverse formats, based on users' problems and tasks. The investigation of information needs can enable the user's involvement in the early stage of solutions development. A collaborative workshop enriches the method. However, a digital setting can impede the realisation of collective intelligence benefits.

For research, the conceptual schema of Gregor et al. (2020) and evaluation framework of Iivari et al. (2020) offered a relevant structure to write and evaluate design principles. Moreover, more research of OGD from an information needs perspective is needed. It could allow the users to have a more active and influential role in data release and reuse. It highlights its importance, calls the researchers to go beyond the exclusive data provider – intermediary interaction abundantly discussed in the literature, and opens new research avenues to improve the realisation of the OGD benefits.

## 6.6. Conclusion

OGD is a resource allowing many opportunities of reuses. Most of the methods known in practice and research to stimulate reuses and develop solutions are technology-driven and attract skilled intermediaries in open innovation contests. The concept of users' information needs in OGD is scarce and previous methods to capture information needs, developed in that context, are covering other objectives (e.g., the citizen empowerment, intermediaries engagement, solution development). To reduce the gap between current open data solutions and user needs, in an inclusive and novel approach, we developed a set of design principles to enable people to develop tailor-sized methods to capture users' information needs. By taking

users information needs as a starting point, we enabled people without specific skills to express their needs with the support of a mixed method.

The contribution of this research is an artefact developed with DSR that can be used by intermediaries and data providers, to take a user-centred approach in their work. The design principles provide prescriptive knowledge to develop mixed methods and inform the implementer about the main data collection method, how to adapt to its purpose and context, the intellectual process to follow, and how to suit purpose of the results with the actor's needs and objectives. Despite it is possible to implement the design principles in face-to-face and digital settings, the first is easier for the implementer and was experienced as more enjoyable by the participants. The group, the guidance in the intellectual process and a good preparation of the participants enabled them to produce fine-grained insights, with perceptible degrees of importance, in a limited time. The imagined solutions, final results of a collective identification of information needs, are contextualized with a role, tasks, and information needs. They can enrich the intermediaries and data providers' perception of the user needs and inform them about what information and services users value most. However, we learned from the evaluation that off-site activities, such as sensitization kits and self-organised user engagement activities come with the risk to not be executed by the participants as their value is perceived afterwards. The design principles challenge especially how the data providers define the scope of their role and relation with the users, which opens interesting research avenues.

## Data Accessibility Statement

In this paper, the data are the developed design principles and tools made accessible in the popular science article (at http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-181126). The data to produce them (e.g., evaluations and participants' written notes specific to each iteration) are less relevant for the readers and, therefore, unpublished.

## Ethics and Consent

All the participants were informed of the research purpose and the use of collected empirical material. They consented to participate in the research activities.

## Acknowledgements

This research is part of the Wal-e-cities project funded by the European Union (FEDER H2020) and Walloon Region (Belgium) and part of a Belgian-Swedish research collaboration surrounding the utility of OGD. We would like to thank also the anonymous reviewers for their relevant feedback received at the 17th Scandinavian Workshop on E-Government (SWEG 2020) and at the EGOV-CeDEM-ePart 2020 conference (submitted as ongoing research).

## Competing interests

The authors declare that they have no competing interests.

All the participants and respondents were introduced to the purpose of the research and agreed on participating in the activities, interviews, or questionnaires without financial compensation. Food and beverages were offered to the participants of the face-to-face workshop. The demographic data (only sensitive data) are anonymized.

# Chapter 7. Discussion

## 7.1. Discussion of the findings: importance of the contributions and critical outlook

In this chapter, I discuss the findings and the contributions regarding the thesis's research problem and purpose. I pinpoint where we can see misalignments in the empirical data, and I synthesise what the roles helped us to learn. Then, I show the relevance of studying roles and value chains in open data by drawing implications for research and practice based on the contributions. The last section reflects on the research methods and contributions.

### 7.1.1. Misalignments and lessons learned through the study of roles

The studies did not aim at proving the occurrences of misalignments in a systematic way; hence, I cannot say that they are a recurrent feature of the open data initiatives in cities. Nevertheless, I can see in my empirical data that they do happen. It is not surprising given the large body of previous literature that exists about barriers and impediments (e.g., Janssen et al., 2012; Zuiderwijk et al., 2012; Ma & Lam, 2019; Martin, 2014; Smith & Sandberg, 2018; Toots, McBride, Kalvet, & Krimmer, 2017). Based on my findings, I also argue that the concept of role can bring insights into new perspectives to reduce or prevent misalignments in open data value chains. I point out the lessons learned, per misalignment, in the following paragraphs.

**(1) Misalignment between the actor's perception of other roles (in terms of definition, scope, and purpose)**

This misalignment happens when the actors engaged with open data perceive differently what a role is supposed to do. I focus on the roles of the municipalities as perceived by themselves (Study 2), and the role of the citizens as perceived by the municipalities (Study 2) and the initiatives' leaders (Study 1).

**Regarding the municipalities' roles**, we have identified seven alternative ways for the municipalities to take or make their open data role (Study 2). The findings show that the municipalities, *a priori* a homogeneous group of actors, do not perceive their role in open data in a homogenous way. They do not agree on what the role they make or take in open data encompasses. Through the typology, we elucidate part of the confusion around the scope and tasks of the municipalities' roles in open data (Mesquita et al., 2020). We show that, according to their views on their role, the municipalities can stay detached from open data and take a role sent by others, make new roles for open data inside their organisation, or assimilate open data to their primary role and public mission.

It means that **the chosen approach to their open data role is specific to a context and can be a new source of misalignment, as it affects the way they engage (or not) with the rest of the actors in the value chain (re-user, end-user)**. Ruijer & Meijer (2020) even argue that to maintain interaction with the rest of the actors engaged and give a chance to open data to last and scale-up, the public actors should integrate open data in their work processes, binding it to a permanent position within the organisation. In their research project, they experienced that, once the open data enthusiast inside their partner organisation changed job, the open data initiative stopped.

**Regarding the citizen's roles** as perceived by the initiatives' leaders, the findings of Study 1 show that the concept of "citizen" in participatory initiatives covers several practices. They vary according to the initiative leaders' rationale for open data and participation (categorised into

three approaches). We observed the same in Study 2, where municipalities can conceive the citizens as anyone willing to access open data ("Rare Birds"), or user of digital solutions, enhanced services, and informed decision-makers ("Smart Citizens"). I understand that the citizens are considered as a whole, but with different attributes according to the actors using the concept. I already noticed in my initial literature review (Background, Section 2.2.3.) this difficulty in role conceptualisation. What was new to me, is that the conceptualisation of role was not exclusively relatable to the actors' sector of activity (i.e., public, private, and third sectors): we found initiatives' leaders from the public sector (municipalities) taking a service innovation approach and involving citizens to provide data (Studies 1 and 2). Similarly, we found private entrepreneurs taking a civic innovation approach and expecting the citizens to be more active in their communities and city, or to change their behaviours for the common good (Study 1).

It means that **the actors' purpose in a specific initiative (project) defines the chosen approach to open data and the roles expected from the citizens**, not the actors' primary role and sector.

In sum, we have learned that **the actors** (initiatives' leaders in Study 1, municipalities in Study 2) **perceive others' roles through the lens of their goal for reusing open data, which results in role combinations and ways to engage with other actors unique to their initiative.** Moreover, both studies show that **the misalignments are context-related and are difficult to generalise between groups of actors**.

**(2) Misalignment between the municipalities' expected and realised outcomes**

This misalignment happens when the municipalities' expected outcomes differ from the realised outcomes, as we can see through the claimed objectives and displayed outputs (cases of reuse).

We saw in the empirical data of Study 2 that the municipalities set the **purposes of governance and transparency at the same level as innovation**. This is reflected in research (Charalabidis et al., 2018a) and in the EU open data policies (Lassinantti, 2019), where these purposes are repetitively cited together despite their conceptual distance. The survey of the municipalities' activities and cases of reuse they report on their websites (Study 2) and the citizen participation forms (Study 1) show, however, evidence of the misalignment, with an unequal achievement of these purposes.

Regarding the **expected outcome of innovation** (improved or new services for the citizens), Study 2 provides empirical evidence that very few municipalities in Belgium and Sweden report innovative solutions and many provide low-quality open data. It supports what previous research already stated: a misalignment between the ambitions of the expected outcomes and the input provided to realise it (open datasets and portals quality) (Elbadawi, 2012). In our empirical data, when municipalities report cases of reuse, either they are at the stage of prototypes, or they are well-functioning web applications provided by themselves. Moreover, in Study 1, all the open data initiatives we identified in Belgium, except one led by civic activists, were directly or indirectly supported by public money (awarded by a prize in a federal call for projects, paid or commissioned by public organisations, or led by organisations supported by other public funding mechanisms such as subsidies). The municipalities' roles that we identified as more involved in the innovation process (the Partner) or developing avant-garde and strategic use of open data (the Orchestrator) were only found in empirical cases from previous

literature (Study 2). The one we call the "Partner" provides briefing, guidance, feedback, and expertise to the innovators for the solutions being developed (seen in e.g., Ruijer et al., 2017; Toots, McBride, Kalvet, Krimmer, et al., 2017), and the "Orchestrator" has a global approach of data production, management and reuse, for itself and re-users (seen in e.g., Bakici, Almirall, & Wareham, 2013; Gupta, Panagiotopoulos, & Bowen, 2020).

Regarding the **expected outcome of transparency and participation in governance**, Study 1 shows one dominant form of citizen participation that is not about democratic processes. We called it *instrumental participation,* because the purpose of this participation form is the project development or service improvement, contributing to the *data value* and the *user value,* not the public or social value directly. In other words, the citizens' roles we observed contribute mostly to the development of solutions (outputs) and are more related to practical and technical tasks. For example, they provide and reuse data to develop new solutions (similar to user innovation; von Hippel, 2005a), and they develop and test solutions (with practices inspired by User-centred design; see Section 2.1.1., see Table 3, p. 36) . It is not about social and political tasks (e.g., give one´s voice in public issues). It corroborates the observation of Reggi and Dawes (2016): in the studied participatory initiatives, citizen participation for the purpose of democratic or civic engagement, is occasional and seldom supported by real public accountability mechanisms (e.g., off-line events, decision-making processes). Similar to the practices developed in open-source and innovation communities that influence the open data movement, citizen participation in open data more about collaboration than participation in society and decision-making (Harrison et al., 2012).

It means that **it is possible that some roles** (e.g., the Orchestrator and the Partner municipality for public innovation, the citizen monitoring the public action and stepping in governance processes) **are social constructs made by researchers in collaboration with practitioners**. This calls into question the realism of the municipalities' expected outcomes. When achieved in practice, in our empirical data, they are often backed by public funds, which suggests that this success may be artificially constructed and not sustainable.

Moreover, I argue that **the concept of "role expectations"** (Herrmann et al., 2004) that we rose in Study 2 **is a key concept to explain and reveal the misalignment between the municipalities' expected and realised outcomes.** The two most represented municipalities' roles (Study 2) were the "Compliant Data Provider" and the "Stand-Alone Publisher". They expected the re-users to do wonders with a data provision called "supply-driven" (Janssen et al., 2012) or "data-over-the-wall" (Sieber & Johnson, 2015), approaches criticized in research since a decade. The municipalities in these roles have broad objectives and address their data to re-users they expect being potentially "anyone". At the same time, they tend to work in isolation, provide low-quality open data, and are not reporting cases of reuse. I also understand that the "Compliant Data Providers" and the "Stand-Alone Publishers" might see open data as a sufficient tool for democratic participation, which seems to be an unrealistic expectation. **Municipalities have inflated expectations.** An explanation can be that **they misunderstand the re-users´ role, need, wants, and capabilities**. It can be a result of the application of the open data principles, which encourage the protection the re-users' anonymity and data access without registration or membership requirements (Sunlight Foundation, 2010). **They could also purposely limit their roles and alignment** because of limited resources (Study 3), or limited motivation to engage in open data (e.g., when their main motivation is to comply with the law, Study 2). In this stance, I understand they do not feel accountable for producing solutions,

providing better data or support for the re-users or citizens to re-use data. I conclude that unrealistic and unfitting role expectations hinder the realisation of expected outcomes with regards to open data.

(3) Misalignment between the publisher's input and the re-user's requirements and purposes

This misalignment happens when the input provided does not satisfy the re-users for its purpose.

Alike the misalignment 2 (between the expected and realised outcome), the scarce cases of reuse observed in our empirical data showed evidence of the misalignment between the publisher's input and the re-user's requirements and purposes.

The publisher's input is of variable form and quality, depending on how the municipality perceive their role of publisher. It can generate *role issues* between the municipality and the intended user of the input. Study 2 shows that a Compliant Data Provider is detached: it releases the minimum data with the least work, which increases the workload of the re-user. Embracing open data is not enough for a Stand-Alone Publisher: their lack of focus creates a risk of role ambiguity and discontinuity in the process. The Dedicated Publisher produces more documentation, tools, and support for the re-users, but they still expect big outcomes compare to the input provided, which can generate role overload in both sides.

The re-users' requirements and purposes are specific to the actors, their specificities (skills, wishes), and their other roles, which creates unique demands for open data, as previously stated by van Loenen (2018). It appeared clearly in our empirical data. In Study 1, a professional developer reported to not use open data and open-source tools for not meeting the quality standards of its end-users, and another needed to supplement open data with citizen-generated data to develop a functioning mobile application. Students interviewed in the first collaborative paper (Crusoe et al., 2019) also reported their frustration for to not be able to develop their dreamed solution by lack of relevant data. At the same time, we noticed that the purpose of this group of re-users was to complete an academic assignment more than developing a functioning solution. To my knowledge, none of the solutions went beyond the prototype stage after the course ended. Students and hobbyist developers are a large group of expected re-users, but their purpose is to develop their skills and have fun, not providing finished solutions (Smith & Sandberg, 2018).

It means that understanding the value creation as a technical process starting with a generic input (open data) where it is assumed that "anyone" can just "plug and play" creates misalignments. Previous research focused on the technical nature of the reuse process. Park & Gil-Garcia (2021) suggest understanding the re-user experience to solve the misalignment by matching the output to their tasks, and van Loenen (2018) suggest to identify the re-user needs to release re-user-centred open data. Other seminal papers treated these recurrent problems as socio-technical barriers to overcome (Zuiderwijk et al., 2012), or myths in the rhetoric that idealize open data's importance and simplicity (Janssen et al., 2012). I suggest considering the multiple roles held by the actors and the role-related issues unravelled by Role Theory. The multiple roles an actor can have besides being a publisher or a re-user allows us seeing the limits and opportunities in terms of value creation for the actors involved, in their respective roles, beyond the duo publisher-re-user. For example, it could open a new way to manage the value creation, where the output could be reached by targeting and combining the

involvement of actors chosen for certain skills and tasks, instead of expecting that re-users cover all the tasks and deliver their singular and finished outputs, on their own. The process could be initiated by one actor. This view also **suggests the need for a supervising role**, such as an expert in "reuse process management", or "open data value chain designer". It could be an intermediary similar to the "interpreter" of Verganti (2008), a coordinator similar to the role taken by the platform provider in the open data ecosystems described by Linåker & Runeson (2020), or a process designer similar to the researchers leading experiments in a Living Lab dedicated to open data, in Ruijer & Meijer (2020). This role could translate and harmonize the needs, purposes, and capabilities of the parties involved in a performant value chain.

**(4) Misalignment between the publishers' input, the re-user's output and the end-users' needs**

This misalignment happens when the end-users' needs are not considered by the actors upstream the value chain.

As a preliminary remark, the specificity of this misalignment is that the end-user´s needs are under-researched in open data. Hence, in Study 3, our first contribution is to bring the end-users' needs as a matter of concern and consider them in terms of information needs. Welle Donker & van Loenen (2017) and Lupi et al. (2020) show in their findings that there are misalignments between the supplied data and the re-users' requirements and information needs. In our study, instead of trying to make data accessible or train the lay users to identify and re-use data (to become re-users), we wanted to understand how to give them a central role in the identification of their information needs. We were inspired by participatory design and service design, which focus on encouraging participants to communicate their tacit needs to designers with the help of generative tools (Pallot et al., 2011; Sanders, 2002; Sanders & Stappers, 2008).

We saw evidence of risks of misalignment **between the publisher' input and end-users' needs**. At the demonstration and the evaluation of the design principles of Study 3, the representatives of municipalities (civil servants and town councillors) explained the distance they kept with the end-users, and the low intention to involve them more, due to the lack of time, lack of direct contact with their re-users, and discomfort in applying (end-)user-centred techniques. They believed that it would force them to step into a role that some of the respondents assumed to be the duty of developers (re-users). In their logic, asking the opinion of the end-users (citizens) make them feel committed to providing the solution as it was a participatory project.

It means that **the distance the municipalities as publishers keep with the end-users is due to conflicting logics underlying their open data roles** (service logic and democratic logic), **their assumptions about other roles, and assumptions carried by the open data principles** (openness, anonymity).

We saw evidence of risks of misalignment **between the re-user' input and the end-user's needs** in Study 1 and 3. In Study 1, only one case involved the end-users in the ideation stage of the development, giving them a role that allows them expressing their needs for data and wishes for reuses, close to the philosophy of participatory design (Steen et al., 2007) and social innovation (Cajaiba-Santana, 2014). The other cases of involvement were at the stage of development and test of the solution: we did not know how the needs were identified at the beginning. Similarly, in Study 3, the re-users, developers in a start-up, explained to use user

research methods (close to the philosophy of user-centred design, see Section 2.1.1.) to validate their own ideas and improve the user experience of the solution's interface, not to collect information needs. They assessed part of the information needs collected by the implementation of the design principles as new and interesting. On the other hand, other information needs were evaluated as impossible to fulfil with the available open data or needing more investigation to prove their relevance and scalability into a solution that meets a market demand. Through our research, we understood also that information needs are peculiar. In everyday contexts, the information needs depend on the end-users' previous knowledge, roles (T. D. Wilson, 1981), personal habits and preferences (Nicholas & Herman, 2010), and can disappear once they are fulfilled (once the person knows and do not need to recall).

It means that to satisfy the end-users' information needs with open data solutions and solve the misalignment, **the roles of the persons needing the information, the context of use, and nature of the information needs** (recurrent and repetitive, or unique and temporary) **are important features the publisher and re-user should consider before publishing open data and developing solutions**. Information needs can affect the demand for open data, the duration of the solutions' service life, and opportunities to scale them up to larger groups of end-users. Not all open datasets are promised to a frequent reuse, and not all information needs are worth the development of a work-intensive mobile application.

### 7.1.2. Critical outlook on open data and the use of an open data value chain

Our findings show that the actors engaged in open data (e.g., municipalities publishing open data, initiative leaders reusing data) **do not contribute together to a common value chain**. They tend to **make their roles and send expectations** to other actors they rely on (e.g., developers, citizens) to contribute to the realisation of their purpose. I understand that role-making and role-sending are **based on assumptions**: assumptions of openness and anonymity as part of the open data principles, assumptions about people's willingness to participate (Studies 1 and 2), counterpart roles' duties (Studies 2 and 3), and end-users' needs (Study 3), and in general, a certain **technological determinism** surrounding open data. I argue, in the following paragraphs, that this path is not helping to create a sustainable and broader value with open data.

The publishers and re-users **do not share a vision of the open data value chain structure (roles and division of tasks into value-adding activities) and a vision of the value to achieve (expected outcome)**. We have seen a few cases where actors send and facilitate the enactment of roles (the role-taking; Herrmann et al., 2004) that follow directly what they believe is their role, to achieve their purpose. For example, enabling municipalities (publishers) can provide the developers (re-users) with documentation and training to encourage reuse. They can organise innovation contests based on challenges/problems selected by themselves and direct the creativity of the re-users. Re-users with a clear goal in mind (a project of re-use), can expect people collecting data, and for that provide a platform and organise an event. I also saw actors taking over the whole value chain. For example, publishers becoming re-users of their own data, "internal open data", or developers of solutions for their end-users, in certain cases to the point that they integrate open data in their operations (the Orchestrator). I do not think that we can consider that as *open data best practices*, in the logic of the open data principles. I understand that these municipalities solve the lack of reuse by managing the whole value creation process by themselves, which does not prove that the principles of openness are workable and bringing any value.

I believe that **the assumptions of openness and anonymity** (of the resource, its usage, and user) encourage the actors to **make the roles** of publisher and re-user **in isolation** and **impede a real "value chain thinking"** (Piboonrungroj et al., 2017) **amongst the actors.** In the open data initiatives in cities, the actors do not see their role as unique and complementary, adding incremental value in a process that is, in the end, delivering value to an end-user. Publishers avoid enforcing a specific type of outcome on the re-users and, too often, consider the re-user as the last link in the chain. Re-users claim their right to reuse data for any purpose, without mandatory registration and reporting. Without provoked encounters between publisher and re-user focusing on the purpose of reuse and the data content, when a feedback is given by re-users to the publishers, it mostly concerns the data quality (Reggi & Dawes, 2016), which says nothing about if open data is reused, how and for what purpose. Lacking interaction, publishers and re-users are influenced by their environment and following their paradigm, isolated in their world. For example, municipalities wish that people used the data for innovation or democratic monitoring. Re-users and open data advocates wish that municipalities provided data following their standards, principles, or practices. We can see that open data moves towards more regulations, in the public sector (new PSI oriented to formats and reuse, regional charters) and civil society (multiplication of principles and charters, see Section 2.2.2.). I perceive these regulations as one world trying to enforce its mechanisms of optimization and its values towards others. The EU directives (European Parliament & European Council, 2013) follow a logic of public values (democracy, access to information) and EU economic strategies (Lassinantti, 2019). The open data principles (e.g, Sunlight Foundation, 2010) are written by supporters of open governments.

Another problem is that **open data**, understood as a means to realise higher outcomes, **is not a self-sufficient and universal building block in a universal value chain**. As argued by Evans & Campos (2013), there is a certain **technological determinism** among the proponents of open data, who try to convey that open data will bring progress and benefits. However, the assumption of openness and the current tactics to create value with open data omit two facts. (1) Open data can be transformed into many outputs (Davies, 2010), which means that the datasets can have as many value chains (i.e., ways to create value) as potential uses (outputs). (2) The models inspiring open data *do not* rely on raw resources and unguided processes. For example, the idea of open governance and citizen participation with open data originates from the right to access *information* (Charalabidis et al., 2018a; Safarov, 2019), which means *transformed data* (into texts). The idea of open innovation with open data came from open innovation practices (Chesbrough, 2011) in the private sector. However, in open innovation, companies open their knowledge, patents, and technologies, for instance, to solve problems in their industries, for their consumers. The reuse of the shared resource is constrained by a field and the needs of a market, sometimes stimulated by innovation contests based on an identified problem, and facilitated by user toolkits (Franke & von Hippel, 2003). Knowledge, innovation contests or toolkits are addressed to a profile of innovators, for a defined purpose. Similarly, open-source communities work on a defined project, for instance the development of software or computer games, with defined roles (Nakakoji et al., 2002), to meet the needs of their community. Put another way, I suspect that open data is facing a **problem of sweeping generalisation**: general accepted rules (open data definition and principles) are incorrectly used in particular instances, ignoring evidence (or confusing purposes and evidence of results; Zuiderwijk et al., 2019) and the need for a more complex understanding of the instances to prove the rules true. It could mean that previous research that assume the applicability of open

data principles might be unable to help us understand and solve the persistent issues of open data. I argue that researchers must open themselves to new sets of theories and approaches.

Hence, I argue that the **belief of "*open data for any purpose, any re-users*" is misleading the actors and hindering the role continuity in the value chains**, which should be recognised as singular. The belief that the more "raw" the data is, the more reusable it is (O'Reilly, 2011) makes the implementation of solutions difficult because the openness, as a key feature of a resource, is extended to the process of reusing it. This is not working. Zuiderwijk et al. (2019) found out that the main benefits perceived by the publishers were first operational and technical (e.g., ability to discover and reuse data in the organisation, improvement of the administrative processes), followed by the economical and then societal. There is a paradox: **the publishers prepare the data for reuse, but they do not know who their re-users and end-users are and what they want, which makes the envisioning of a value chain and alignment impossible**. One cannot develop a solution to an unknown consumer base. The idea of value chain is that each activity and actor add incremental value towards something that at the end is valuable for the customer (Harmon, 2011). It implies that the expected outcome is defined because *if no one knows what is intended, it is not possible to combine the required knowledge and resources, plan the activities to be done, and complete the process* (Piboonrungroj et al., 2017).

In the short term, acting in isolation can be comfortable. Municipalities can limit their role to data provision, as the re-use and innovation are somebody else's (tough) job. Developers and innovators are not inhibited in their creativity. In the long term, it brings inefficiencies and can impede the achievement of ambitious expected outcomes. When assumptions and beliefs drive the actors, misalignments happen. Misalignments are not systematic and depend on the contexts, the way the actors involved understand their roles, and the way they define their purpose with open data. I believe misalignments are not a bad thing if they can trigger discussion, negotiation of the roles, and increase of the value gained by the actors involved, as the latter is not obvious in open data (Lindman & Nyman, 2014).

Hence, is a "**value chain thinking**" (Piboonrungroj et al., 2017) the solution to increase the open data value, then? No, I do not believe a value chain is an easy model that can solve wicked problems. First because a value chain carries an intrinsic economic perspective (Zamora, 2016). Actors in open data are not only driven by the search and improvement of their competitive advantage in a market logic, but also by public and democratic values. The model is less relevant for issues of transparency, for example. Second, it is an analytical tool originally designed for segmented markets, in a defined industry, with the aim to increase the consumer value (Zamora, 2016). In open data, the actors come from several "industries", with their paradigms. A value chain perspective can also collide with the foundations of open data and ideals of the open data enthusiasts, as it assumes to define the outcome to achieve. Lastly, value is a concept difficult to define and measure. Therefore, it is difficult to align actors on it.

However, in combination, roles and value chain perspectives open new research avenues. I argue and conclude that **it is useful to understand the roles in open data initiatives in cities, and relevant to conceive them as transactional and negotiated and to think the value creation in terms of singular value chains instead of universal access.** It can advance research and challenge established practices, which I develop in the implications for research and practice that follow.

## 7.2. Implications for research

Open data is not giving enough results the way it is currently implemented. I showed that the concept of role brings new understandings in the misalignments and perspective of research. I believe that to harvest the value of open data the roles need to be adjusted to each other and need to be provided with the right tools and facilitating mechanisms to perform. Biddle (1986) and Herrmann et al (2004) suggest the symbolic interactionist role's perspective, where roles are defined in terms of interactions and agreements between actors, specific to a context. The use of roles, as a position in a process, determined by paradigms and rationales, in a specific context, can help researchers to reach a better conceptual clarity about the roles of the municipalities, citizens, and end-users in open data. The concept of role also allows us to study role combinations, important factors that influence the reasons why and the ways publishers release open data and re-users reuse it. An actor should not be reduced to its role of "publisher" or "re-user"; as they are also a municipality, or a civil engineer passionate about transport data, for example. Researchers should dig into the actor's multiple roles to understand the barriers and success factors of open data initiatives, beyond the technical nature of open data. These roles shape their expectations, the enactment of their open data role, *taken or made*, and to what extent they succeed in realising their expected outcome.

A value chain perspective is also an interesting analytical tool that can question the established open data definition and research streams by applying some of its concepts. First, similar to the idea of role interaction, a value chain assumes the need for governance to coordinate the actors' contribution towards a common outcome (Piboonrungroj et al., 2017). I argue that singular value chains and roles for singular outcomes need to be researched to improve the value creation with open data. Previous attempts to coordinate the actors per industry and change the open data paradigm have already been observed, for example in "data collaboratives" (Susha et al., 2017). There, the actors try to align the published data, the purpose of re-use and the expected outcome, the re-users and end-users' needs, in facilitated collaborations. Open data as a coordinated enterprise towards a shared outcome, and the paths to scale up the coordination mechanisms, is worth exploring.

## 7.3. Implications for practice: managerial recommendations

Our contributions answer the practical problem. For the practitioners (municipalities and their employees as publishers, initiative leaders as re-users, actors involved in the stimulation of data release and reuse), I suggest using the framework and artefacts as tools to reflect on their roles and possible value chains.

In Study 1, the suggested open data value chain is a visual framework intended to be a triggering tool. The value chain does not explain how to organise participation in open data initiatives, but it shows the possible entry points to involve people with or without data literacy, and the type of value resulting from the activities. It could be used by the re-users to try out new participatory activities and better align their claimed objectives and expected outcomes with the actual participation forms.

The typology of Study 2 can be used by municipalities as a tool to reflect on their desired or current approach to open data. It can help to become aware of how municipalities' roles call for certain users' roles and can help actors to reflect on whether a pair of roles is appropriate to realise a specific outcome, in their context. It can trigger question such as: "Do we have such profiles in our community? Do we believe it is the role we want to take?". Awareness is a

prerequisite to developing good strategies. The typology proposes role categories that have been labelled in a way that eases dissemination and understanding by a non-academic audience.

The design principles of **Study 3** are delivered in a popular science article, presented as a script for facilitating their implementation. The popular science article comes along with examples of tools and protocols to take an end-users centred approach in open data, developed for the instantiations. The design principles aim at encouraging the publishers and re-users to consider more the end-users' needs and integrate them into the input. It means open data published for meeting the needs of a larger group of people.

In sum, the **compilation thesis** aims at exploring how the concept of roles can be used as a catalyser to capture possible misalignments and envision new solutions. Eventually, the contributions could challenge the established open data practices and rhetoric, arouse the curiosity of the practitioners on other ways of thinking open data, and convey a more coordinated approach to open data release and reuse through the lenses of the roles in value chains.

In general, to the actors involved in open data (municipalities/publishers, re-users, digital agencies and similar actors active in the support and development of open data), I recommend:

- **Envisioning open data release and reuse as actors participating in a value chain where every actor needs to coordinate to create and capture value,** because the reasoning behind openness that encourages the actors to make roles based on assumptions and send inflated expectations hinders value creation.
- **Interacting with other actors to understand their other roles in society, their rationales, and purposes,** because it defines the way they approach open data and perceive other's roles, and because understanding the value creation as a technical process is not sufficient to solve misalignments.
- **Considering end-users' information needs,** because the roles of the people needing the information, the context of use, and the nature of the information needs (recurrent and repetitive, or unique and temporary) are important features influencing the demand for open data and the type of re-uses to satisfy the needs.

You should avoid:

- **Assuming that others will contribute to the achievement of your expected outcome,** because sent roles might never be taken. We have seen that ideal roles (e.g., the Orchestrator and the Partner municipality for public innovation or governance, the citizen monitoring the public action and stepping in governance processes) might be based on unique cases, result from misunderstanding of others' purposes and needs, and suffering from too high expectations and lack of support.

This means, specifically,

*For the less experienced municipalities (future and nascent publishers)*

Publish with a defined purpose, rely on the assets available in your community, and be result-oriented to initiate specific open data value chains. The first steps could be:

- Define what you want to achieve with open data (expected outcomes, types of innovations or type of governance made possible with open data). Think about the support, policies, and other investments it would require to be achieved and whether they are affordable for your organisation.
- At the same time, investigate the potential reuses based on the data you have available (what possible public and private interests they can satisfy). Leave aside the assumption that we cannot know the possible reuses and try to engage with people and organisations interested in your data.
- Investigate which potential actors to involve in the community (who has skills or resources, who has an interest in data, what are their purposes and needs, who could benefit from data reuses, could these actors supplement each other to deliver a sustainable outcome?).
- Define measurable objectives with re-users and end-users. Keep in mind that your expected outcomes can be different of the ones of independent re-users, and that each outcome can require specific support, policies and strategies to be achieved.
- Publish data required for the targeted outcomes and identified re-users.

*For the more experienced municipalities (publishers and possibly re-users)*

Interact with your active re-users, learn from the cases of reuse, and adapt your data release to your re-users and end-users´ information needs. The steps could be:

- Identify the actors involved in open data and analyse the cases of reuse: who was involved, what were the main tasks, roles and milestones in the value chain, what do the actors get from their involvement, are the reuses successful or not, and why?
- Stimulate the interactions with the actual re-users and end-users, and potential re-users, to understand their information needs. A suggested method is provided in Study 3. It could help to make the difference between data that can be re-used in simple visualisations, which a municipality can easily provide to help the lay citizens understand public affairs and participate (achieve public and social value), and data promised to complex reuses that could generate revenue and must satisfy higher technical requirements of re-users (economic value).
- Explore with the re-users and end-users the possible collaboration and governance mechanisms for open data. It implies negotiation and organisation of each other's resources, capabilities, and captured value, to reach identified outcomes.
- Adapt and make your data role, instead of taking a role sent that could not fit your organisation, main purposes and resources.
- Publish new data and review published data accordingly.

*For the re-users (developers, start-ups)*

- Report your cases of reuse and share your concerns and requirements to the publishers to help them seize better your profile and needs and stimulate the negotiation of roles.
- Involve the end-users to identify information needs and ideate new solutions. The user involvement should not be limited to the improvement of user-friendly interfaces since the value of a solution depends on the fulfilment of a need in the first place. A suggested method is provided in Study 3.
- If the participation is about participation in society, it can imply a substantial investment in incentives, provision of facilitation tools (also off-line), integration into other democratic mechanisms, and ideally tools to collect empirical evidence of the participation. Use external expertise and collaborate with public actors.

*For the actors supporting and developing open data (digital agencies and similar)*

- Envision your role as a process designer, that could help the publishers not only in the release of data, but also in the definition of achievable outcomes and concrete ways to realise them, the identification of external resources and needed roles in the value chain, the reflection on value capture for the engaged stakeholders. This role would require:
    - understanding the world (paradigm) of the municipalities and the world of the re-users and other stakeholders,
    - facilitating the negotiation towards a common outcome, governance mechanisms, the value captured by each party, and allocation of work and resources between roles, from the identification of the needs to the launch of the finished solution.
- This role could also facilitate the access to external expertise to design participatory activities with end-users if the actors do not have the skills in-house. Experts can help assess participants' motivations, manage expectations (to what extent their output will be used and when), and increase the performance of the activities.

## 7.4. Reflection

Before concluding, I want to reflect on pragmatism, the methods used, and the development of my knowledge and perspective on open data through my research project.

*Being pragmatic in open data research*

As defined by Goldkuhl (2012) and Avenier & Thomas (2015), the goal of knowledge creation in a pragmatic constructivist paradigm is to build intelligible models of human experience, to provide insights for organising the human experience, intentionally acting, and make a purposeful difference in practice. However, I want to stay cautious with the production of prescriptive or normative knowledge considered essential in pragmatism according to Goldkuhl (2012). I have seen that the rhetoric around the benefits of open has a significant influence on research and practice. Open data rhetoric conveys strong assumptions about what "open" should be (open principles) to reach positively connoted benefits (e.g., value,

innovation, and participation), and therefore, do what it takes to improve its implementation (rise barriers, increase adoption).

I think that the role of a researcher is to provide insights to understand a phenomenon and encourage critical thinking in both research and practice, not to prove by all means the value of one specific technological opportunity (following a diffusion paradigm). I do not position myself as an open data activist or a fierce critic. I would define my research, at least at the beginning of the project, as sceptical optimistic: I am curious about open data, realistic about its limits, critical against the research claiming its benefits without empirical evidence. In an optimistic and pragmatic way, I question and look for paths to improve the expected outcomes. I am eager to understand and find applicable solutions, more than convincing people about all the good of open data.

Being pragmatic, in my point of view, is also guaranteeing that my research output is relevant and appropriate to the current practitioners' issues and research gaps. If we observe that open data is not reaching fruition in practice, I want to understand the reasons and produce knowledge that stimulates reflection on both the purpose and the process. It means that being pragmatic, I could argue that open data is worth or not for a given problem after collecting enough evidence.

Saying that, I consciously limit my contribution, in this thesis, to understanding, explaining, and providing actionable research output, leaving the "acting" and "change-making" parts (Goldkuhl, 2012) in the hands of the practitioners. I want to make a purposeful difference by making sure that my contributions are accessible and relevant to practice and by raising their awareness and questions, not presenting them *the* answer. The design principles of Study 3 are more prescriptive by nature, but also abstract and flexible enough to allow their users to adapt them to their environment.

*Conducting a pragmatic research project: main challenge and legitimation of the knowledge produced*

All along the research process, access to the field was the main challenge. Besides the Covid-19 pandemic and confinement that lasted for half of my thesis time, overall, reaching out to open data practitioners to collect data or discuss the findings was not simple. Open data release and reuse happen now and then, most likely as part of the job of a civil servant, a student project, a hobby of a developer, or a side project of an entrepreneur. The open data maturity in Belgium and Sweden is low (van Hesteren et al., 2021). I saw in my empirical data that there are seldom contact details on the municipal web pages and open data portals, some prototypes never reach the market launch stage, and in digital projects, participation can be managed from the personal devices of the users. The organisation of participatory workshops for Study 3 also required time and energy for a low response rate.

We used different accepted strategies to check the quality of the findings and legitimate the knowledge. Study 1 was explanatory. We screened an extended number of empirical cases, conducted exploratory and confirmatory interviews, and went back and forth in the literature to ensure that our understanding was based on comprehensive data and literature (triangulation). In Study 2, we tested the typology on empirical cases and used an analytical evaluation method (Hevner et al., 2004). In Study 3, we involved end-users and publishers in the development, and re-users and publishers in the evaluation of the artefact (the design

principles). We set clear boundaries delimiting the objectives, the domain of applicability, and the audience of the findings. Our methodology raised a few questions about the recruitment process (representativeness). We argued that we developed design principles to collect information needs in the context of service development. In this case, representativeness is easier to achieve because based on user profiles, not on a population as it is in policy-making or democratic decision-making processes.

*Knowledge development and change of perspectives through the research*

As a concluding reflective thought, I would like to go back to the epistemological consideration of the constructivist pragmatic paradigm. The paradigm assumes that there is interdependence between the knowing subject and the object of study in knowledge building (Avenier & Gavard-Perret, 2012). Knowledge building starts with what the researcher knows and wants to know, and evolves as the researcher's understanding grows (see Section 3.1.2.). I mean that I am aware that the theoretical lenses I used (roles and value chain) influenced my perception of the topic an my conclusions, and that what I found changed my perception. I started my research with what I called a sceptical optimism towards open data. At the end of my thesis, I see it with more realism and pragmatism. Open data has started around 2009, a lot of effort and resources have been put into it since then. Yet, we see little of its promises kept. I believe now that open data is more a political tool than a governance or innovation means. Open data does serve a political agenda: I have perceived political or ideological motivations in municipalities that release open data and re-users that claim re-using open data. However, as a governance and innovation means, open data is still far from delivering the expected value and local successes seldom scale up. As Zuiderwijk et al. (2016) said, economic value creation with open data is a wicked problem. I think the reflection can be extended to the public and social value creation. I believe that the difficulties together with the lack of relatable, concrete benefits could lead to a stagnation of open data.

Today, I would approach open data by evaluating its opportunities with more considerations. My research ethics also questions my responsibilities as a researcher in regards with my contribution to a field and what it means, to me, to have a positive impact on society. Do I believe my work and conclusions positively impact the practitioners? I am concerned about findings that would make me recommend open data to be integrated in the work processes of municipalities, because the municipalities who did seem to achieve results (e.g., the orchestrator in Study 2), and those who not, might end-up interrupting their initiatives (as experienced by Ruijer & Meijer, 2020). However, is it ethical to ask such an investment from small municipalities with limited resources, while, as public organisations they have to fulfil their public mission in an efficient, accountable, and equal manner (Stoker, 2006)? What kind of research to conduct if I want to stay away of the current open data rhetoric, nurturing the belief that uncertain benefits are worth the effort? Today I think that, instead of asking practitioners *"do you want to reap the benefits of open data?"* and conduct research to find ways to do it, we should ask: *"what goal do you want to achieve in the first place, and how does open data compare to alternative options for successfully achieving it?"* I do not renounce to open data, but I believe that municipalities, having to work with increasingly limited means, have a duty to use public money intelligently and efficiently, for the benefit of the greatest number. Keeping these ethics principles in mind, open data has not yet proven to be the best alternative to realise the results expected by municipalities.

# Chapter 8. Conclusion

In this last chapter, I conclude by connecting the theoretical and practical research problem, and the research purpose to the key contributions. I show how I answered the research questions, present a few limitations, and suggest new leads to future research.

## 8.1. Key learnings and answers to the research problem

The **theoretical research problem** of my thesis was the lack of understanding of roles in open data initiatives in cities, specifically the roles in the shadow of the publishers and re-users: the citizens, the municipalities and their expected users, and the end-users. I conceptualised this problem in four possible misalignments centred on the actors and roles in an open data value chain. I suspected misalignments and conceptualised them at four levels, between:

(1) the actor's perception of other roles (in terms of definition, scope, and purpose);
(2) the municipalities' expected and realised outcomes;
(3) the publisher's input and the re-user's requirements and purposes;
(4) the publisher's input, the re-user's output and the end-users' needs.

I identified also a **practical problem:** the municipalities lack understanding of the other stakeholders' needs and roles. when the municipalities' input does not match the re-users or end-users' needs, it limits its reuse and can hinder the realisation of the expected outcomes (Zuiderwijk & Janssen, 2014). I observed also in practice that the lack of reuse and real impact of open data is contrasting with the optimistic rhetoric of the open data advocates. This rhetoric is vague about the steps to follow to achieve success. It encourages practitioners to mimic each other's initiatives (Zuiderwijk et al., 2019) and can leave the development of open data in a form of stagnation, blinded by ideology.

My research did not aim at proving the value of open data or identifying and solving misalignment in a systematic way. I used a pragmatic perspective on the process of open data value creation and suggested that the first step to improving it is to understand better its components (actors, roles, activities) and role mechanisms (role expectations, roles' contributions in the process).

Therefore, my **research purpose** was to explore how the concept of role can bring new insights to capture the possible misalignments in the open data value chain and envision new solutions. To address this purpose, I used roles as a lens and an open data value chain as a conceptual framework. I elucidated and described the citizens' and municipalities' diverse roles (Study 1 and 2), and I explored and prescribed possible methods to give the end-users an active role in the capture of their information needs (Study 3).

In these studies, the concept of role and value chain contributed to bring **new insights into the misalignments** we saw happening in the studied initiatives. I synthesise these insights and learnings into five key contributions.

1. The publishers' and re-users' **purpose** in a specific open data initiative **defines the way they approach open data** (whether they make or take their open data role) and the way they perceive others' roles and contributions to their initiative. As a result,
    - It influences the roles sent and expected from the citizens and the way municipalities engage (or not) with the other the actors in the value chain (re-users, end-users), which can be a source of new misalignments.

- It creates roles and role combinations unique to their initiative. Misalignments are contextual and are difficult to generalise between groups of actors.

2. I understand that the reasoning behind openness encourages actors to **make roles based on assumptions and send expectations**. For example, municipalities publishing data have inflated expectations towards the re-users, either because they misunderstand the re-users, or because they purposely limit their roles and involvement in open data. As a result,
    - The concept of "role expectations" and role-related issues are useful concepts to capture misalignments: they can be an explanation for the mismatch between input and user requirements, and between the publisher's purposes and the realised outcomes, for instance.

3. A few **ideal roles related to expected outcomes appear repetitively in previous research but not in our empirical data** (e.g., those we called the "Orchestrator" and the "Partner" municipality for public innovation, the citizen supposed to monitor the public action and step in governance processes). Similarly, we observed that many reported finished and functioning cases of reuse were funded or supported by public money. As a result,
    - It is possible that some ideal roles are social constructs made by researchers, based on ideals, research-generated through partnership collaborations and project grants, or unique cases.
    - Enthusiastic public actors self-confirm the value of open data by reusing themselves their data or sponsoring re-users;
    - In both cases, the expected outcomes are proven by artificial mechanisms, integrated and controlled processes, or best practices specific to a single case, which make them difficult to replicate in other circumstances and do not prove the workability of open data principles.

4. To **envision the value creation with open data as a technical process**, built on standardisable resources and tasks, **is not sufficient to solve the misalignments**, because actors combine roles and have purposes influenced by their other roles in society. Their other roles shape their open data role. For example, I understand that the distance the municipalities as publishers keep with the end-users is due to conflicting logic between their roles (service logic and democratic logic) and the assumptions of open data (openness, anonymity). As a result,
    - Considering the multiple roles synergies and conflicts allows us to see the limits and opportunities in terms of value creation for the actors involved, in their respective roles, and start managing the value creation process to increase the chance to achieve the expected outcomes.
    - It calls for a supervising role, such as an "open data value chain designer".

5. The **end-users**, at the very end of the value chain, **can take an active role** and help the re-user and publisher capture their information needs if the intellectual process of discovering an information need is properly facilitated. However, information needs are peculiar to the persons, their roles, and contexts. They are critical but complex

factors that publishers and re-users should integrate to develop useful and long-lasting open data solutions. As a result,

- It is possible to consider people's information needs to ideate open data solutions, which could be a way to reduce the misalignment between the publisher's input, re-users' solutions and end-users' needs.
- The nature of the information needs (personal, temporary) can affect the demand for open data and the nature of open data (limited to the activities and priorities of the publisher) raises issues for the ideals of open data (data that would serve anyone, for any purpose). In other words, not all open datasets are promised to a frequent reuse, and not all information needs are worth the development of a work-intensive mobile application.

The insights reveal **new approaches to solve misalignments**. I conclude that we should understand the roles as transactional and negotiated and the value chain as a whole entity that requires actors to agree on and plan together the expected outcomes. The assumption of openness conflicts with the minimum coordination and planning required to create value when multiple actors are involved. The misalignments are opportunities to discuss and adjust to each other along the process. I argue that we should think in terms of singular value chains for determined outcomes instead of universal access and standardisation of the resource. Hence, I argue that new ways of publishing, reusing, and conducting research on open data, with more critical eyes on its rhetoric and use of evidence, are needed. A value chain thinking is not the solution to open data´s wicked problems, but it helps to ask interesting questions, such as the management of the value generated and captured by the actors involved to achieve sustainable outcomes.

## 8.2. Limitations and future research

My contributions have some **limitations**. I already mentioned the methodological limitations in the Research Design Chapter (see Section 3.4.). In this concluding section, I discuss the limitations related to the theoretical lens and the contributions.

I used a process approach, and the linear view of a value chain is limited to show the inherently iterative nature of data re-use, the dynamics of actors' interactions, role negotiation (role-taking or role-making; Herrmann et al., 2004), and value creation or capture. My focus was on the roles of the citizens, municipalities and end-users in open data, in cities. I do not cover other significant roles, such as for example, the role and influence of the open data platforms providers on the publishers. The thesis does not cover the questions of economic value creation and the public and social value creation are part of the general context, not the findings per se. My contributions do not explain how to measure value creation.

I also see a limited generalisation of the findings due to the contexts. In Study 1 and 2, I have studied open data in three national cases, which two have a low open data maturity. The municipalities' practices in open data and initiative leaders' approaches to citizen participation can be different in other countries, especially non-European countries where the political landscape and strategies are different of the EU's.

In Study 3, we delimited our study to the capture of information needs for ideation purposes. It limits the findings to a specific stage of the service development that does not include the technical considerations (e.g., user requirements, interface development, information structure and formats). Therefore, the researchers and practitioners implementing the design

principles of Study 3 should carefully assess the value of the end-users' insights. Participatory design produces insights centred on the participants' understanding of a current situation or problem and result on incremental innovation (Norman & Verganti, 2014), participation is demanding for the people involved (Charles, 2012), and people are not always willing to contribute (Papadopoulos, 2013). Experts and lead users, with their technical and field knowledge, can envision future solutions and practices (Steen et al., 2007). The users of the design principles should critically assess these inputs, to separate the needs from the wishes, and reinterpret the possible solutions to increase their potential disruptiveness.

Hence, I see **future research** in several directions.

First, I suggest to further research roles in open data. Future research could explore the dynamics between the identified roles (Study 1 and 2), by using other perspectives such as ecosystems, complex systems, or Coordination Theory. The application of methods that include longitudinal and in-depth case studies, observations, and discussion of the identified roles with practitioners could help exploring the roles combinations, development, and enactment. For example, future research questions could be: what are the motivations of citizens in taking a role in open data initiatives? If the roles are not taken, why and what strategies to implement? Do the citizens combine roles? Do the municipalities change roles over time? Are there trajectories or patterns in these moves? What are their rationales for changing roles? Case studies through the lens of Coordination Theory and coordination mechanisms could bring interesting findings on if and how negotiation and interaction happen between the stakeholders and affect the outcomes (e.g., mutual adjustment, standardisation of Mintzberg, 1980). It could inform the strategic design and governance of open data value chains and could be one avenue to overcome the misalignments identified in this thesis.

Second, I see interesting questions to explore with the application of a value chain model in open data. We could investigate misalignments in the value chain in a more systematic way. Additionally, a value chain assumes that the search for competitive advantage is the main concern of the actors to improve the quality of their contribution, enter, and stay in the system, and that they rewarded only when the chain as a whole is effective (Walters & Lancaster, 2000). What can we learn by applying the concept of competitive advantage as a way to improve each actor´s contribution? Can open data produce any valuable outcomes if the actors involved do not see their contribution as a part of a bigger whole, which success depend on an effective coordination between all? The concepts of value creation and value capture challenge also the simplistic vision of the duo publisher-re-user dominant in previous research. What is the real value captured by a hobbyist re-user? If the hobbyist and students are the most frequent profiles of re-users, how can the municipality adjust its expected outcomes and the ways to achieve them?

Finally, I see future research based on the contributions of the studies. The usability of the framework (Study 1) and the typology (Study 2) could be further tested and developed. How usable for practitioners are the artefacts developed in the studies? Future contributions could focus on the method to develop tailor-sized open data value chains, based on the practitioners' purposes, expected outcomes and environment, and tools to identify misalignments specific to a case. More research could investigate the means and tools to assess the success of the participatory activities (Study 1) and the approaches to and experienced barriers when integrating end-users' information needs in the publishers' and re-users' work (Study 3). The

design principles are currently focussing on the capture of information needs, they could be extended to methods helping publishers to integrate the information needs in their work.

# References

Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2017). A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems. *Cities*, *64*, 47–53. https://doi.org/10.1016/j.cities.2017.01.011

Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2019). The process of open data publication and reuse. *Journal of the Association for Information Science and Technology*, *70*(3), 296–300. https://doi.org/10.1002/asi.24116

Abras, C., Maloney-krichmar, D., & Preece, J. (2004). User-Centered Design. In *Encyclopedia of Human-Computer Interaction* (pp. 1–14). Thousand Oaks: Sage Publications.

Agrawal, D., Kettinger, W. J., & Zhang, C. (2014). The Openness Challenge: Why Some Cities Take It On and Others Don't. In *Twentieth Americas Conference on Information Systems* (pp. 1–7). Savannah.

Ahlin, K., & Crusoe, J. (2022). Why Should You Believe in Open Data? A Document Study Examining Persuasion Rhetoric of OGD Benefits. In *Electronic Government Proceedings of the 21th IFIP 8.5 WG International Conference, EGOV 2022, September 6–8* (p. [ACCEPTED]). Linköping, Sweden: Electronic Government. Springer International Publishing, Cham.

Ahmadi Zeleti, F., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. *Government Information Quarterly*, *33*(3), 535–551. https://doi.org/10.1016/j.giq.2016.01.008

Al-Shboul, M. K., & A.Abrizah. (2014). Information Needs: Developing Personas of Humanities Scholars. *The Journal of Academic Librarianship*, *40*(5), 500–509. https://doi.org/10.1016/j.acalib.2014.05.016

Almirall, E., Lee, M., & Majchrzak, A. (2014). Open innovation requires integrated competition-community ecosystems: Lessons learned from civic open innovation. *Business Horizons*, *57*(3), 391–400. https://doi.org/10.1016/j.bushor.2013.12.009

Anderson, H., Havila, V., Andersen, P., & Halinen, A. (1998). Position and role-conceptualizing dynamics in business networks. *Scandinavian Journal of Management*, *14*(3), 167–186. https://doi.org/10.1016/S0956-5221(97)00037-7

Antikainen, M., Mäkipää, M., & Ahonen, M. (2010). Motivating and supporting collaboration in open innovation. *European Journal of Innovation Management*, *13*(1), 100–119. https://doi.org/10.1108/14601061011013258

Anttiroiko, A.-V. (2016). City-as-a-Platform: The Rise of Participatory Innovation Platforms in Finnish Cities. *Sustainability*, *8*(9), 922. https://doi.org/10.3390/su8090922

Armingeon, K., & Guthmann, K. (2014). Democracy in crisis? The declining support for national democracy in European countries, 2007-2011. *European Journal of Political Research*, *53*(3), 423–442. https://doi.org/10.1111/1475-6765.12046

Arnstein, S. R. (1969). A Ladder Of Citizen Participation. *Journal of the American Planning Association*, *35*(4), 216–224. https://doi.org/10.1080/01944366908977225

Ashforth, B. E. (2000). *Role transitions in organizational life: An identity-based perspective*. Mahwah NJ Lawrence Erlbaum Associates.

Assemblée Nationale, & Sénat. (2016). *LOI n° 2016-1321 du 7 octobre 2016 pour une*

*République numérique*. Retrieved from https://www.legifrance.gouv.fr/jorf/id/JORFARTI000033203423

Attard, J., Orlandi, F., & Auer, S. (2016a). Data Driven Governments: Creating Value through Open Government Data. In A. Hameurlain, J. Küng, R. Wagner, A. Anjomshoaa, P. C. K. Hung, D. Kalisch, & S. Sobolevsky (Eds.), *Transactions on Large-Scale Data-and Knowledge-Centered Systems XXVII* (Vol. 9860, pp. 84–110). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/https://doi.org/10.1007/978-3-662-53416-8_6

Attard, J., Orlandi, F., & Auer, S. (2016b). Value creation on open government data. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2016-March*, 2605–2614. https://doi.org/10.1109/HICSS.2016.326

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, *32*(4), 399–418. https://doi.org/10.1016/j.giq.2015.07.006

Avenier, M.-J., & Gavard-Perret, M.-L. (2012). Inscrire son projet de recherche dans un cadre épistémologique. In *Méthodologie de la recherche en sciences de gestion* (pp. 11–62). Pearson. https://doi.org/10.1016/B978-2-294-75066-3/00001-2

Avenier, M.-J., & Thomas, C. (2015). Finding one's way around various methodological guidelines for doing rigorous case studies: A comparison of four epistemological frameworks. *Systèmes d'information & Management*, *20*(1), 61. https://doi.org/10.3917/sim.151.0061

Bacharach, S. B., Bamberger, P., & Sonnenstuhl, W. J. (2018). The Organizational Transformation Process : The Micropolitics of Dissonance Reduction and the Alignment of Logics of Action Author ( s ): Samuel B . Bacharach , Peter Bamberger and William J . Sonnenstuhl Published by : Sage Publications , Inc . on behalf, *41*(3), 477–506.

Baker, Lynda, M. (2016). Observation: A complex research method. *Library Trends*, *55*(1), 171–189. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/3659/ResearchProcess.pdf?sequence=2

Baker, J., & Singh, H. (2019). The roots of misalignment: Insights on strategy implementation from a system dynamics perspective. *Journal of Strategic Information Systems*, *28*(4). https://doi.org/10.1016/j.jsis.2019.101576

Bakici, T., Almirall, E., & Wareham, J. (2013). A Smart City Initiative: The Case of Barcelona. *Journal of the Knowledge Economy*, *4*(2), 135–148. https://doi.org/10.1007/s13132-012-0084-9

Barber, K., & MacLellan, D. (2019). Examining Open Data at the Urban Level: An Exploration of "Wellbeing Toronto." *Journal of Urban Technology*, *26*(1), 107–121. https://doi.org/10.1080/10630732.2018.1558573

Barbosa Tavares, R., Hepworth, M., & De Souza Costa, S. M. (2011). Investigating citizens' information needs through participative research: a pilot study in Candangolândia, Brazil. *Information Development*, *27*(2), 125–138. https://doi.org/10.1177/0266666911401167

Barns, S. (2018). Smart cities and urban data platforms: Designing interfaces for smart

governance. *City, Culture and Society*, *12*(November 2017), 5–12. https://doi.org/10.1016/j.ccs.2017.09.006

Belkin, N. J., & Vickery, A. (1985). Interaction in information systems: A review of research from document retrieval to knowledge - based systems. *CIMMYT*, *35*.

Benington, J., & Moore, H. M. (2010). *Public value: Theory and practice.* Palgrave Macmillan.

Berends, J., Carrara, W., Engbers, W., & Vollers, H. (2017). *Re-using Open Data: a study on companies transforming Open Data into economic & societal value*. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0895717711001476

Berends, J., Carrara, W., & Vollers, H. (2020). *European Data Portal, Analytical Report 6. Open Data in Cities 2*.

Berners-Lee, T. (2012). 5 Star Open Data. Retrieved February 3, 2022, from https://5stardata.info/en/

Bhatt, P., Ahmad, A. J., & Roomi, M. A. (2016). Social innovation with open source software: User engagement and development challenges in India. *Technovation*, *52–53*, 28–39. https://doi.org/10.1016/j.technovation.2016.01.004

Biddle, B. (1986). Recent Developments in Role Theory. *Annual Review of Sociology*, *12*(1), 67–92. https://doi.org/10.1146/annurev.soc.12.1.67

Birkinshaw, P. (2006). Transparency as a human right. In C. Hood & D. Heald (Eds.), *Transparency: The key to better governance?* (pp. 47–58). Oxford: Oxford University Press.

Björgvinsson, E., Ehn, P., & Hillgren, P. A. (2010). Participatory design and "democratizing innovation." *ACM International Conference Proceeding Series*, (Ehn 1988), 41–50. https://doi.org/10.1145/1900441.1900448

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.

Boell, S. K., & Cecez-Kecmanovic, D. (2014). A hermeneutic approach for conducting literature reviews and literature searches. *Communications of the Association for Information Systems*, *34*(1), 257–286. https://doi.org/10.17705/1cais.03412

Bogers, M., & West, J. (2012). Managing distributed innovation: Strategic utilization of open and user innovation. *Creativity and Innovation Management*, *21*(1), 61–75. https://doi.org/10.1111/j.1467-8691.2011.00622.x

Bommert, B. (2010). Collaborative Innovation in the Public Sector. *International Public Management Review*, *11*(1), 15–33. Retrieved from http://journals.sfu.ca/ipmr/index.php/ipmr/article/view/73

Borgen, S. O., & Aarset, B. (2016). Participatory Innovation: Lessons from breeding cooperatives. *Agricultural Systems*, *145*(0442), 99–105. https://doi.org/10.1016/j.agsy.2016.03.002

Boztepe, S. (2007). User Value: Competing Theories and Models User Value. *International Journal of Design*, *1*(2007), 55–63.

Brown, T. (2009). *Change by Design: How Design Thinking Transforms Organizations and*

*Inspires Innovation*. New York: HarperBusiness.

Brown, T., & Katz, B. (2011). Change by Design. *Journal of Product Innovation Management*, *28*(3), 381–383. https://doi.org/10.1111/j.1540-5885.2011.00806.x

Brown, T., & Wyatt, J. (2010). Design Thinking for Social Innovation. *Development Outreach*, *12*(1), 29–43. https://doi.org/10.1596/1020-797X_12_1_29

Buur, J., & Larsen, H. (2010). The quality of conversations in participatory innovation. *CoDesign*, *6*(3), 121–138. https://doi.org/10.1080/15710882.2010.533185

Buur, J., & Matthews, B. (2008). Participatory Innovation. *International Journal of Innovation Management*, *12*(03), 255–273. https://doi.org/10.1142/S1363919608001996

Cabinet du Ministre de l'Agenda numérique. (2015). *Projet de note stratégique pour une stratégie fédérale Open Data*. Retrieved from http://digitalbelgium.be/wp-content/uploads/2017/07/compressed_FR-strategisch-dossier.pdf

Cabitza, F., Locoro, A., & Batini, C. (2018). Making Open Data more Personal Through a Social Value Perspective: a Methodological Approach. *Information Systems Frontiers*, 1–18. https://doi.org/10.1007/s10796-018-9854-7

Cajaiba-Santana, G. (2014). Social innovation: Moving the field forward. A conceptual framework. *Technological Forecasting and Social Change*, *82*, 42–51. https://doi.org/10.1016/j.techfore.2013.05.008

Cambridge Dictionary. (n.d.-a). Innovation (noun). Retrieved February 3, 2022, from https://dictionary.cambridge.org/dictionary/english/innovation?q=innovation+

Cambridge Dictionary. (n.d.-b). Misalignment (noun). Retrieved February 3, 2022, from https://dictionary.cambridge.org/dictionary/english/misalignment

Cambridge Dictionary. (n.d.-c). Role (noun). Retrieved February 3, 2022, from https://dictionary.cambridge.org/dictionary/english/role

Canares, M. P. (2016). Smart Cities Need to be Open: The Case of Jakarta, Indonesia. *Open Data Lab Jakarta*, 167–182. Retrieved from file:///Z:/Dropbox/TransparencyCitavi-Shared/Citavi Attachments/Canares 2016 - Smart Cities Need to be.pdf

Caragliu, A., Bo, C. Del, & Nijkamp, P. (2009). Smart Cities in Europe. In *3rd Central European Conference in Regional Science* (pp. 1–15). https://doi.org/10.1080/10630732.2011.601117

Carr, S. J., & Lassiter, A. (2017). Big Data, Small Apps: Premises and Products of the Civic Hackathon. In P. Thakuriah, N. Tilahun, & M. Zellner (Eds.) (pp. 543–559). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40902-3_29

Carrara, W., Chan, W. S., Fischer, S., Steenbergen, E. van, Capgemini Institute, & European Commission. (2015). *Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources*. *European Commission*. https://doi.org/10.2759/328101

Carroll, J. M. (1999). Five reasons for scenario-based design. *Proceedings of the Hawaii International Conference on System Sciences*, 123. https://doi.org/10.1109/hicss.1999.772890

Carroll, N., Richardson, I., & Whelan, E. (2012). Service Science: An Actor-Network Theory

Approach. *International Journal of Actor-Network Theory and Technological Innovation*, *4*(3), 51–69. https://doi.org/10.4018/jantti.2012070105

Chan, C. M. L. (2013). From open data to open innovation strategies: Creating e-services using open government data. *Proceedings of the Annual Hawaii International Conference on System Sciences*, (January 2013), 1890–1899. https://doi.org/10.1109/HICSS.2013.236

Chandra, L., Seidel, S., & Gregor, S. (2015). Prescriptive Knowledge in IS Research: Conceptualizing Design Principles in Terms of Materiality, Action, and Boundary Conditions. In *2015 48th Hawaii International Conference on System Sciences* (pp. 4039–4048). IEEE. https://doi.org/10.1109/HICSS.2015.485

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018a). The Open Data Landscape. In *The World of Open Data* (Vol. 28, pp. 1–10). https://doi.org/10.1007/978-3-319-90850-2

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018b). *The World of Open Data* (Vol. 28). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-90850-2

Charles, J. (2012). Les charges de la participation. *SociologieS*, *Premiers t*, 0–17. https://doi.org/10.4000/sociologies.4151

Chesbrough, H. (2003). The era of open innovation. *MIT Sloan Management Review*, *44*(3), 35–41.

Chesbrough, H. (2011). The Era of Open Innovation. *MIT Sloan Management Review*, *Top 10 les*(January), 35–41.

Collins Dictionary. (n.d.). Actor (noun). Retrieved February 3, 2022, from https://www.collinsdictionary.com/dictionary/english/actor

Conradie, P., Mulder, I., & Choenni, S. (2012). Rotterdam open data: Exploring the release of Public Sector Information through co-creation. In *2012 18th International ICE Conference on Engineering, Technology and Innovation* (pp. 1–10). IEEE. https://doi.org/10.1109/ICE.2012.6297651

Conseil d'Etat. (2019). *Arrêté royal du 2 juin 2019 relatif à la réutilisation des informations du secteur public*. Retrieved from http://www.ejustice.just.fgov.be/eli/arrete/2019/06/02/2019013333/justel

Conseil d'Etat, & Chambre des représentants. (2016). *Loi du 4 mai 2016 relative à la réutilisation des informations du secteur public*. Retrieved from http://www.ejustice.just.fgov.be/eli/loi/2016/05/04/2016009236/justel

Cooper, R. G. (2018). The drivers of success in new-product development. *Industrial Marketing Management*, (January), 1–12. https://doi.org/10.1016/j.indmarman.2018.07.005

Cordella, A., & Paletti, A. (2019). Government as a platform, orchestration, and public value creation: The Italian case. *Government Information Quarterly*, *36*(4), 101409. https://doi.org/10.1016/j.giq.2019.101409

Corrales-Garay, D., Mora-Valentín, E.-M., & Ortiz-de-Urbina-Criado, M. (2019). Open Data for

Open Innovation: An Analysis of Literature Characteristics. *Future Internet*, *11*(3), 77. https://doi.org/10.3390/fi11030077

Corrales-Garay, D., Ortiz-De-Urbina-Criado, M., & Mora-Valentin, E. M. (2020). A Research Agenda on Open Data Impact Process for Open Innovation. *IEEE Access*, *8*, 34696–34705. https://doi.org/10.1109/ACCESS.2020.2974378

Corsaro, D., & Snehota, I. (2011). Alignment and misalignment in business relationships. *Industrial Marketing Management*, *40*(6), 1042–1054. https://doi.org/10.1016/j.indmarman.2011.06.038

Creswell, J. W., & Creswell, J. D. (2018). *Research Design. Qualitative, Quantitative, and Mixed Methods Approaches* (Fifth Edit). Sage Publications.

Crusoe, J. (2021). *Open Government Data as a Reform and Ecosystem : A conceptual framework for evolution and health* (Vol. 810). Linköping: Linköping University Electronic Press. https://doi.org/10.3384/diss.diva-177827

Crusoe, J., & Ahlin, K. (2019). Users ' activities for using Open Government Data - a process framework. *Transforming Government: People, Process and Policy*, 1–33. https://doi.org/10.1108/TG-04-2019-0028

Crusoe, J., Gebka, E., & Ahlin, K. (2020). Open Government Data from the Perspective of Information Needs - A Tentative Conceptual Model. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 12219 LNCS, pp. 250–261). https://doi.org/10.1007/978-3-030-57599-1_19

Crusoe, J., Simonofski, A., Clarinval, A., & Gebka, E. (2019). The Impact of Impediments on Open Government Data Use: Insights from Users. In *2019 13th International Conference on Research Challenges in Information Science (RCIS)* (pp. 1–12). IEEE. https://doi.org/10.1109/RCIS.2019.8877055

Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science*, *32*(5), 554–571. https://doi.org/10.1287/mnsc.32.5.554

Davies, T. (2010). Open data, democracy and public sector reform: A look at open government data use from data. gov. uk, 1–47. Retrieved from http://www.opendataimpacts.net/report/

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*, *33*(1), 15–27. https://doi.org/10.1016/j.giq.2016.01.003

De Tuya, M., Cook, M., Sutherland, M. K., & Luna-Reyes, L. F. (2017). Information requirements to create public value: sharing and opening data to address urban blight. *Transforming Government: People, Process and Policy*, *11*(1), 79–98. https://doi.org/10.1108/TG-12-2015-0054

De Vries, H., Bekkers, V., & Tummers, L. (2016). Innovation in the public sector: A systematic review and future research agenda. *Public Administration*, *94*(1), 146–166. https://doi.org/10.1111/padm.12209

Dedehayir, O., Mäkinen, S. J., & Ortt, J. R. (2016). Roles during innovation ecosystem genesis :

A literature review. *Technological Forecasting & Social Change*. https://doi.org/10.1016/j.techfore.2016.11.028

Dell'Era, C., & Landoni, P. (2014). Living lab: A methodology between user-centred design and participatory design. *Creativity and Innovation Management*, *23*(2), 137–154. https://doi.org/10.1111/caim.12061

Demil, B., & Lecocq, X. (2006). Neither Market nor Hierarchy nor Network: The Emergence of Bazaar Governance. *Organization Studies*, *27*(10), 1447–1466. https://doi.org/10.1177/0170840606067250

Denscombe, M. (2010). *The Good Research Guide for small-scale social research projects* (Fourth Edi). Open University Press.

Desouza, K. C., & Bhagwatwar, A. (2012). Citizen Apps to Solve Complex Urban Problems. *Journal of Urban Technology*, *19*(3), 107–136. https://doi.org/10.1080/10630732.2012.673056

Devadason, F. J., & Lingam, P. P. (1997). A Methodology for the Identification of Information Needs of Users. *IFLA Journal*, *23*(1), 41–51. https://doi.org/10.1177/034003529702300109

Dresch, A., Lacerda, D. P., & Antunes, J. A. V. (2015). *Design science research: A method for science and technology advancement*. *Design Science Research: A Method for Science and Technology Advancement*. https://doi.org/10.1007/978-3-319-07374-3

Drori, I., Sheaffer, Z., Weber, Y., & Landau, D. (2009). Misalignment as legacy of an industrial leader. *European Journal of International Management*, *3*(1), 75–91. https://doi.org/10.1504/EJIM.2009.022635

Eckert, C. M., & Stacey, M. K. (2010). What is a Process Model? Reflections on the Epistemology of Design Process Models. In *Modelling and Management of Engineering Processes* (pp. 3–14). London: Springer London. https://doi.org/10.1007/978-1-84996-199-8_1

Edwards-Schachter, M. (2018). The nature and variety of innovation. *International Journal of Innovation Studies*, *2*(2), 65–79. https://doi.org/10.1016/j.ijis.2018.08.004

Ehrlich, T. (2000). *Civic responsibility and higher education*. Greenwood Publishing Group.

Elbadawi, I. A. (2012). The state of open government data in GCC countries. In *Proceedings of the 12th European Conference on eGovernment. Vol. 36. No. 41.* (p. 193).

Emerson, J., Wachowicz, J., & Chun, S. (2000). *Social Return on Investment: Exploring Aspects of Value Creation in the Nonprofit Sector*. *The Roberts Enterprise Development Fund*.

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning. *Ergonomics*, *43*(7), 960–974. https://doi.org/10.1080/001401300409143

Erdelez, S. (2005). Information Encountering: It's More Than Just Bumping into Information. *Bulletin of the American Society for Information Science and Technology*, *25*(3), 26–29. https://doi.org/10.1002/bult.118

European Commission. (2020a). *Digital Economy and Society Index (DESI) Sweden*. *Shaping Europe's digital future*. Retrieved from

https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=66932

European Commission. (2020b). *Open Data Maturity: Country factsheet Sweden*. Retrieved from https://data.europa.eu/sites/default/files/country-factsheet_sweden_2020.pdf

European Open Data Portal. (2019). What is Open Data? Retrieved August 14, 2019, from https://data.europa.eu/elearning/en/module1/#/id/co-01

European Parliament, & European Council. (2003). *Directive on public access to environmental information*. Directive 2003/4/EC. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32003L0004

European Parliament, & European Council. (2013). *Directive on the re-use of public sector information*. Directive 2013/37/EU. Retrieved from https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32013L0037

European Parliament, & European Council. (2019). *Directive on open data and the re-use of public sector information*. Directive (EU) 2019/1024. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024

Evans, A. M., & Campos, A. (2013). Open Government Initiatives: Challenges of Citizen Participation. *Journal of Policy Analysis and Management*, *32*(1), 172–185. https://doi.org/10.1002/pam.2165

Farnham, David, & Horton, S. (1995). The Political Economy of Public Sector Change. In *Managing the New Public Services* (pp. 3–26). Basingstoke: Macmillan.

Foulonneau, M., Martin, S., & Turki, S. (2014). How Open Data Are Turned into Services? In *International Conference on Exploring Services Science* (pp. 31–39). https://doi.org/10.1007/978-3-319-04810-9_3

Franke, N., & von Hippel, E. (2003). Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software. *Research Policy*, *32*(7), 1199–1215. https://doi.org/10.1016/S0048-7333(03)00049-0

Franke, N., von Hippel, E., & Schreier, M. (2006). Finding Commercially Attractive User Innovations: A Test of Lead-User Theory*. *Journal of Product Innovation Management*, *23*(4), 301–315. https://doi.org/10.1111/j.1540-5885.2006.00203.x

Futurocité. (2021). Séminaire Smart Gouvernance 2021: Webinaires. Retrieved February 3, 2022, from http://www.futurocite.be/activites/seminaire-smart-gouvernance/seminaire-smart-governance-2021/webinaires/webinaire-1-lopen-data-un-levier-concret-pour-les-politiques-communales/

Gascó-Hernández, M., Martin, E. G., Reggi, L., Pyo, S., & Luna-Reyes, L. F. (2018). Promoting the use of open government data: Cases of training and engagement. *Government Information Quarterly*, *35*(2), 233–242. https://doi.org/10.1016/j.giq.2018.01.003

Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory*. Chicago: Aldine.

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, *21*(2), 135–146. https://doi.org/10.1057/ejis.2011.54

Goldkuhl, G., & Sjöström, J. (2018). Design Science in the Field: Practice Design Research (pp.

67–81). https://doi.org/10.1007/978-3-319-91800-6_5

Gonzalez-Zapata, F., & Heeks, R. (2015). The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly*, *32*(4), 441–452. https://doi.org/10.1016/j.giq.2015.09.001

Gouvernement de la Région de Bruxelles-Capitale. (2016). *Ordonnance du 27 octobre 2016 visant à l'établissement d'une politique de données ouvertes (Open Data) et portant transposition de la Directive 2013/37/UE du Parlement européen et du Conseil du 26 juin 2013 modifiant la Directive 2003/98/CE du Parlement e*. Retrieved from http://www.ejustice.just.fgov.be/eli/ordonnance/2016/10/27/2016031739/justel

Gouvernement de la Région de Bruxelles-Capitale. (2018). *Arrêté du Gouvernement de la Région de Bruxelles-Capitale du 1er février 2018 portant exécution de l'ordonnance du 27 octobre 2016 visant à l'établissement d'une politique de données ouvertes (open data) et portant transposition de la directive 2013/37/UE*. Retrieved from http://www.ejustice.just.fgov.be/eli/arrete/2018/02/01/2018010589/justel

Gray, J. (2014). Towards a Genealogy of Open Data. *SSRN Electronic Journal*, *42*(3–4), 323–363. https://doi.org/10.2139/ssrn.2605828

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, *37*(2), 337–355. Retrieved from https://www.jstor.org/stable/43825912

Gregor, S., Kruse, L., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, *21*, 1622–1652. https://doi.org/10.17705/1jais.00649

Guba, E. ., & Lincoln, Y. S. (1989). *Fourth Generation Evaluation*. London: Sage Publications.

Guba, E. G., & Lincoln, Y. S. (1998). Competing Paradigms in Qualitative Research. In *The Landscape of Qualitative Research* (pp. 195–220). London: Sage Publications.

Gupta, A., Panagiotopoulos, P., & Bowen, F. (2020). An orchestration approach to smart city data ecosystems. *Technological Forecasting and Social Change*, *153*(January), 119929. https://doi.org/10.1016/j.techfore.2020.119929

Ham, J., Koo, Y., & Lee, J. (2019). Provision and usage of open government data: strategic transformation paths. *Industrial Management & Data Systems*, *119*(8), 1841–1858. https://doi.org/10.1108/IMDS-04-2019-0218

Hancock, D. R., & Algozzine, B. (2006). *Doing Case Study Research: A Practical Guide for Beginning Researchers*. *Dalton Transactions* (Vol. 47). Teachers College Press. https://doi.org/-

Harmon, P. (2011). How Do Processes Create Value? *BPTrends*, (February), 1–13. Retrieved from www.bptrends.com

Harrison, T. M., Guerrero, S., Burke, G. B., Cook, M., Cresswell, A., Helbig, N., … Pardo, T. (2012). Open government and e-government: Democratic challenges from a public value perspective. *Information Polity*, *17*(2), 83–97. https://doi.org/10.3233/IP-2012-0269

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). *Big Data for Big Business?*

*Cambridge Service Alliance*. https://doi.org/10.1016/j.im.2014.08.008

Heath, D., Singh, R., & Shephard, B. (2013). Approaching strategic misalignment from an organizational view of business processes. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 4055–4064. https://doi.org/10.1109/HICSS.2013.99

Heikkinen, M. T., Mainela, T., Still, J., & Tähtinen, J. (2007). Roles for managing in mobile service development nets. *Industrial Marketing Management*, *36*(7), 909–925. https://doi.org/10.1016/j.indmarman.2007.05.014

Hellberg, A.-S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. *Transforming Government: People, Process and Policy*, *9*(1), 35–51. https://doi.org/10.1108/TG-04-2014-0013

Hemerly, J. (2013). Public policy considerations for data-driven innovation. *Computer*, *46*(6), 25–31. https://doi.org/10.1109/MC.2013.186

Hemmersam, P., Martin, N., Westvang, E., Aspen, J., & Morrison, A. (2015). Exploring Urban Data Visualization and Public Participation in Planning. *Journal of Urban Technology*, *22*(4), 45–64. https://doi.org/10.1080/10630732.2015.1073898

Herrmann, T., Jahnke, I., & Loser, K. U. (2004). The Role Concept as a Basis for Designing Community Systems. *Cooperative Systems Design, Scenario-Based Design of Collaborative Systems*, 163–178.

Hess, A., Diebold, P., & Seyff, N. (2017). Towards Requirements Communication and Documentation Guidelines for Agile Teams. In *2017 IEEE 25th International Requirements Engineering Conference Workshops (REW)* (pp. 415–418). IEEE. https://doi.org/10.1109/REW.2017.64

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, *19*(2), 87–92. https://doi.org/http://aisel.aisnet.org/sjis/vol19/iss2/4

Hevner, A. R., March, S. T., Park, J., Ram, S., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Hey, J. (2004). The Data, Information, Knowledge, Wisdom Chain: The Metaphorical link. *Intergovernmental Oceanographic Commission*, *26*, 1–18.

Hivon, J., & Titah, R. (2017). Conceptualizing citizen participation in open data use at the city level. *Transforming Government: People, Process and Policy*, *11*(1), 99–118. https://doi.org/10.1108/TG-12-2015-0053

Hjalmarsson, A., Johannesson, P., Jüll-Skielse, G., & Rudmark, D. (2014). Beyond innovation contests: A framework of barriers to open innovation of digital services. In *Ecis 2014 22nd european conference on information systems*. Tel Aviv.

Hogan, M., Ojo, A., Harney, O., Ruijer, E., Meijer, A., Andriessen, J., … Malandrino, D. (2017). Governance, Transparency and the Collaborative Design of Open Data Collaboration Platforms: Understanding Barriers, Options, and Needs (pp. 299–332). https://doi.org/10.1007/978-3-319-63743-3_12

Hörold, S., Mayas, C., & Krömker, H. (2012). Identifying the information needs of users in

public transport. *Advances in Human Aspects of Road and Rail Transportation*, *1*, 331–340.

Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Huang, R., Lai, T., & Zhou, L. (2017). Proposing a framework of barriers to opening government data in China. *Library Hi Tech*, *35*(3), 421–438. https://doi.org/10.1108/LHT-01-2017-0014

Hunnius, S., & Krieger, B. (2014). The Social Shaping of Open Data through Administrative Processes. In *Proceedings of The International Symposium on Open Collaboration* (pp. 1–5). New York, NY, USA: ACM. https://doi.org/10.1145/2641580.2641601

Huyer, E., & van Knippenberg, L. (2020). *The Economic Impact of Open Data. Opportunities for value creation in Europe*. Retrieved from https://data.europa.eu/sites/default/files/the-economic-impact-of-open-data.pdf

Iivari, J. (2015). Distinguishing and contrasting two strategies for design science research. *European Journal of Information Systems*, *24*(1), 107–115. https://doi.org/10.1057/ejis.2013.35

Iivari, J., Rotvit Perlt Hansen, M., & Haj-Bolouri, A. (2020). A proposal for minimum reusability evaluation of design principles. *European Journal of Information Systems*, 1–18. https://doi.org/10.1080/0960085X.2020.1793697

Iivari, J., & Venable, J. (2009). Action research and design science research - Seemingly similar but decisively dissimilar. *17th European Conference on Information Systems, ECIS 2009*.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, *29*(4), 258–268. https://doi.org/10.1080/10580530.2012.716740

Janssen, M., & Zuiderwijk, A. (2014). Infomediary Business Models for Connecting Open Data Providers and Users. *Social Science Computer Review*, *32*(5), 694–711. https://doi.org/10.1177/0894439314525902

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2013). The Generative Mechanisms Of Open Government Data. In *ECIS 2013 Proceedings* (Vol. 12). https://doi.org/10.5278/ojs.persk..v12i23.489

Jetzek, T., Avital, M., & Bjorn-Andersen, N. (2014). Data-driven innovation through open government data. *Journal of Theoretical and Applied Electronic Commerce Research*, *9*(2), 100–120. https://doi.org/10.4067/S0718-18762014000200008

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). Generating Value from Open Government Data. In *Thirty Fourth International Conference on Information Systems* (pp. 1–20). Milan. https://doi.org/10.1007/978-3-662-43459-8_5

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2014). Generating Sustainable Value from Open Data in a Sharing Society. In *Creating Value for All Through IT* (pp. 62–82).

Johnson, P., & Robinson, P. (2014). Civic Hackathons: Innovation, procurement, or civic engagement? *Review of Policy Research*, *31*(4), 349–357.

https://doi.org/10.1111/ropr.12074

Juell-Skielse, G., Hjalmarsson, A., Johannesson, P., & Rudmark, D. (2014). Is the public motivated to engage in open data innovation? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8653 LNCS*, 277–288. https://doi.org/10.1007/978-3-662-44426-9_23

Kahn, K. B. (2018). Understanding innovation. *Business Horizons*, *61*(3), 453–460. https://doi.org/10.1016/j.bushor.2018.01.011

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). *Organizational stress: Studies in role conflict and ambiguity*. John Wiley.

Kaplinsky, R., & Morris, M. (2001). A handbook for value chain research, (January 2001).

Karat, J., & Karat, C. M. (2003). The evolution of user-centered focus in the human-computer interaction field. *IBM Systems Journal*, *42*(4), 532–541. https://doi.org/10.1147/sj.424.0532

Karlsson, F., Holgersson, J., Söderström, E., & Hedström, K. (2012). Exploring user participation approaches in public e-service development. *Government Information Quarterly*, *29*(2), 158–168. https://doi.org/10.1016/j.giq.2011.07.009

Khayyat, M., & Bannister, F. (2017). Towards a model for facilitating and enabling co-creation using open government data. *Information Polity*, *22*(4), 211–231. https://doi.org/10.3233/IP-170406

Kohler, Fueller, Matzler, Stieger, & Füller. (2011). Co-Creation in Virtual Worlds: The Design of the User Experience. *MIS Quarterly*, *35*(3), 773. https://doi.org/10.2307/23042808

Komninos, N., Pallot, M., & Schaffers, H. (2013). Special Issue on Smart Cities and the Future Internet in Europe. *Journal of the Knowledge Economy*, *4*(2), 119–134. https://doi.org/10.1007/s13132-012-0083-x

Krishna, R., Kummitha, R., & Crutzen, N. (2017). How do we understand smart cities ? An evolutionary perspective. *Cities*, *67*(March), 43–52. https://doi.org/10.1016/j.cities.2017.04.010

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the Association for Information Science and Technology*, (42), 361–371.

Kujala, S. (2003). User involvement: A review of the benefits and challenges. *Behaviour and Information Technology*, *22*(1), 1–16. https://doi.org/10.1080/01449290301782

Lassinantti, J. (2019). Re-Use of Public Sector Open Data. *International Journal of Enterprise Information Systems*, *1*(May), 1–29.

Lassinantti, J., Ståhlbröst, A., & Runardotter, M. (2018). Relevant social groups for open data use and engagement. *Government Information Quarterly*, (October), 1–14. https://doi.org/10.1016/j.giq.2018.11.001

Latour, B. (2005). *Reassembling the social. An Introduction to Actor-Network-Theory*. New York: Oxford University Press.

Le Moigne, J.-L. (1995). *Les Épistémologies constructivistes*. Paris: PUF.

Linåker, J., & Runeson, P. (2020). Collaboration in Open Government Data Ecosystems: Open Cross-sector Sharing and Co-development of Data and Software (pp. 290–303). https://doi.org/10.1007/978-3-030-57599-1_22

Linders, D. (2012). From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly*, *29*(4), 446–454. https://doi.org/10.1016/j.giq.2012.06.003

Lindman, J., & Nyman, L. (2014). The Businesses of Open Data and Open Source: Some Key Similarities and Differences. *Technology Innovation Management Review*, *4*(1), 12–17. https://doi.org/10.22215/timreview/757

Liu, Y., Jiang, C., & Li, S. (2017). Research on the Evaluation of Urban Open Data. *World Journal of Engineering and Technology*, *05*(03), 122–134. https://doi.org/10.4236/wjet.2017.53B014

Lupi, L., Antonini, A., De Liddo, A., & Motta, E. (2020). Actionable Open Data: Connecting City Data to Local Actions. *Journal of Community Informatics*, *16*, 3–25.

Ma, R., & Lam, P. T. I. (2019). Investigating the barriers faced by stakeholders in open data development: A study on Hong Kong as a "smart city." *Cities*, *92*(June 2018), 36–46. https://doi.org/10.1016/j.cities.2019.03.009

Magalhaes, G., & Roseira, C. (2017). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*, pp. 1–10. https://doi.org/10.1016/j.giq.2017.08.004

Malone, T. W., & Crowston, K. (1990). What is coordination theory and how can it help design cooperative work systems? In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work - CSCW '90* (pp. 357–370). New York, New York, USA: ACM Press. https://doi.org/10.1145/99332.99367

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2

March, & Storey. (2008). Design Science in the Information Systems Discipline: An Introduction to the Special Issue on Design Science Research. *MIS Quarterly*, *32*(4), 725. https://doi.org/10.2307/25148869

Markham, S. K., Ward, S. J., Aiman-Smith, L., & Kingon, A. I. (2010). The valley of death as context for role theory in product innovation. *Journal of Product Innovation Management*, *27*(3), 402–417. https://doi.org/10.1111/j.1540-5885.2010.00724.x

Martin, C. (2014). Barriers to the Open Government Data Agenda: Taking a Multi-Level Perspective. *Policy & Internet*, *6*(3), 217–240. https://doi.org/10.1002/1944-2866.POI367

Matheus, R., Janssen, M., & Maheshwari, D. (2018). Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. *Government Information Quarterly*, (November 2017), 101284. https://doi.org/10.1016/j.giq.2018.01.006

Mattimoe, R., & Seal, W. (2011). Pricing in a Service Sector Context: Accounting and Marketing Logics in the Hotel Industry. *European Accounting Review*, *20*(2), 355–388.

https://doi.org/10.1080/09638180.2010.493657

McAvinia, C. (2016). Activity Theory. In *Online Learning and its Users* (pp. 59–100). Elsevier. https://doi.org/10.1016/B978-0-08-100626-9.00003-4

McBride, K., Aavik, G., Toots, M., Kalvet, T., & Krimmer, R. (2019). How does open government data driven co-creation occur? Six factors and a 'perfect storm'; insights from Chicago's food inspection forecasting model. *Government Information Quarterly*, *36*(1), 88–97. https://doi.org/10.1016/j.giq.2018.11.006

McGann, M., Blomkamp, E., & Lewis, J. M. (2018). The rise of public sector innovation labs: experiments in design thinking for policy. *Policy Sciences*, *51*(3), 249–267. https://doi.org/10.1007/s11077-018-9315-7

Meijer, A. (2012). The Do It Yourself State. *Information Polity*, *17*(3,4), 303–314. https://doi.org/10.3233/IP-2012-000283

Meijer, A., Curtin, D., & Hillebrandt, M. (2012). Open government: connecting vision and voice. *International Review of Administrative Sciences*, *78*(1), 10–29. https://doi.org/10.1177/0020852311429533

Meijer, A., & Potjer, S. (2018). Citizen-generated open data: An explorative analysis of 25 cases. *Government Information Quarterly*, *35*(4), 613–621. https://doi.org/10.1016/j.giq.2018.10.004

Mergel, I., Kleibrink, A., & Sörvik, J. (2018). Open data outcomes: U.S. cities between product and process innovation. *Government Information Quarterly*, *35*(4), 622–632. https://doi.org/10.1016/j.giq.2018.09.004

Mertens, D. M. (2019). An introduction to research. In *Research and Evaluation in Education and Psychology Integrating Diversity With Quantitative, Qualitative, and Mixed Methods* (5th ed., pp. 1–46). Gallaudet University, USA.

Mesquita, M. A., Luciano, E. M., Rafael, M., & Wiedenhöft, G. C. (2020). Discussing the Twofold Role of Government - Provider and User - in the Open Government Data Ecosystem. In *EGov-CEDEM-EPart Conference Proceedings*. Linköping.

Miles, R. H. (1977). Role-Set Configuration as a Predictor of Role Conflict and Ambiguity in Complex Organizations. *Sociometry*, *40*(1), 21. https://doi.org/10.2307/3033542

Ministère de la Communauté française. (2017). *Décret conjoint de la Région wallonne et de la Communauté française relatif à la réutilisation des informations du secteur public et visant à l'établissement d'une politique de données ouvertes ("Open Data")*. Retrieved from http://www.ejustice.just.fgov.be/eli/decret/2017/07/19/2017031098/justel

Mintzberg, H. (1980). Structure in 5 ' s : A Synthesis of the Research on Organization Design Author ( s ): Henry Mintzberg Published by : INFORMS Stable URL : http://www.jstor.org/stable/2630506 JSTOR is a not-for-profit service that helps scholars , researchers , and student. *Management Science*, *26*(3), 322–341.

Mirri, S., Prandi, C., Salomoni, P., Callegati, F., & Campi, A. (2014). On combining crowdsourcing, sensing and open data for an accessible smart city. In *Proceedings - 2014 8th International Conference on Next Generation Mobile Applications, Services and Technologies, NGMAST 2014* (pp. 294–299). https://doi.org/10.1109/NGMAST.2014.59

Montano, E., Silva, D. P. d., Pawelke, A., Alonso, J. M., Shekhar, S., Zigoni, C., & Canares, M. (2014). *Exploring emerging impacts of open data in developing countries: Phase One.* Retrieved from http://webfoundation.org/docs/2016/07/ODDC1-Followup-Evaluation-Final-Version.pdf

Moore, M. H. (1995). *Creating public value: Strategic management in government*. Harvard university press.

Moore, N. (1998). *The information society. World Information Report 97/98*.

Morgan, D. L. (2007). Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, *1*(1), 48–76. https://doi.org/10.1177/2345678906292462

Mulder, I., & Stappers, P. J. (2009). Co-creating in practice: Results and challenges. In *2009 IEEE International Technology Management Conference (ICE)* (pp. 1–8). IEEE. https://doi.org/10.1109/ITMC.2009.7461369

Muller, M. J., & Kuhn, S. (1993). Participatory design. *Communications of the ACM*, *36*, 24–26.

Nakakoji, K., Yamamoto, Y., Nishinaka, Y., Kishida, K., & Ye, Y. (2002). Evolution Patterns of Open-Source Software Systems and Communities. *International Workshop on Principles of Software Evolution (IWPSE)*, (January 2001), 76–85. https://doi.org/10.1145/512054.512055

Nam, T. (2012). Suggesting frameworks of citizen-sourcing via Government 2.0. *Government Information Quarterly*, *29*(1), 12–20. https://doi.org/10.1016/j.giq.2011.07.005

Nam, T., & Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th Annual International Digital Government Research Conference on Digital Government Innovation in Challenging Times - dg.o '11* (p. 282). https://doi.org/10.1145/2037556.2037602

Neves, F. T., de Castro Neto, M., & Aparicio, M. (2020). The impacts of open data initiatives on smart cities: A framework for evaluation and monitoring. *Cities*, *106*(November 2018). https://doi.org/10.1016/j.cities.2020.102860

Nicholas, D., & Herman, E. (2010). *Assessing Information Needs in the Age of the Digital Consumer*. *Assessing Information Needs in the Age of the Digital Consumer*. Routledge. https://doi.org/10.4324/9780203855799

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336–359. https://doi.org/10.1057/ejis.2012.26

Nielsen Norman Group. (n.d.). The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities. Retrieved February 22, 2022, from https://www.nngroup.com/articles/participation-inequality/

Norman, D. A. (1986). Cognitive engineering. In *User-Centered System Design: New Perspectives on Human Computer Interaction* (pp. 31–65). London: Lawrence Erbaum Associated. Retrieved from https://pdfs.semanticscholar.org/57f1/76992f92ae559d9c110211d7f04c5143cb44.pdf

Norman, D. A., & Verganti, R. (2014). Incremental and Radical Innovation: Design Research vs.

Technology and Meaning Change. *Design Issues*, *30*(1), 78–96. https://doi.org/10.1162/DESI_a_00250

Nørreklit, H., Nørreklit, L., & Mitchell, F. (2010). Paradigms and pragmatic constructivism: A reply. *Accounting, Auditing and Accountability Journal*, *23*(6), 764–773. https://doi.org/10.1108/09513571011065862

Nørreklit, H., Raffnsoe-Moller, M., & Mitchell, F. (2016). Editorial: A pragmatic constructivist approach to accounting practice and research. *Qualitative Research in Accounting and Management*, *13*(3), 266–277. https://doi.org/10.1108/QRAM-05-2016-0039

Nørreklit, L. (2013). Reality as a construct: outline of a pragmatic constructivist perspective. *Proceedings of Pragmatic Constructivism*, *3*(2), 57–66.

Noveck, B. S. (2009). *Wiki government: How technology can make government better, democracy stronger, and citizens more powerful*. (Brookings Institution Press, Ed.). Washington DC.

Nyström, A.-G., Leminen, S., Westerlund, M., & Kortelainen, M. (2014). Actor roles and role patterns influencing innovation in living labs. *Industrial Marketing Management*, *43*(3), 483–495. https://doi.org/10.1016/j.indmarman.2013.12.016

O'Connor, K., Janenova, S., & Knox, C. (2019). Open Government in Authoritarian Regimes. *International Review of Public Policy*, *1*(1), 65–82. https://doi.org/10.4000/irpp.325

O'Reilly, T. (2011). Government as a Platform. *Innovations: Technology, Governance, Globalization*, *6*(1), 13–40. https://doi.org/10.1162/INOV_a_00056

Öberg, C. (2010). Customer roles in innovations. *International Journal of Innovation Management*, *14*(6), 989–1011. https://doi.org/10.1142/S1363919610002970

Ojo, A., Curry, E., & Zeleti, F. A. (2015). A tale of open data innovations in five smart cities. In *Proceedings of the Annual Hawaii International Conference on System Sciences* (Vol. 2015-March, pp. 2326–2335). https://doi.org/10.1109/HICSS.2015.280

Okamoto, K. (2017). Introducing Open Government Data. *Reference Librarian*, *58*(2), 111–123. https://doi.org/10.1080/02763877.2016.1199005

Online Etymology Dictionary. (n.d.). Innovation (n.). Retrieved February 3, 2022, from https://www.etymonline.com/word/innovation#etymonline_v_30060

Open Knowledge Belgium, & Noms peut-être! (2020). EqualStreetNames.Brussels. Retrieved May 16, 2022, from https://equalstreetnames.brussels/fr/index.html#11.14/50.8389/4.363

Open Knowledge Foundation. (n.d.). Mission Statement. Retrieved February 3, 2022, from https://okfn.org/about/

Open Knowledge Foundation. (2015). The Open Data Handbook. Retrieved December 9, 2019, from http://opendatahandbook.org/guide/en/

OpenDataSoft. (2021). Visualisez vos données. Retrieved May 16, 2022, from https://www.opendatasoft.com/en/visualisez-vos-donnees/

Osborne, D., & Gaebler, T. (1992). *Reinventing Government*. New York: Plume.

Osborne, S. P. (2006). The New Public Governance? *Public Management Review*, *8*(3), 377–387. https://doi.org/10.1080/14719030600853022

Osborne, S. P. (2018). From public service-dominant logic to public service logic: are public service organizations capable of co-production and value co-creation? *Public Management Review*, *20*(2), 225–231. https://doi.org/10.1080/14719037.2017.1350461

Osborne, S. P., Radnor, Z., & Strokosch, K. (2016). Co-Production and the Co-Creation of Value in Public Services: A suitable case for treatment? *Public Management Review*, *18*(5), 639–653. https://doi.org/10.1080/14719037.2015.1111927

*Oslo Manual, Guidelines for Collecting and Interpreting Innovation Data*. (2005). Paris.

Oxford Dictionary. (n.d.). Actor (noun). Retrieved February 3, 2022, from https://www.oxfordlearnersdictionaries.com/definition/english/actor?q=actor

Pallot, M., Trousse, B., Senach, B., Scapin, D., Pallot, M., Trousse, B., … Scapin, D. (2011). Living Lab Research Landscape : From User Centred Design and User Experience towards User Cocreation To cite this version : HAL Id : inria-00612632 Living Lab Research Landscape : From User Centred Design and User Experience towards User Cocreation. Retrieved from https://hal.inria.fr/inria-00612632

Panagiotopoulos, P., Klievink, B., & Cordella, A. (2019). Public value creation in digital government. *Government Information Quarterly*, *36*(4). https://doi.org/10.1016/j.giq.2019.101421

Papadopoulos, Y. (2013). *Democracy in crisis?: politics, governance and policy*. Macmillan International Higher Education.

Park, S., & Gil-Garcia, J. R. (2021). Open data innovation: Visualizations and process redesign as a way to bridge the transparency-accountability gap. *Government Information Quarterly*, (May 2021), 101456. https://doi.org/10.1016/j.giq.2020.101456

Paulus, D., Meesters, K., & van de Walle, B. (2018). Turning data into action: supporting humanitarian field workers with open data. In *Proceedings of the 15th ISCRAM Conference* (pp. 1030–1040).

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Pereira de Castro Casa Nova, S., Costa Lourenço, I., & Ferreira LeitãoAzevedo, R. (2018). The misalignment between accounting faculty perceptions of success and organizational image during a process of institutional change. *International Journal of Organizational Analysis*, *26*(5), 812–841. https://doi.org/10.1108/IJOA-08-2017-1216

Pereira, G. V., Macadar, M. A., Luciano, E. M., & Testa, M. G. (2017). Delivering public value through open government data initiatives in a Smart City context. *Information Systems Frontiers*, *19*(2), 213–229. https://doi.org/10.1007/s10796-016-9673-7

Piboonrungroj, P., Williams, S. J., & Simatupang, T. M. (2017). The emergence of value chain thinking. *International Journal of Value Chain Management*, *8*(1), 40. https://doi.org/10.1504/ijvcm.2017.10003558

Piller, F. T., Ihl, C., & Vossen, A. (2010). A Typology of Customer Co-Creation in the Innovation Process. *SSRN Electronic Journal*, (December). https://doi.org/10.2139/ssrn.1732127

Porter, M. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: Free Press.

Poulain, H., Goetz, J., & Lapoix, S. (n.d.). Data Gueule. Retrieved May 16, 2022, from https://www.youtube.com/user/datagueule/featured

Publications Office of the European Union. (2012). data.europa.eu. Retrieved February 3, 2022, from https://data.europa.eu/en

Reggi, L., & Dawes, S. (2016). Open Government Data Ecosystems: Linking Transparency for Innovation with Transparency for Participation and Accountability. In *Proceedings of the 18th Annual International Conference on Digital Government Research - dg.o '17* (Vol. Part F1282, pp. 74–86). New York, New York, USA: ACM Press. https://doi.org/10.1007/978-3-319-44421-5_6

Région de Bruxelles Capitale. (2021). *Circulaire précisant les modalités pratiques de publication en open data des inventaires des marchés publics et des subventions visés par les décret et ordonnance conjoints du 16 mai 2019*. CM 2021-12-16/16, art. M, 002. Retrieved from http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=fr&la=F&cn=2021011410&table_name=loi

Roberts, D., Hughes, M., & Kertbo, K. (2014). Exploring consumers' motivations to engage in innovation through co-creation activities. *European Journal of Marketing*, *48*(1/2), 147–169. https://doi.org/10.1108/EJM-12-2010-0637

Rose, Göbel, Cronholm, Holgersson, Söderström, & Hallqvist. (2019). Theory-Based Design Principles for Digital Service Innovation. *E-Service Journal*, *11*(1), 1. https://doi.org/10.2979/eservicej.11.1.01

Ruijer, E., Grimmelikhuijsen, S., & Meijer, A. (2017). Open data for democracy: Developing a theoretical framework for open data use. *Government Information Quarterly*, *34*(1), 45–52. https://doi.org/10.1016/j.giq.2017.01.001

Ruijer, E., & Meijer, A. (2020). Open Government Data as an Innovation Process: Lessons from a Living Lab Experiment. *Public Performance & Management Review*, *43*(3), 613–635. https://doi.org/10.1080/15309576.2019.1568884

Sadiq, S., & Indulska, M. (2017). Open data: Quality over quantity. *International Journal of Information Management*, *37*(3), 150–154. https://doi.org/10.1016/j.ijinfomgt.2017.01.003

Safarov, I. (2019). Institutional Dimensions of Open Government Data Implementation: Evidence from the Netherlands, Sweden, and the UK. *Public Performance and Management Review*, *42*(2), 305–328. https://doi.org/10.1080/15309576.2018.1438296

Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, *22*(1), 1–24. https://doi.org/10.3233/IP-160012

Sanders, E. B.-N. (2002). From User-Centered to Participatory Design Approaches.

Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Sangiambut, S., & Sieber, R. (2017). The civic open data and crowdsourcing app ecosystem: Actors, materials, and interventions. *URISA Journal*, *28*(1), 49–62.

Saunders, M., & Lewis, P. (2018). *Doing Research in Business and Management. An essential guide to planning your project.* (Second edi). Pearson Education Limited.

Scholl, H. J. (2014). The EGOV Research Community: An Update on Where We Stand (pp. 1–16). https://doi.org/10.1007/978-3-662-44426-9_1

Schuler, D., & Namioka, A. (1993). *Participatory Design: Principles and Practices*. New Jersey: CRC Press.

Sein, Henfridsson, Purao, Rossi, & Lindgren. (2011). Action Design Research. *MIS Quarterly*, *35*(1), 37. https://doi.org/10.2307/23043488

Selig, C. J., & Baltes, G. H. (2018). Clarifying the roles in corporate entrepreneurship. *2017 International Conference on Engineering, Technology and Innovation: Engineering, Technology and Innovation Management Beyond 2020: New Challenges, New Approaches, ICE/ITMC 2017 - Proceedings*, *2018-Janua*, 879–887. https://doi.org/10.1109/ICE.2017.8279976

Shenton, A. K. (2007). Viewing information needs through a Johari Window. *Reference Services Review*, *35*(3), 487–496. https://doi.org/10.1108/00907320710774337

Sieber, R. E., & Johnson, P. A. (2015). Civic open data at a crossroads: Dominant models and current challenges. *Government Information Quarterly*, *32*(3), 308–315. https://doi.org/10.1016/j.giq.2015.05.003

Silvis, E., & M. Alexander, P. (2014). A study using a graphical syntax for actor-network theory. *Information Technology & People*, *27*(2), 110–128. https://doi.org/10.1108/ITP-06-2013-0101

Smart Flanders. (2018). Open Data Charter. Retrieved February 3, 2022, from https://smart.flanders.be/open-data-beleid/open-data-charter/

Smith, G., & Sandberg, J. (2018). Barriers to innovating with open government data: Exploring experiences across service phases and user types. *Information Polity*, *23*(3), 249–265. https://doi.org/10.3233/IP-170045

Soh, C., & Sia, S. K. (2004). An institutional perspective on sources of ERP package-organisation misalignments. *Journal of Strategic Information Systems*, *13*(4 SPEC. ISS.), 375–397. https://doi.org/10.1016/j.jsis.2004.11.001

Sørensen, E., & Torfing, J. (2011). Enhancing Collaborative Innovation in the Public Sector. *Administration & Society*, *43*(8), 842–868. https://doi.org/10.1177/0095399711418768

Staron, M., Meding, W., & Baniasad, P. (2019). Information Needs for SAFe Teams and Release Train Management: A Design Science Research Study. In *International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement (IWSM Mensura 2019)* (pp. 55–70). Retrieved from http://ceur-ws.org/Vol-2476/paper4.pdf

Steen, M., Kuijt-Evers, L., & Klok, J. (2007). Early user involvement in research and design projects–A review of methods and practices. In *23rd EGOS Colloquium* (pp. 1–21). https://doi.org/10.1080/15710880701875068

Stoker, G. (2006). Public value management: A new narrative for networked governance? *American Review of Public Administration*, *36*(1), 41–57. https://doi.org/10.1177/0275074005282583

Sunlight Foundation. (2010). Ten Principles for Opening Up Government Information. Retrieved February 3, 2022, from https://sunlightfoundation.com/policy/documents/ten-open-data-principles/

Susha, I., Grönlund, Å., & Janssen, M. (2015). Organizational measures to stimulate user engagement with open data. *Transforming Government: People, Process and Policy*, *9*(2), 181–206. https://doi.org/10.1108/TG-05-2014-0016

Susha, I., Janssen, M., & Verhulst, S. (2017). Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development. In *Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 2691–2700). https://doi.org/10.24251/HICSS.2017.325

Sveriges regering - Infrastrukturdepartementet. (2019). *Sweden's Open Government Partnership Action Plan 2019–2021*. Annex to decision II 3 at Government meeting on 29 August 2019, I2019/02308/DF. Retrieved from https://www.regeringen.se/4ad5d7/contentassets/0e4dc8996f374895b54f9f994e6c5fd2/sweden-ogp-action-plan-2019-2021.pdf

Terwiesch, C., & Xu, Y. (2008). Innovation Contests, Open Innovation, and Multiagent Problem Solving. *Management Science*, *54*(9), 1529–1543. https://doi.org/10.1287/mnsc.1080.0884

The Open Data Charter. (2015). International Open Data Charter's Principles. Retrieved February 3, 2022, from https://opendatacharter.net/principles/

The Open Government Partnership. (n.d.). Policy Area: Right to Information. Retrieved February 3, 2022, from https://www.opengovpartnership.org/policy-area/right-to-information/

The White House. (2009). *Transparency and Open Government. Memorandum for the heads of executive departments and agencies*. Retrieved from https://obamawhitehouse.archives.gov/the-press-office/transparency-and-open-government

The White House. (2013). *Executive Order -- Making Open and Machine Readable the New Default for Government Information*. Retrieved from https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-

The World Bank. (n.d.). World Bank Open Data. Retrieved February 3, 2022, from https://data.worldbank.org/

Timmerman, J. G., Boer, J., Hisschemöller, M., & Mulder, W. (2001). Specifying information needs: improving the working methodology. *Regional Environmental Change*, *2*(2), 77–84. https://doi.org/10.1007/s101130100030

Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory*, *30*(3), 167–186. https://doi.org/10.1177/0735275112457914

Toots, M., McBride, K., Kalvet, T., & Krimmer, R. (2017). Open data as enabler of public service co-creation: Exploring the drivers and barriers. *Proceedings of the 7th International Conference for E-Democracy and Open Government, CeDEM 2017*, (October), 102–112. https://doi.org/10.1109/CeDEM.2017.12

Toots, M., McBride, K., Kalvet, T., Krimmer, R., Tambouris, E., Panopoulou, E., … Tarabanis, K. (2017). A Framework for Data-Driven Public Service Co-production. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10428 LNCS, pp. 264–275). https://doi.org/10.1007/978-3-319-64677-0_22

Toro, J. F., Carrion, D., Albertella, A., & Brovelli, M. A. (2019). Cross-Border Open Data Sharing: Gioconda Project. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(4/W14), 233–238. https://doi.org/10.5194/isprs-archives-XLII-4-W14-233-2019

U.S. Federal Government. (2009). Data.gov. Retrieved February 3, 2022, from https://www.data.gov/

Ubaldi, B. (2013). *Open Government Data. Towards Empirical Analysis of Open Government Data Initiatives*.

United Nations. (1948). Universal Declaration of Human Rights. Retrieved February 2, 2022, from https://www.un.org/en/about-us/universal-declaration-of-human-rights

Urquhart, C. (2001). Bridging information requirements and information needs assessment: do scenarios and vignettes provide a link? *Information Research*, *6*(2). Retrieved from http://informationr.net/ir/6-2/paper102.html

Vaishnavi, V., Kuechler, B., & Petter, S. (2004). Design Science Research in information Systems, (1), 1–66. https://doi.org/1756-0500-5-79 [pii]\r10.1186/1756-0500-5-79

van Hesteren, D., van Knippenberg, L., Weyzen, R., Huyer, E., & Cecconi, G. (2021). *Open Data Maturity Report 2021*. Retrieved from https://data.europa.eu/sites/default/files/landscaping_insight_report_n7_2021_0.pdf

van Knippenberg, L. (2020). *Open Data Maturity Report*. Retrieved from https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n6_2020.pdf

van Loenen, B. (2018). Towards a User-Oriented Open Data Strategy. In *Open Data Exposed, Information Technology and Law* (pp. 33–53). T.M.C. Asser Press. https://doi.org/10.1007/978-94-6265-261-3_3

van Schalkwyk, F., Willmers, M., & McNaughton, M. (2016). Viscous Open Data: The Roles of Intermediaries in an Open Data Ecosystem. *Information Technology for Development*, *22*, 68–83. https://doi.org/10.1080/02681102.2015.1081868

Verganti, R. (2008). Design, Meanings, and Radical Innovation: A Metamodel and a Research Agenda *. *Journal of Product Innovation Management*, *25*(5), 436–456. https://doi.org/10.1111/j.1540-5885.2008.00313.x

Vigoda, E. (2002). From responsiveness to collaboration: Governance, citizens, and the next generation of public administration. *Public Administration Review*, *62*(5), 527–540. https://doi.org/10.1111/1540-6210.00235

Vigoda, E., & Golembiewski, R. T. (2001). Citizenship Behavior and the Spirit of New Managerialism. *The American Review of Public Administration*, *31*(3), 273–295. https://doi.org/10.1177/02750740122064956

Ville de Namur. (n.d.-a). Open Data Namur. Retrieved May 16, 2022, from https://data.namur.be/map/+734a13246fba4515/edit/

Ville de Namur. (n.d.-b). Présentation Open Data. Retrieved February 3, 2022, from https://data.namur.be/pages/presentationopendata/

Virkar, S., & Viale Pereira, G. (2018). Exploring Open Data State-of-the-Art: A Review of the Social, Economic and Political Impacts (pp. 196–207). https://doi.org/10.1007/978-3-319-98690-6_17

Visser, F. S., Stappers, P. J., van der Lugt, R., & Sanders, E. B.-N. (2005). Contextmapping: experiences from practice. *CoDesign*, *1*(2), 119–149. https://doi.org/10.1080/15710880500135987

Vlaams Parlement. (2015). *Décret du 12 juin 2015 modifiant le décret du 27 avril 2007 portant réutilisation des informations du secteur public et modifiant le décret du 18 juillet 2008 relatif à l'échange électronique de données administratives*. Retrieved from http://www.ejustice.just.fgov.be/eli/decret/2015/06/12/2015035784/justel

von Glasersfeld, E. (1988). Introduction à un constructivisme radical. In *L'Invention de la réalité. Contributions au constructivisme* (pp. 19–43). Paris: Seuil.

von Glasersfeld, E. (2001). The Radical Constructivist View of Science. *Foundations of Science*, *6*, 31–43.

von Hippel, E. (1986). Lead Users : A Source of Novel Product Concepts. *Management Science*, *32*(7), 791–805.

von Hippel, E. (1994). "Sticky Information" and the Locus of Problem Solving: Implications for Innovation. *Management Science*, *40*(4), 429–439. https://doi.org/10.1287/mnsc.40.4.429

von Hippel, E. (2001). Innovation by User Communities: Learning from Open-Source Software. *MIT Sloan Management Review*, *42*(4), 82–86.

von Hippel, E. (2005a). Democratizing innovation: The evolving phenomenon of user innovation. *MIT Press,* 63–78. https://doi.org/10.1007/s11301-004-0002-8

von Hippel, E. (2005b). *Democratizing Innovation*. Cambridge: The MIT Press. Retrieved from http://library.oapen.org/handle/20.500.12657/26093

von Hippel, E., & von Krogh, G. (2003). Open Source Software and the " Private-Collective" Innovation Model: Issues for Organization Science. *Organization Science*, *14*(2), 209–224.

Walker, J., Frank, M., & Thompson, N. (2015). User centred methods for measuring the value of open data. In *Open Data Research Symposium* (p. 31). Ottawa, Canada.

Walters, D., & Lancaster, G. (2000). Implementing value strategy through the value chain. *Management Decision*, *38*(3), 160–178. https://doi.org/10.1108/EUM0000000005344

Wang, H. J., & Lo, J. (2016). Adoption of open government data among government agencies. *Government Information Quarterly*, *33*(1), 80–88. https://doi.org/10.1016/j.giq.2015.11.004

Warfield, J. N. (1994). *Science of generic design: managing complexity through systems design*. Iowa State Press.

Warner, E. S., Murray, A. D., & Palmour, V. E. (1973). *Information Needs of Urban Residents*. *U.S. Department of Health, Education, and Welfare Office of Education Bureau of Libraries and Learning Resources*.

Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: an empirical view from the Citizen's perspective. *Information Systems Frontiers*, *19*(2), 285–300. https://doi.org/10.1007/s10796-016-9679-1

Wei, H. L., Wang, E. T. G., & Ju, P. H. (2005). Understanding misalignment and cascading change of ERP implementation: A stage view of process analysis. *European Journal of Information Systems*, *14*(4), 324–334. https://doi.org/10.1057/palgrave.ejis.3000547

Welle Donker, F., & van Loenen, B. (2017). How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, *10*(3), 284–306. https://doi.org/10.1080/17538947.2016.1224938

Westbrook, L. (1993). User Needs: A Synthesis and Analysis of Current Theories for the Practitioner. *RQ*, *32*(4), 541–549. Retrieved from https://www.jstor.org/stable/25829371

Wilson, B., & Chakraborty, A. (2019). Planning Smart(er) Cities: The Promise of Civic Technology. *Journal of Urban Technology*, *26*(4), 29–51. https://doi.org/10.1080/10630732.2019.1631097

Wilson, T. D. (1981). On user studies and information needs. *Journal of Documentation*, *37*(1), 3–15. https://doi.org/10.1108/00220410610714895

Wilson, T. D. (1994). Information needs and uses: fifty years of progress. *Fifty Years of Information Progress: A Journal of Documentation Review*, *28*(1), 15–51.

Wing, J., Andrew, T., & Petkov, D. (2017). The changing nature of user involvement in information system development projects. *2017 Conference on Information Communication Technology and Society, ICTAS 2017 - Proceedings*. https://doi.org/10.1109/ICTAS.2017.7920655

Wirtz, B. W., Piehler, R., Thomas, M.-J., & Daiser, P. (2016). Resistance of Public Personnel to Open Government: A cognitive theory view of implementation barriers towards open government data. *Public Management Review*, *18*(9), 1335–1364. https://doi.org/10.1080/14719037.2015.1103889

Worthy, B. (2015). The impact of open data in the Uk: Complex, unpredictable, and political. *Public Administration*, *93*(3), 788–805. https://doi.org/10.1111/padm.12166

Yoon, A., Copeland, A., & McNally, P. J. (2018). Empowering communities with data: Role of data intermediaries for communities' data utilization. *Proceedings of the Association*

*for Information Science and Technology*, *55*(1), 583–592. https://doi.org/10.1002/pra2.2018.14505501063

Yu, H., & Robinson, D. G. (2012). The New Ambiguity of "Open Government." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2012489

Zamora, E. A. (2016). Value Chain Analysis: A Brief Review. *Asian Journal of Innovation and Policy*, *5*(2), 116–128. https://doi.org/10.7545/ajip.2016.5.2.116

Zhang, Y., Hua, W., & Yuan, S. (2017). Mapping the scientific research on open data: A bibliometric review. *Learned Publishing*, *31*(2), 95–106. https://doi.org/10.1002/leap.1110

Zuiderwijk, A., Helbig, N., Gil-García, J. R., & Janssen, M. (2014). Special issue on innovation through open data - A review of the state-of-the-art and an emerging research agenda: Guest editors' introduction. *Journal of Theoretical and Applied Electronic Commerce Research*, *9*(2). https://doi.org/10.4067/S0718-18762014000200001

Zuiderwijk, A., & Janssen, M. (2014). The negative effects of open government data - Investigating the dark side of open data. In *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2612733.2612761

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., Sheikh_Alibaks, R., & Alibaks, R. S. (2012). Socio-technical impediments of open data. *Electronic Journal of EGovernment*, *10*(2), 156–172. https://doi.org/10.1641/b570402?ref=search-gateway:885882d1830675b0f27af0760faeaef8

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. *Information Polity*, *19*(1–2), 17–33. https://doi.org/10.3233/IP-140329

Zuiderwijk, A., Janssen, M., Poulis, K., & Van De Kaa, G. (2015). Open data for competitive advantage: Insights from open data use by companies. *ACM International Conference Proceeding Series*, *27-30-May-*, 79–88. https://doi.org/10.1145/2757401.2757411

Zuiderwijk, A., Janssen, M., Van De Kaa, G., & Poulis, K. (2016). The wicked problem of commercial value creation in open data ecosystems: Policy guidelines for governments. *Information Polity*, *21*(3), 223–236. https://doi.org/10.3233/IP-160391

Zuiderwijk, A., Shinde, R., & Janssen, M. (2019). Investigating the attainment of open government data objectives: is there a mismatch between objectives and results? *International Review of Administrative Sciences*, *85*(4), 645–672. https://doi.org/10.1177/0020852317739115

Zuiderwijk, A., Susha, I., Charalabidis, Y., Parycek, P., & Janssen, M. (2015). Open data disclosure and use: critical factors from a case study. In P. Parycek & N. Edelmann (Eds.), *CeDEM 2015: Proceedings of the International Conference for E-Democracy and Open Government 2015* (pp. 197–208). Krems: Edition Donau-Universität Krems.

Zuiderwijk, A., Volten, C., Kroesen, M., & Gill, M. (2018). Motivation perspectives on opening up municipality data: Does municipality size matter? *Information (Switzerland)*, *9*(11), 1–26. https://doi.org/10.3390/info9110267

# Appendices

# Appendix I – Study 1

## Appendix I. A. List of selected and analysed initiatives

| **Initiative Code, Sector:**<br>CS: Civil society<br>B: Business<br>P: Public sector<br>(B) Belgium<br>(F) France<br>And Code Name (for anonymization) | **Nature of the Initiative** and description | **Theme** and Expected Outcome (social value) | **Approach** (if several, by perceived importance after coding) | Number of Coded Documents per type of Sources (IS: initiative sources, TS: third source) |
|---|---|---|---|---|
| **1/B (B)**<br>Local Information App | **Mobile application** delivering local information to the citizen | **Accessibility to local services:** Up-to-date local information for the citizens and development of the local economy | **Service Innovation**<br>Civic innovation | PS: 9<br>TS: 4 |
| **2/C (B)**<br>Civic Air Sensors 1 | **Air quality sensors in kits and visualisation tool** developed by the open data community, active in several open data and web projects. | **Environment :** Improve the air quality and challenge the politics | **Open Data Innovation**<br>Civic innovation | PS: 2 |
| **3/B (B)**<br>Public Transport Route Planner | **Mobile application** for smart mobility aggregating the real-time data of the main public transport companies in Belgium | **Mobility:** Encourage the use of public transport and reduce the traffic congestion in cities | **Service Innovation** | PS: 3<br>TS: 3 |
| **4/CS (B)**<br>Open Mobility Dev Tools | **Open source tools** developed as a community project to ease the development of mobility apps | **Mobility:** to empower anyone to build any kind of route planner, without having to invest in data integration, so that more Intermodal route planning solutions can be and used to convince people to opt for more sustainable ways of transport. | **Open Data Innovation**<br>Civic innovation | PS: 4<br>TS: 2 |
| **5/B (B)**<br>Multi Transport planner | **Mobile app** for smart mobility aggregating the information of mobility solutions in one place | **Mobility:** Encourage the use of alternative mobility solutions and reduce the traffic congestion in cities | **Service Innovation** | PS: 1<br>TS: 2 |

| | | | | |
|---|---|---|---|---|
| **7/CS (B)** Bike Commuters app | **Mobile app used as personal sensor and dashboard** to encourage commuting in bicycle (for workers and companies) | **Mobility/environment:** Encourage the use of bicycles to commute to work and reduce the traffic congestion in cities and air pollution | **Civic innovation** Service Innovation | PS: 1 TS: 2 |
| **8/CS (B)** Traffic Counter | **Home sensors and visualisation dashboard** to count traffic in the streets | **Governance in a public issue:** Put citizens in control of local mobility policy and encourage them to change their mobility habits | **Civic innovation** | PS: 1 TS: 5 |
| **9/B (B)** Pupils on Bike app | **Mobile app used as personal sensor and dashboard** to encourage commuting by bicycle (for students) | **Mobility/environment:** Encourage the use of bicycle to commute to school and reduce the traffic congestion in cities and air pollution | **Service Innovation** Civic innovation | PS: 3 TS: 1 |
| **11/CS (B)** Bicycle Parking app | **Web map** to localize bike parking spots location and availabilities | **Mobility/environment:** Encourage the use of bicycles by helping the riders to find parking places | **Service Innovation** Open Data innovation | PS: 4 TS: 2 |
| **15/CS (F)** Civic Data-driven Governance | **Participatory methodology** to identify and encourage personal data sharing reused to empower citizens in the governance of public issues | **Governance in a public issue:** Exploit the potential of the personal data for the governance of territories | **Open Data Innovation** Civic innovation | PS: 5 TS: 1 |
| **20/B (B)** Local Consumption app 1 | **Mobile app** to promote responsible consumption | **Responsible consumption:** Encourage the consumption of local products and local economic development | **Service Innovation** Civic innovation | PS: 3 TS: 2 |
| **23/B (F)** Participatory Urban Planning app 1 | **Mobile app and web platform** to report empty cells and vote on the allocation of commercial spaces | **Governance in urban planning:** Encourage participation in the allocation of commercial spaces and local economic development | **Civic innovation** | PS: 1 TS: 2 |
| **28/B (F)** Garage Locator | **Web platform** to localize, find and rate garage and auto repair services | **Accessibility to local services:** Help the community members to have access to local services | **Service Innovation** | PS: 1 TS: 1 |
| **31/B (F)** **Local** Consumption app 2 | **Mobile app** to connect local producers and consumers | **Responsible consumption:** Encourage the consumption of local products (environmental friendly) and local economic development | **Service Innovation** Civic innovation | PS: 2 |

| | | | | |
|---|---|---|---|---|
| **38/B (F)** Recycling Together app | **Mobile app** to connect waste owners and available drivers | **Cleanliness/environment:** Encourage the recycling of waste (environmental issue) | **Service Innovation** Civic innovation | PS: 4 TS: 5 |
| **39/P (F)** Handicap-friendly Map 1 | **Web map** to help disabled people to navigate in the city | **Navigation of disabled persons:** Improve the navigation experience of the disabled citizens | **Service Innovation** | PS: 1 TS: 1 |
| **40/P (F)** Touristic Map | Touristic collaborative **web map** | **Promotion of the territory:** Improve the attractiveness of a territory and the engagement of the citizens in its promotion | **Service Innovation** Open Data innovation | PS: 1 TS: 1 |
| **41/B (F)** Handicap-friendly Map 2 | **Web map** to help disabled people to navigate in the city and report obstacles | **Navigation of disabled persons:** Improve the navigation experience of the disabled citizens | **Service Innovation** Open Data Innovation | PS: 1 TS: 1 |
| **43/P (F)** Local Services Map | **Web map** to find local services | **Accessibility to local services:** Improve the well-being and living experience of the citizens by the accessibility to local services | **Service Innovation** Open Data innovation | PS: 1 TS: 1 |
| **44/CS (F)** Environmental Monitoring app | **Mobile app and web platform** to report damages to the environment and positive initiatives | **Environment:** Solve environmental issues and improve citizen involvement in it | **Civic innovation** | PS: 2 TS: 1 |
| **45/P (B)** Street Damages Reporting app 1 | **Mobile app** to report damages to the public space | **Improvement of the public space:** Improve the quality of the public space and public service | **Service Innovation** Civic innovation Open Data Innovation | PS: 1 TS: 2 |
| **46/CS (B)** Small Paths app | **Mobile app and maps** to encourage pedestrians to go on small paths | **Navigation of pedestrians:** Improve the safety and experience of the pedestrians by promoting forgotten and hidden paths | **Service Innovation** Open Data innovation | PS: 2 TS: 2 |
| **47/CS (B)** Civic Air Sensors 2 | **Home sensors** to measure air quality | **Environment:** Improve the air quality and challenge the politics | **Civic innovation** | PS: 2 TS: 1 |
| **48/B (F)** Talking with the City app | **Mobile app** to enable communication between citizens and municipalities to report problems, provide ideas or congratulate initiatives | **Governance in urban planning:** Improve the communication between citizens and municipalities, the public services and the well-being in the cities | **Civic Innovation** | PS: 2 TS: 2 |

| | | | | |
|---|---|---|---|---|
| **49/CS (F)** Recycling centres Locator | **Web map** to find recycling centres | **Cleanliness/environment:** Encourage the recycling of waste (environmental issue) | **Service Innovation** Civic Innovation | PS: 1 TS: 1 |
| **50/B (F)** Participatory urban planning app 2 | **Web app** to inform citizens about urban changes and planning, and enable them to participate in the decision process | **Governance in urban planning:** Enable citizens to participate in the governance of the territory | **Civic innovation** | PS: 1 TS: 1 |

## Appendix I. B. Coding scheme of the inductive approach (conventional content analysis)

| Code | Definition |
|---|---|
| Initiative's purpose | Mission, objectives and expected outcome |
| Locus of the innovation | Main value of the innovation |
| Open data use | Rationale of using data:<br>Objectives and reasons for using open data<br>Source of the data |
| Citizen participation | Rationale of citizen participation:<br>Objectives and underlying logic of the participation<br>(motivation and expected output of the participation) |

## Appendix I. C. Coding scheme of the deductive approach (based on the conceptual framework and previous literature): role of the citizen

| Code (category of role) | Definition and example of related activities | Reference |
|---|---|---|
| *Citizen Data provider* | Consciously generates data that is openly available for use in the public domain. | (Meijer & Potjer, 2018) |
| *Supplier* | Allows others to use and re-use Open Data | (Berends et al., 2017) |
| *Aggregator* | Collects and aggregates Open Data and, sometimes, other proprietary data and finds correlations, identifies efficiencies, and visualizes complex relationships | (Berends et al., 2017) |
| *Developer* | Designs, builds, and[/or] sells web-based, tablet, smartphone applications | (Berends et al., 2017) |
| *Enricher* | Uses Open Data to enhance their existing products and services through better insight (typically larger companies)<br>Enabler | (Berends et al., 2017) |
| *Enabler* | Facilitates the supply of Open Data or use of Open Data | (Berends et al., 2017) |
| *Idea Generator* | Expresses ideas based on needs | (Cooper, 2018)<br>(Sanders & Stappers, 2008) |
| *Co-creator* | Participates in the development and co-creation of solutions | (Cand ooper, 2018)<br>(Sanders & Stappers, 2008) |
| *Tester* | Participates in testing and validation phases of solution | (Cooper, 2018) |

## Appendix I. D. Semi-structured Interview guide

- *Does you project use open data?*
- *Tell me about your project in a few words. What is your purpose?*
- *How and why do you use open data?*
- *How is the citizen involved, and for what purpose?*
- *What are your current challenges?*
- *What are the future developments of your project?*

## Appendix I. E. List of initiatives interviewed sorted by innovation approach

| Approach | Initiative code and code name | Motivation for the interview | Respondent's function in the initiative | Interview duration |
|---|---|---|---|---|
| **Open data innovation** | 2/CS Civic Air Sensors 1 | Example of open data enthusiasts' movement | Local Chapter Coordinator | 55 minutes |
| **Service Innovation** Civic Innovation | 1/B Local Information app | Example of digitalization of public services (driven by a private actor). Leader in open data reuse in Wallonia | Founder and CEO of the start-up | 48 minutes |
| **Service Innovation** | 5/B Multi Transport Planner | Example of digital entrepreneur (new service for the public interest) | Head Of Business Development | 25 minutes |
| **Service Innovation** | 36/B Street Damages Reporting app 2 | Example of digitalization of public services (driven by a private actor). *The interviewed revealed that the initiative did not use OD and was removed of the final selection.* | Product Manager | 28 minutes |
| **Service Innovation** Civic innovation | 20/B Local Consumption app 1 | Example of digital social entrepreneurship (for profit and social value oriented) | Co-Founder & Project Manager | 48 minutes |
| **Service Innovation** Civic Innovation | 38/B Recycling Together app | Example of digital social entrepreneurship (for profit and social value oriented) | Entrepreneur/ developer | 22 minutes |

| | | | | |
|---|---|---|---|---|
| **Service Innovation** | 41/CS Handicap-friendly Map 2 | Example of civic hacker project (hobbyist) | Entrepreneur/ developer | 28 minutes |
| **Service Innovation** Open Data innovation | 46/CS Small Paths app | Example of CGD initiative to support the purpose of an existing association | Communication and Project Manager | 25 minutes |
| **Civic Innovation** Service innovation | 7/CS Bike Commuters app | Example of open data reuse to support the purpose of an existing organisation (to promote a mobility switch) | Project Manager | 37 minutes |
| **Civic innovation** | 8/CS Traffic Counter | Example of CGD initiative to support the purpose of an existing association (to enable participation in policy making) | Research and development team leader | 56 minutes |
| **Civic innovation** | 47/CS Civic Air Sensors 2 | Example of CGD initiative to support the purpose of an existing activist movement (air quality) | Project leader and activist | 43 minutes |
| **Civic innovation** | 44/CS Environmental monitoring app | Example of CGD initiative to support the purpose of an existing organisation (environmental issues) | Project Coordinator | 20 minutes |

## Appendix II – Study 3

### Appendix II. A. Design of Iteration 1

| | |
|---|---|
| **Characteristics** | Face-to-face workshop |
| **Duration** | 1:30 hours, including evaluation of the instantiation with the participants. |
| **Participants' role** | Users with information needs. |
| **Participants' profiles** | 9 Researchers as inhabitants or commuters, and users of services in the city of Namur (Belgium). Their fields are Information Management (1 Participant (Ps)), Computer Science (2 Ps), Business Informatics (1 Ps), Education (1 Ps), Management and Creativity (1 Ps), Management (1 Ps), Education and Management (1 Ps), and Media Industry (1 Ps). |
| **Justification of profiles** | The participants belonged to a homogeneous group of users and perceived to share similar needs regarding information related to the city they work in. Being researchers, they could also be critical and knowledgeable participants contributing to the improvement of the method at the early development stage. |
| **Recruitment method** | Direct contact through e-mail. |
| **Vignette** | *"You have been accepted to your dream job and moved into an apartment in the big city. You find yourself in a new environment without any prior knowledge about it. What would you do to settle in? What would you like to know?"* |
| **Question scheme** | (1) Identify a task/need related to your role in the given situation (what do I want to do?), (2) What information do I need? (3) Where could I get the information? and (4) A solution: shortly describe it. What information would it provide? |

## Appendix II. B. Evaluation of iteration 1 at the instantiation level

| | |
|---|---|
| **Purpose** | Evaluate the mechanisms used against their objectives as perceived by the participants, with a scientific overlook from our peers. |
| **Method** | Open questions and discussions at the end of the face-to-face workshop. Data was collected with field notes written by the researchers. |
| **Respondents** | Participants of the workshop. |
| **Question scheme** | (1) What are your thoughts about the wording of the questions? (2) What are your thoughts about the sequences of work, and work dynamics? (3) What are your thoughts about the use of scenario? (4) What went well that we should keep for a next workshop? and (5) What was difficult and could be improved for a next workshop? |

## Appendix II. C. Design of iteration 2

| | |
|---|---|
| **Characteristics** | Face-to-face workshop |
| **Duration** | 1:30 hours, including evaluation of the instantiation with the participants. |
| **Participants' role** | Users with information needs. |
| **Participants' profiles** | 11 students, as inhabitants or commuters and users of services in the city of Namur (Belgium). Students from Management (3 Ps), Informatics (1 Ps), Political sciences (1 Ps), Project Management (1 Ps), Analytics and Digital business (4 Ps), and unspecified (1 Ps). |
| **Justification of profiles** | The students belonged to a homogeneous group of users and perceived to share similar needs regarding information related to the city they study in. |
| **Recruitment method** | Posters in the student areas, communication in the classrooms, and direct recruitment (presentation of the research project in direct interaction, in the halls and student rooms, during the lunch break). |
| **Vignette** | *"You have been accepted to your dream university and moved into an apartment in the big city. You find yourself in a new environment without any prior knowledge about it. What would you do to settle in? What would you like to know?"*<br>• In addition, the groups selected one activity domain to focus on: (a) socialize and connect with people, (b) explore and know the city, (c) eat and drink, and (d) be involved in the life of the city. |
| **Question scheme** | (1) Identify a task/need related to your role in the given situation (what do I want to do?), (2) What information will I need? (3) Existing solutions: what information needs can it satisfy? What challenges/problem can you see with it? and (4) Imagined solution: shortly describe it. What information will it satisfy? What problems will it solve? |

## Appendix II. D. Evaluation of iteration 2 at the instantiation level

| | |
|---|---|
| **Purpose** | Evaluate the mechanisms used against theirs objectives as perceived by the participants. |
| **Method** | Open questions and discussions at the end of the face-to-face workshop. Data was collected with field notes written by the researchers. |
| **Respondents** | Participants of the workshop. |
| **Question scheme** | (1) What are your thoughts about the wording of the questions? (2) What are your thoughts about the sequences of work, and work dynamics? (3) Were the objectives of the workshop clear? (4) What went well that we should keep for a next workshop? and (5) What was difficult and could be improved for a next workshop? |

## Appendix II. E. Evaluation of iteration 2 at the artefact level

| | |
|---|---|
| **Purpose** | Evaluate the relevance, perceived reusability (accessibility, actionability), and interest for the method under development given the produced insights and given the facilitation script. |
| **Method** | A popular science article presenting the method and examples of activities from the instantiations (a facilitation script with extensive description of steps, tasks, and roles to conduct the method without prior knowledge) was sent to the respondents a few days in advance. It was followed by a short presentation of the workshop and its objectives, the produced insights, and an interview with the practitioners. The data was recorded as field notes written by the researcher conducting the interviews (lasting 1:00 to 2:00 hours). |
| **Respondents**<br>• **Intermediaries**<br>• **Future Data Providers:** | Potential implementers of the method and beneficiaries of its insights (audience of the artefact), i.e. intermediaries and future data providers.<br><br>The CEO, a UX Designer, and a front-end developer from a start-up. The start-up is a pioneer in Wallonia (Belgium) in using OGD. The company has developed an application to provide local information to the citizens (relevant field knowledge) and works almost exclusively with OGD since 7 years. They were contacted through direct e-mail.<br><br>4 representatives of 2 municipalities (2 per site). A local deputy and a civil servant responsible for IT systems and OGD (per site). The municipalities are planning to release OGD. The artefact was presented as way to stimulate reflection and help them prioritize datasets with users involvement (citizens). They were contacted through direct e-mail. |
| **Question scheme**<br>• **Intermediaries**<br>• **Future Data Providers:** | Do the presented insights, produced by the workshop, help you to grasp the users' information needs in the given context? (2) What could you do with these insights, in your job? Are they useful? (3) What is missing? What would you like to know more? Do the insights need to be processed? and (4) Do they help you to envision new reuses and identify potential OGD?<br><br>(1) Do the presented insights, produced by the workshop, help you to foreseen potential data to release? (2) Does it look useful and relevant to you, as future data providers? (3) Given the material (facilitation script), would you feel ready to implement the workshop? and (4) If not, what is holding you back? |

## Appendix II. F. Design of iteration 3

| | |
|---|---|
| **Characteristics** | Virtual workshops. |
| **Duration** | 2:00 hours. Repeated 3 times with three groups (not including the Evaluation, as a post-questionnaire). |
| **Participants' role** | Future data providers; potential implementer of the method (audience of the artefact). |
| **Participants' profiles** | Civil servants of 4 Belgian public agencies and 12 municipalities (28 in total, split into three groups of 8, 15 and 7 participants). |
| **Justification of profiles** | Municipalities and public agencies are planning to release OGD. The artefact was implemented as way to start to stimulate reflection and help them prioritize datasets and projects of reuses with user involvement. |
| **Recruitment method** | The workshop was the initial session of a training program, aiming at coaching the municipalities and in the release of OGD, organised by a public agency. The researchers had full responsibility and autonomy regarding the objectives, content, and facilitation of the workshop. |
| **Vignette** | Cases were brought by the participants. |
| **Question scheme** | (1) Activity: what do my persona in the situation (problem)? What does he/she try to solve in first place? (2) Information needs: what information does he/she need to perform that activity or task? and (3) Potential difficulties: related to the use of an existing solution, to a context, and to organisational constraints. |

## Appendix II. G. Evaluation of iteration 3 at the instantiation level

| | |
|---|---|
| **Purpose** | Compare the mechanisms used in the instantiation against their objectives, as perceived by the participants. |
| **Method** | Online forms with Linkert scales and open questions sent by e-mail to the participants a few days after the virtual workshop. |
| **Respondents** | Civil servants as future data providers (16 out of the 28 participants responded). |
| **Question scheme** | **Structured question scheme:**<br>The question scheme used Likert scales (from 1 "not agree at all" to 10 "totally agree"). The participants were asked to rate their agreement of the following statements; (1) Thanks to the plenary presentations, documents, and wording of questions, I understood the main concepts: (open) data, (need for) information, solutions reusing open data; (2) Through plenary presentations, documents, and wording of questions, I understood what was expected of me; (3) Through the plenary presentations, preparatory activities, and wording of questions, I felt prepared and able to participate in the activities; (4) I felt included during the webinar (room to interact, easy to give ideas and speak among other participants); (5) Preparatory activities and webinar questions helped me frame and position my project (taking into account the context, problem, user role and needs, existing solutions); (6) The webinar questions helped me connect user needs, resulting information needs and my project idea; (7) The written productions (preparatory activities and brainstorming) of this webinar were useful in developing my project idea.<br><br>**Open question scheme:**<br>(1) TOPS - What went well that we should keep for a next webinar? and (2) FLOPS - What was one difficulty that could be improved for a next webinar? (Regarding the method, the tools, or unmet expectations) |

## Appendix II. H. Evaluation of iteration 3 at the artefact level

| | |
|---|---|
| **Purpose** | Evaluate the relevance, perceived reusability, and interest for the final design principles (called "guidelines" for the practitioners) based on the evaluation framework of Iivari et al. (2020). |
| **Method** | Popular science article presenting the design principles and example of activities from the instantiations, together with a questionnaire (online form) sent to experienced providers. They were offered to either fill the form at their own pace or suggest us a time at their convenience for an interview, based on the given questionnaire. |
| **Respondents** | Potential implementers of the method (audience of the artefact). Sent to 6 experienced practitioners, 3 responded, 2 accepted an interview (1:00 long), 1 answered the questionnaire. |
| **Profiles** | 3 experienced providers, from the municipality of Gent (Belgium), the region of Västra Götaland (Sweden) and the national traffic agency (Sweden). They were civil servants and open data managers. |
| **Justification** | A municipality, region, and public agency with a long experience with releasing data. They were interviewed to complement the perspectives of the future providers. |
| **Recruitment method** | Direct contact through e-mail. |
| **Question scheme** | **Structured question scheme:**<br>Questions based on the evaluation framework of Iivari et al. (2020) using a 5-grade Linkert scale (from "Totally disagree" to "Totally agree"). The participants were asked to rate their agreement of the following statements; (1) the guidelines make sense to me, (2) the guidelines are easy for me to comprehend, (3) in my view, the guidelines address a real problem in my practice, (4) I find that the guidelines convey new ideas to me, (5) I find the guidelines insightful to my own practice, (6) I think that the guidelines can realistically be carried out in practice, (7) I think that the guidelines can easily be carried out in practice, (8) I think that the guidelines provide sufficient guidance for designing methods, (9) I think that the guidelines provide me with sufficient freedom when designing methods, (10) I find the guidelines useful, (11) compared to the current situation, I believe that the guidelines would enhance the understanding of users' needs, (12) compared to the current situation, I believe that the guidelines would enhance the release of useful datasets, and (13) compared to the current situation, I believe that the guidelines would enhance the development of useful data-driven solutions.<br><br>**Open question scheme:**<br>Other feedback, comments, or insights for future development. |