



UNIVERSITE CATHOLIQUE DE LOUVAIN

LOUVAIN SCHOOL OF MANAGEMENT

The Belgian open data ecosystem and innovation through open data

Supervisor at LSM: Pr Paul Belleflamme

Research Master's Thesis

Submitted by Florian Barthélemy

With a view of getting the degrees

Master in Business Engineering

ACADEMIC YEAR 2015-2016

Acknowledgments

This master thesis, linked to an internship inside PwC Belgium, was written under the direction of Nikolaos Loutas (internship supervisor) and Pr. Paul Belleflamme (academic supervisor).

I would like to thank my internship supervisor, Nikolaos Loutas, for his experience and wise advices, my academic supervisor, Pr. Paul Belleflamme, for his confidence throughout my research and the structure both brought me in my work.

Finally, I would like to thank my parents for the last five years leading to a fresh start in the work life and my girlfriend for her patience and support.

Disclaimer

The views expressed in this document are personal to the authors and do not represent the position of PwC Belgium which takes no responsibility for the content or the opinions shared in the following work.

Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by PwC Belgium.

Contents

Contents	V
Introduction.....	1
Why a study about the Belgian open data ecosystem?	3
1 Setting the context	5
1.1 What is open data?	5
1.2 Why open data?.....	5
1.3 What is an open data ecosystem?	6
2 Methodology.....	10
2.1 Surveys	10
2.2 Semi-structured interviews	11
3 The Belgian open data ecosystem	12
3.1 Actors of the ecosystem	12
3.1.1 Open data initiators, open data providers and open data public agencies.....	16
3.1.2 Open data enablers or infomediaries	19
3.1.3 Open data re-users.....	20
3.2 Interactions in the Belgian open data ecosystem.....	26
4 Innovation and open data in Belgium	31
5 Findings and recommendations	35
5.1 Benefits of open data in Belgium	35
5.2 Organising the Belgian open data ecosystem and supporting open data re-uses.....	38
6 Conclusion.....	46
6.1 Awareness.....	46
6.2 Recommendations	48
6.3 Limits and critical perspective	49
6.4 Guidelines for further research.....	50
7 Tables of figures and tables	52
8 References	53

Introduction

In the context of digital transformation, the public administrations around the world realise the need for adaptation and modernisation of their services in order to meet the changing and ever increasing demands of the whole society. From e-government to open government, the challenges are evolving in a fast pace (Katsonis & Botros, 2015). Nowadays, the economic context, the complexity of our societies, the increased personalisation of products and services and the transparency controversy are all contributing to the need for evolution and adaptation of public services (Arrassi et al., 2016). Open government and open data in a wider perspective are supposed to give an answer to these challenges. The open government data global movement started in the United States in 2008 followed shortly by the United Kingdom but has today expanded far beyond those two countries. Current open government data initiatives can be separated in two main schools of thought: the first one, from an economic perspective, seeking for re-uses of data to innovate and generate economic benefits, while the other one is more open government oriented, approaching things from a social perspective and focusing on creating social value and improving government transparency (Jetzek, Avital, & Bjorn-Andersen, 2014; Lassinantti, Bergvall-Kåreborn, & Ståhlbröst, 2014).

The rapid development of the literature about open data and open government is relatively new (Sandoval, 2011) but has focused on specific topics until recently. First, numerous papers were dedicated to the terminology. Second, most of the literature about open data established what are the main drivers, barriers and challenges faced by governments to open their data: privacy (John Carlo Bertot, Gorham, Jaeger, Sarin, & Choi, 2014; Choenni, van Dijk, & Leeuw, 2010; Meijer, Conradie, & Choenni, 2014; Rohunen, Markkula, Heikkila, & Heikkila, 2014; Scassa, 2014; Anneke Zuiderwijk et al., 2012), transparency (John C. Bertot, Jaeger, & Grimes, 2010; Dawes & Helbig, 2010; Jaeger & Bertot, 2010; Meijer et al., 2014; Anneke Zuiderwijk, Gascó, Parycek, & Janssen, 2014), collaboration and participation of different actors into governance (Harrison, Pardo, & Cook, 2012; M. Janssen, Charalabidis, & Zuiderwijk, 2012; Parycek, Höchtel, & Ginner, 2014), and economic benefits (Carrara, Fischer, Oudkerk, van Steenberg, & Tinholt, 2015; Commission, 2015a; Manyika et al., 2013; Verhulst & Young, 2016). About the last point, economic benefits, two remarks could be outlined. If economic value was pointed out as one of the main drivers since the early days of the open data movement, it has only been in recent years that articles started to tackle this as their main topic. One of the reasons is the lack of metrics and models to measure precisely the value, either economic or social, behind open data (Harrison et al., 2012; Verhulst & Young, 2016). Consequently

to this and to other reasons, the literature restrained itself mostly to the study of how the public sector could embrace the open data ‘movement’. The open data re-users’ side is crucial because re-users represent the actors who are adding value to the data re-used.

Next to metrics and measurement models, the relation between innovation and open data or the types of open data business models existing are sectors for which a deeper understanding is required (Carrara, Fischer, Oudkerk, et al., 2015; M. Janssen & Zuiderwijk, 2014; Anneke Zuiderwijk, Helbig, Gil-García, & Janssen, 2014). Until now, many papers describing open data innovation have taken the perspective of case studies (Commission, 2015a; Ferro & Osella, 2013; Verhulst & Young, 2016). The setting up of the relation between open data and innovation through modelling has not started yet even if a few attempts proposing some factors and recommendations to construct a business using open data already exist (Verhulst & Young, 2016).

While talking about open data, we must not forget the interactions between actors inherent to any open data ecosystem (Harrison et al., 2012). More and more, the concept of ecosystem is put forward to explain the connections between the open data actors. Currently, the concept of open data ecosystem and its role need to be more characterised with concrete examples (Harrison et al., 2012).

In relation with ecosystems, we have to increase our knowledge of the different functionalities open data infrastructures should offer in order to optimise their support capabilities to open data re-uses (Davies, 2011; Anneke Zuiderwijk et al., 2014). Finally, the literature has already assessed the lacks of interoperability, structure of the data and data quality itself, leading to less open data re-used and less business opportunities (Petychakis, Vasileiou, Georgis, Mouzakis, & Psarras, 2013; Anneke Zuiderwijk et al., 2014).

Our objective in this study is triple:

O_1	First, to develop awareness and basic knowledge around the concept of open data ecosystem and how this ecosystem is applied to the Belgian case; and around the factors enhancing innovation through open data
O_2	Second, to be able to give concrete recommendations on how to support the development of the Belgian open data ecosystem and innovation in Belgium for all

	types of open data actors in Belgium
O_3	Third, to outline guidelines for further researches on open data innovation.

Why a study about the Belgian open data ecosystem?

From the current advancement of the literature, we have observed that different domains require a deeper comprehension; especially concerning the role of an ecosystem regarding open data and open government and the relation existing between innovation and open data. Those two elements need a wider approach putting the re-users at the focal point. Moreover, open data studies focusing for example on value measurement are mostly adopting a large scope, e.g. global or European level, or analysing large scale initiatives like early adopters as the UK and the US.

Therefore, it could be interesting to know how open data is applied and developed in smaller countries and regions since many governments who have started open data initiatives or are willing to start should be considered at a different scale from early adopters. In this way, Belgium fulfils certain criteria due to its context:

- A small country with a high density of population and bigger percentage of employees working in high technology sector than the European average (Eurostat, 2015),
- A federal state with three languages and three regions,
- An open data maturity characterised by an early open data activity in Europe but some steps behind the neighbouring countries today (Commission, 2015a, 2015b; K. Janssen, 2011) even if Belgium is catching up according to the two last rankings (Colpaert & Vanagt, 2016; Foundation & development, 2013; Knowledge, 2014).

Second, some Belgian open data actors have just established the first connections of a concrete ecosystem in 2016. As a consequence, it would be relevant to analyse the different challenges, motivations and impacts the construction of an ecosystem has in the specific case of Belgium. It could represent the first step towards a standardised model of open data ecosystem and an example for other initiatives to follow.

In this context, we first take stock of the current situation of the Belgian open data ecosystem with a focus on open data re-use. Then we study more deeply the relation between innovation and open

data. Our methodology focus during a first phase on the literature existing around different key concepts like 'open data', 'open data ecosystem', 'open data innovation' etc. in order to build up a good overview of the current state of open data in Belgium and internationally from the literature's perspective (part 1). In a second phase we conduct surveys for which we first identify the potential re-users of the Belgian open data ecosystem and second, we ask them to answer a questionnaire on their use of open data and their profiles. The third and last phase of our approach aims at interviewing different Belgian open data experts in different groups of actors from a typical open data ecosystem.

Thanks to this methodology (part 2), we elaborate multiple observations and questions for further researches in this report. We synthesise them into a current picture of the Belgian open data ecosystem as detailed as possible (part 3). In part 4, we also describe the different observations we make about the relation between open data and innovation. We propose different factors enabling innovation at the end of this part with examples from Belgian open data re-users and actors. In the fifth and last part, we develop the findings and consequently the recommendations we could conclude from our entire research.

During our study, a logical difficulty was to identify open data re-users who, by definition of open data, are not always specifying that they are using open data. Another one consisted in open data being a young 'movement' with all the consequences resulting from that: the lack of knowledge, the insufficient perspective and experience in the domain, etc.

Thanks to this study and our methodology, we hope first to raise awareness about open data at multiple levels of the society and multiple levels of technical background. The intended audience is mostly constituted of Belgians already conscious or working with open data. The objective is double: by levelling upward the knowledge of different (potential) actors, we contribute to the development of a denser ecosystem with more actors and more connections between them. Besides that, through our description of the Belgian open data ecosystem and the relation between innovation and open data, we make a step towards achieving a European analysis of the ecosystems in place and really evaluating quantitatively the relation innovation-open data and the value of open data.

We provide in the next part the definitions and explanations of the key concepts that we are using during the report.

1 Setting the context

1.1 What is open data?

The definition of **open data** which is used for this report is:

“Information that can be freely used, modified, and shared by anyone for any purpose. It must be available under an open licence and provided in a convenient and modifiable form that is machine readable.” (Commission, 2015a, p. 7).

Examples of data are numerous: nation statistics, law cases and legal procedures, governments' spending, political results, tax evasion, energy market, environment quality etc. (Archer et al., 2014). Open data has to be distinguished from **government data** which are data or information released by the public sector (Lindman, 2014). Open data can thus be non-governmental even if currently most of open data comes from the public sector.

1.2 Why open data?

In this section, we make an attempt to answer the question why open (government) data is seen as a leverage to achieve: more **transparency** in the decision-making, increased **collaboration** from citizens and businesses in the government policies, and **innovation** leading to public and private value (Anneke Zuiderwijk, Janssen, & Davis, 2014).

From the broadest perspective, the “value of open data rests on whether or not it enables us to solve problems and meet important needs of individuals, communities, or society writ large.” (Harrison et al., 2012, p. 912). A recent study describes 19 open data re-use cases where concrete problems were addressed (Verhulst & Young, 2016). For example, cases where government transparency was tackled: the ‘Open Budget Transparency Portal’ in Brazil and the ‘Open Contracting Projects’ in Slovakia. While the first concerns fiscal transparency with “an average of 900,000 unique visitors each month” (Verhulst & Young, 2016, p. 10), the second forces public administrations to publish online all documents linked to the process of public contracts. A third example, ‘OS OpenData’ in the UK provides support to all kind of businesses in the UK which are working with maps. Since the agency is financially independent, it uses a kind of freemium model with data available for free to everybody and other chargeable. The contribution to the GDP was estimated to an increase between 13 and 28.5 million pounds in the first five years (Verhulst & Young, 2016).

In the case of **transparency**, open data can have huge impacts on diminishing corruption or improving the comprehension of citizens towards government decisions as we have just described in the last examples. Transparency value for the society is lying in the reduction of information asymmetry between public administrations and citizens or businesses (Jetzek et al., 2014). However, information asymmetry is not coming naturally from the release of the data alone as highlighted by Thorhildur Jetzek, Michel Avital and Niels Bjorn-Andersen, but from the re-uses made on top of these available data.

From an economic perspective, “the [economic] benefits of open data lie in the creation and delivery of new products and services.”(Carrara, Fischer, Oudkerk, et al., 2015, p. 3). Different revenue models have emerged around open data re-uses like consultancy services, advertising, subscription models, licensing and lead generation (Lindman, 2014). The repartition of respondents to our survey between the different revenue models is given in the next part. What can already be said is that businesses re-using open data are able to develop a revenue model at different stages of the data chain: from producing and collecting data to enabling re-uses and products/services developments on top of data. More examples are given further in the next parts (Carrara, Fischer, Oudkerk, et al., 2015).

The European Commission estimated also the cost savings for the public sector at 0.22% of all expected expenditures in 2020 or in absolute number, total cost savings of about €1.7 billion in 2020 for EU28+ members thanks to open data (Commission, 2015a).

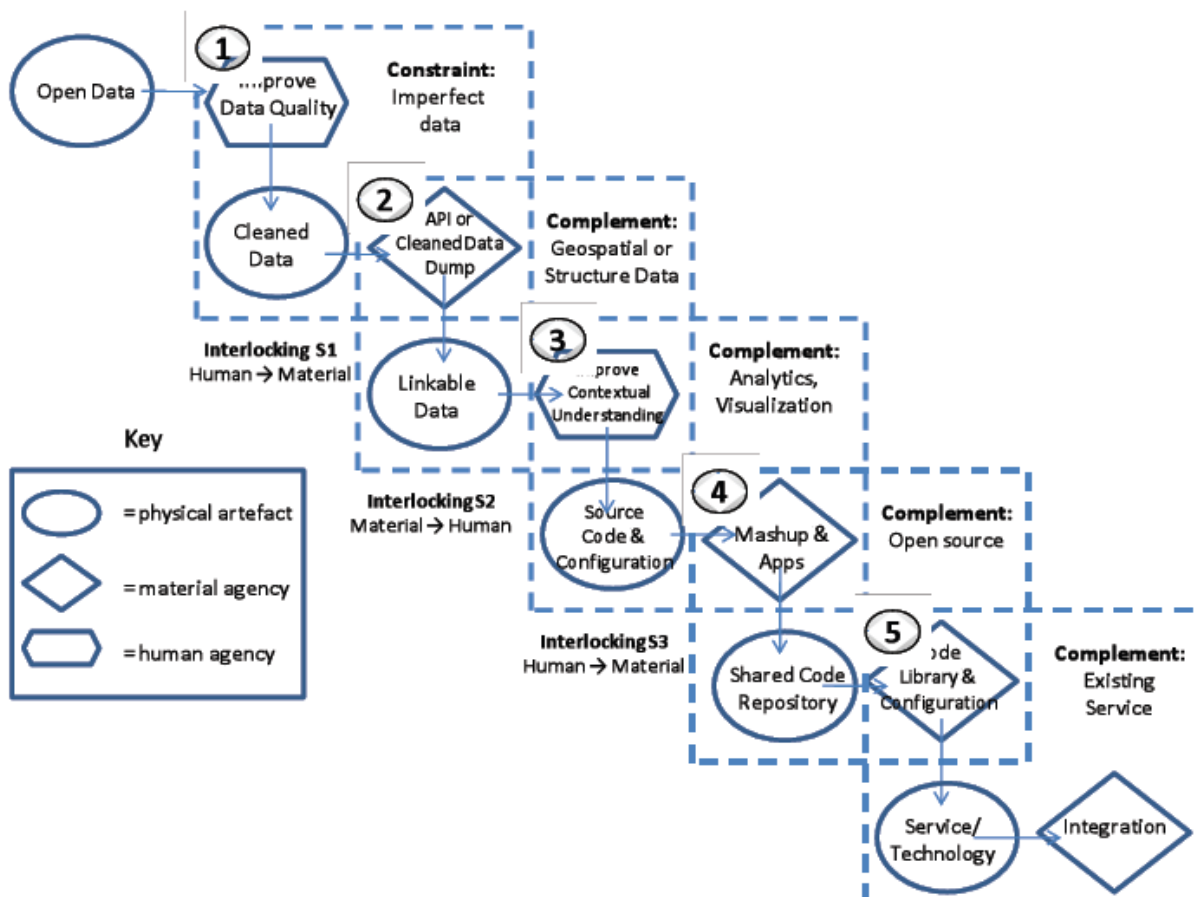
1.3 What is an open data ecosystem?

An **ecosystem** is defined as:

“A system of people, practices, values, and technologies in a particular local environment. Interacting, relatively tightly connected components with substantial interdependencies. The product of such interconnected and interacting components is continuous systemic change since when one element is changed, effects can be felt throughout the whole system.”
(Harrison et al., 2012, p. 906).

When applying the concept of ecosystem to the open data field, we could say that connected components can involve human as well as material agencies¹ - being themselves part of private or public organisations or not- with data and information extracted from raw data representing a big part of the production of this ecosystem. An open data ecosystem consists of multiple data chains interconnected and of various lengths. Kuk and Davies (2011) give a good example of a potential data chain combining inputs from human and material agencies (Figure 1). In this example, 'cleaned data' are transformed into 'linkable data' by a material agency adding geospatial or structured information. From there, human agencies are on the one hand able to construct analytics and visualizations from the linkable data by coding, and on the other hand, they improve the general capabilities by sharing their 'source code and configuration'. Again, a material agency can gather the codes and configurations with the potential of being made open source etc. (Kuk & Davies, 2011).

Figure 1 - Assemblage of open data complementarities (Kuk & Davies, 2011, p. 11)



¹ In their paper, a material agency represents an agency in which the technology has a higher importance than the human input to achieve certain tasks.

Finally, the concept of the ‘open data ecosystem’ is used to explain the three major characteristics experienced in the open data dynamics: autonomy, interactions, and interdependence (Harrison et al., 2012).

First, actors of the open data ecosystem are developing **autonomously** one from the other. This is easy to understand by the nature of open data. Since open data is accessible and re-usable by anyone or anything with no need to notify the publisher, any person or organisation can decide to enter the ecosystem and participate to its elaboration without any control from a third-party.

Interactions between open data actors and, in a broader view, between different networks of actors represent the second characteristic. Modern communications enable each actor to participate and interact in multiple ways with other groups of actors.

The third characteristic is the consequence of the former ones. Provided that actors in the open data ecosystem are autonomous in their decisions, each decision made by an actor can have an impact on others through different direct and indirect interactions with each other. Open data actors are **interdependent**. For instance, if a tool proposing open data aggregation and enrichment from publisher to re-users would have to disappear. Because of the direct and indirect relations with re-users and end-users, it could lead to important effects (Davies, 2011).

But one should not confuse the ecosystem and the infrastructure supporting this ecosystem. While an ecosystem is made of “emergent, autonomous and self-organising components, linked together in local and global feedback loops and developing according to local specialisations and adaptation rather than top-down design.” (Davies, 2011, p. 3). Infrastructures are “basic physical and organisational structures and facilities needed for the operation of a society or enterprise” (Oxford English Dictionary Online cited by Davies, 2011), in this case the operation of the ecosystem.

Until now, the public sector has been the main developer of open data ecosystems by supporting the creation of an infrastructure with the focus on the publication of open data. Basically, open government data ecosystem is often used to describe the central role of the public sector in the open data ecosystem.

For example the revised PSI directive (or Directive 2013/37/EU) in Europe draws a framework around “economic aspects of re-use of information [...]. It encourages the Member States to make as much information available for re-use as possible.” (Pop et al., 2016, p. 11). The term ‘information’ reports

to all material collected by or at disposal of public organisations (organisations funded or under control of public authorities are included) (Commission, 2015c). As a matter of fact, the revised PSI directive has put the public sector as a determinant actor of the open data movement.

Anneke Zuiderwijk, Marijn Janssen, et al. (2014) claim that, when processing open data, a potential re-user uses more than one infrastructure or tool since they represent “only one part of the puzzle for open data users.”(Anneke Zuiderwijk, Marijn Janssen, et al., 2014, p. 17).

Barbara Ubaldi adds that “building an ecosystem that responds to specific demands asking for the provision of open government data are a necessity to create value out of OGD.”(Ubaldi, 2013, p. 27).

To paraphrase her, feedback loops are an absolute necessity for an open data initiative in order to be able to improve and adapt to its ecosystem needs (Pollock, 2011). We could go further by saying that the vision and strategies for open data initiatives should be decided collaboratively by the main actors of the ecosystem and the initiative itself with an equal representativeness (Davies, 2011).

From our review of the literature and the accumulated experience at PwC from multiple perspectives, we conclude that several benefits emerge from the intervention of governments in the open data ecosystems:

- Open data initiator “can address possible fragmentation of the ecosystem”, for example by “engaging [...] [himself] in process of secondary standardisation”(Davies, 2011, p. 4). Moreover, when linked with cooperation and interaction, this type of initiative can avoid redundancy between different open data initiatives seeking at achieving the same objectives. A good example of the so-called ‘**once-only principle**’ is ‘BeSt-Address’ in Belgium, project which is planned to be fully operational in 2018. It demonstrates a concrete cooperation between regions and federal state to create a unique service of addresses (Vanlischout, 2015).
- For open data re-users who “find it hard to identify tools within the ecosystem that they could use, or have already added value that they could benefit from” (Davies, 2011, p. 4), it is useful to have open data supporters involved in making components of an ecosystem “**visible and shared**” (Davies, 2011, p. 4). Two examples are “providing space for sharing source code and scripts [...] [or] link from data catalogues.”(Davies, 2011, p. 4).
- An open data initiative can provide considerable support to key elements of the ecosystem in order to assure the **sustainability** of these elements. Parsons et al. (2011), talking about data

science ecosystem, say that the “need for a sustainable business model highlights the importance of sponsors and funding agencies [...] in the data ecosystem.” (Parsons et al., 2011, p. 15) In our case, the sponsor being replaced by the open data initiator.

- The open government as open data initiator could also ensure an “ethical framework” (Parsons et al., 2011, p. 16) thanks to the cooperation with key actors of the ecosystem. This would support positively the confidence of all stakeholders involved in it and through this of the sustainability of the ecosystem.

2 Methodology

Our methodology contributes answering the three objectives we have defined. Practically the first two steps of our research (literature review and surveys) are part of the first objective to develop awareness and basic knowledge around the concept of open data ecosystem, how this ecosystem is applied to the Belgian case, and around the factors enhancing innovation through open data. For the second objective, the recommendations on how to support the development of the Belgian open data ecosystem and innovation for all types of open data actors in Belgium are mostly based on the surveys and interviews. The third and last objective is exposed throughout our research, every time an interesting point or a confrontation between different ideas is encountered.

2.1 Surveys

Identify Belgian potential open data re-users	• By desk research: keywords on search engine ; list of participants to open data events in Belgium ; developers of applications listed by open data portals. • By networking: ask key organisations and people active in the Belgian open data ecosystem if they have pre-established listing.
Verify the list	Multiple reviews (individual and external) of the newly-created list in order to delete errors in the listing.
Write the questionnaire (Annex 9.1.1)	Based on the literature review, on previous open data surveys conducted by other research organisations (included PwC Belgium) and on the research questions to be explained.
Send and remind	By e-mail in a first step, after the third reminder, by call to the potential re-users for whom we have found the phone number.
Collect the data (Annex	Delete the different cases for which data are missing, organise the data

9.1.2 and online link ¹)	in the right format.
Analyse (Annex 9.1.3)	By using contingency tables, distributions and other statistical tools.
Mine information	Gather and synthesise the different information before drawing conclusions.

2.2 Semi-structured interviews

Identify Belgian open data champions or relevant cases (Annex 9.2.2)	• Belgian open data literature review • Attendance at the event 'Open Belgium 2016' • Contacts established during the surveys • PwC Belgium's network
Write the interview's guidelines (Annex 9.2.1)	Based on the literature and the results of the surveys. Next to this, construction of a graph connecting the main characteristics of an open data ecosystem in order to be confirmed or infirmed by interviewees.
Conduct interviews (Annex 9.2.3)	In the interviewees' office or by video conference/phone. The graph initially developed in the guidelines is updated according to the feedbacks, comments and opinions shared by interviewees.
Qualitative analysis	• Similar topics discussed during the interviews, with opinions convergent or divergent are gathered: organisation of the different information (citations, ideas exposed etc.) per topic discussed independently from the convergence or divergence of the opinions expressed (Miles & Huberman, 1994 cited by Lejeune, 2013) • Synthesis of the information and first conclusions (Miles & Huberman, 1994 cited by Lejeune, 2013) • Informal calls to some interviewees to clarify few details and also to confirm the plausibility of the conclusion proposed (Miles & Huberman, 1994 cited by Lejeune, 2013).

¹ https://drive.google.com/open?id=1jY5P2fRuXtRI2gE5Oxuc9_KZZpo&usp=sharing

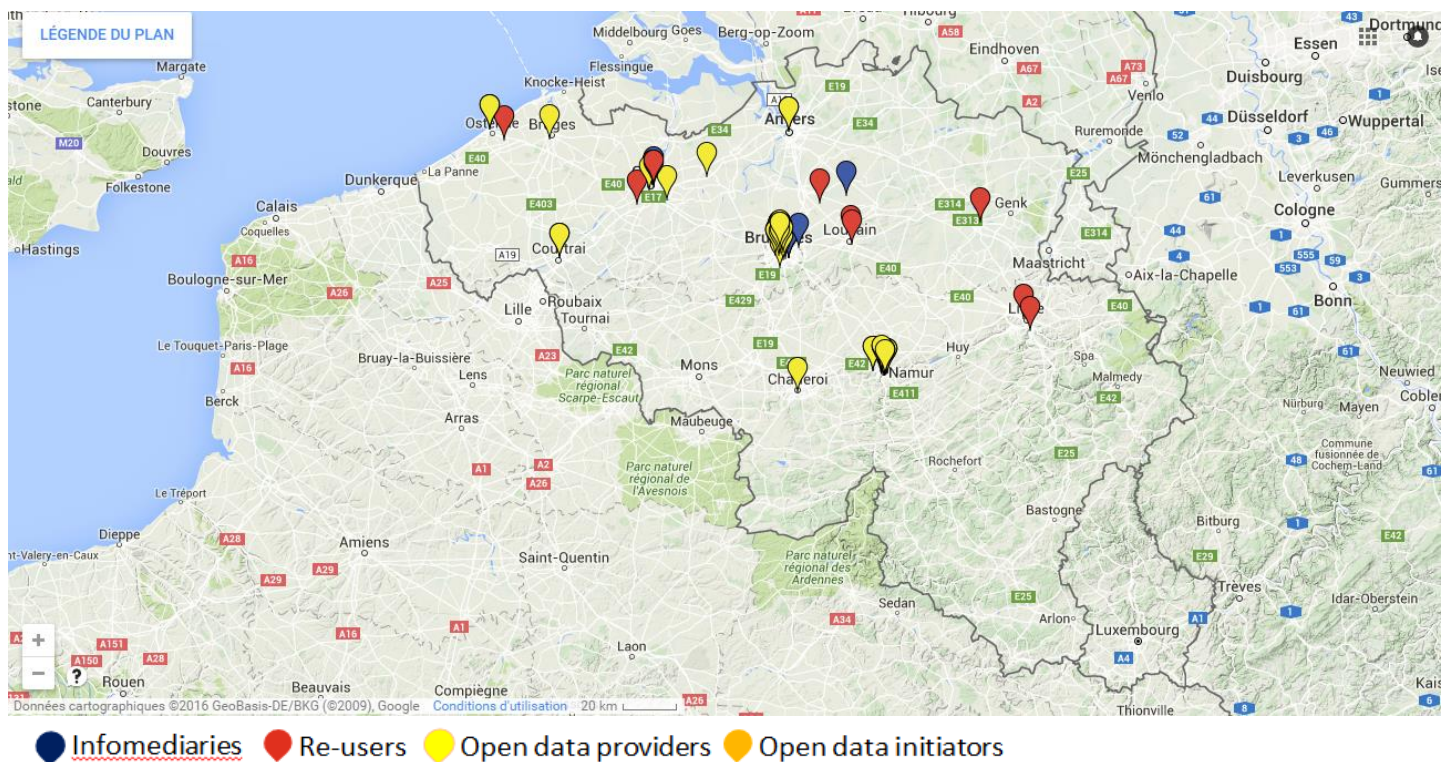
3 The Belgian open data ecosystem

3.1 Actors of the ecosystem

In our researches, we have identified 99 actors of the Belgian open data ecosystem. An interesting point is that all the different groups which are commonly composing open data ecosystems are present in Belgium.

Figure 2 shows a map of the actors of the Belgian open data ecosystem per category of actors. An interactive map is provided online.¹

Figure 2 - Map of the Belgian open data ecosystem*



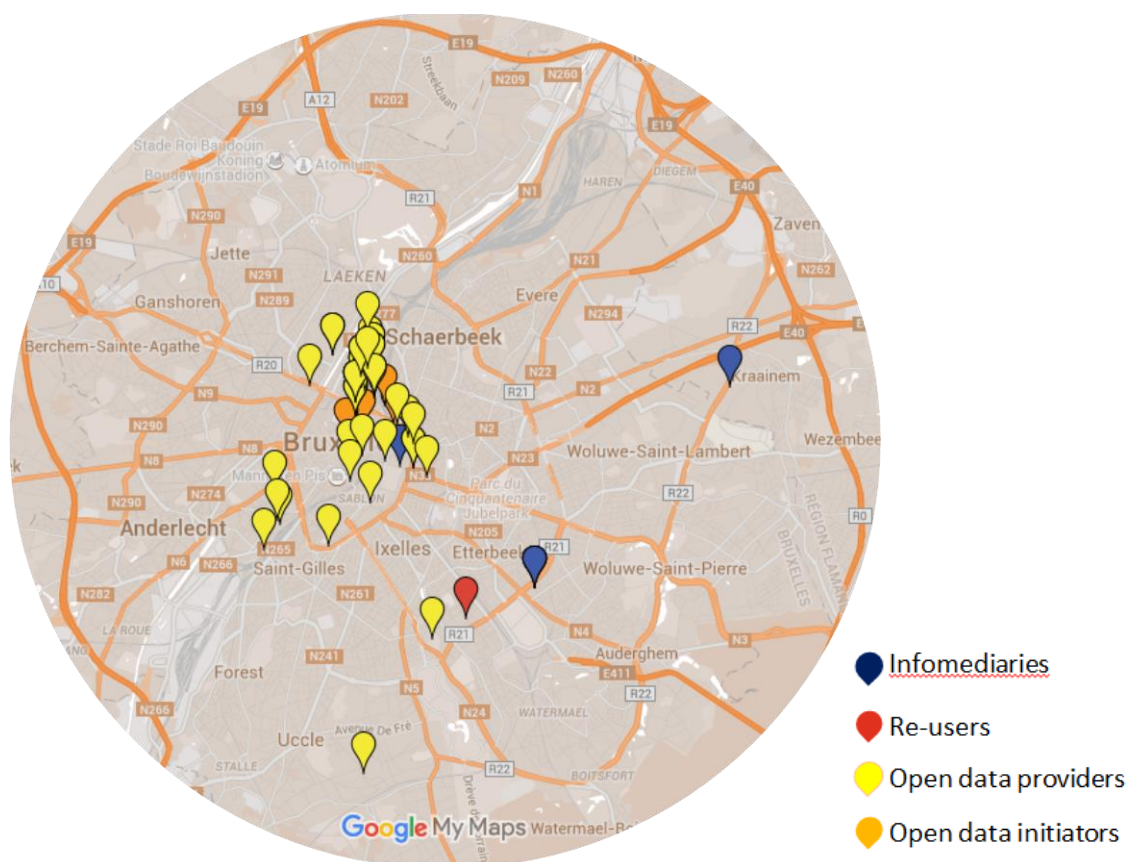
*Belgian cities active in the open data are typically combining multiple roles. In the purpose of making easier the interpretation, we categorised them as 'Infomediaries'.

In Belgium, existing open data ecosystems are mostly centred around three points: the areas of Brussels, Namur and Ghent with respectively 41, 14 and 8 actors. If Namur has more actors than Ghent, its actors have started open data activities more recently and represent mostly public administrations providing open data consequently to the government decision to be involved in the open data movement (Hucq, 2016). We highlighted the Brussels area in Figure 3. The areas of Antwerp and Liège are also showing recent activities related to open data. The different Belgian

¹ Interactive map available at https://drive.google.com/open?id=1jY5P2fRuXtRI2gE5Oxuc9_KZZpo&usp=sharing

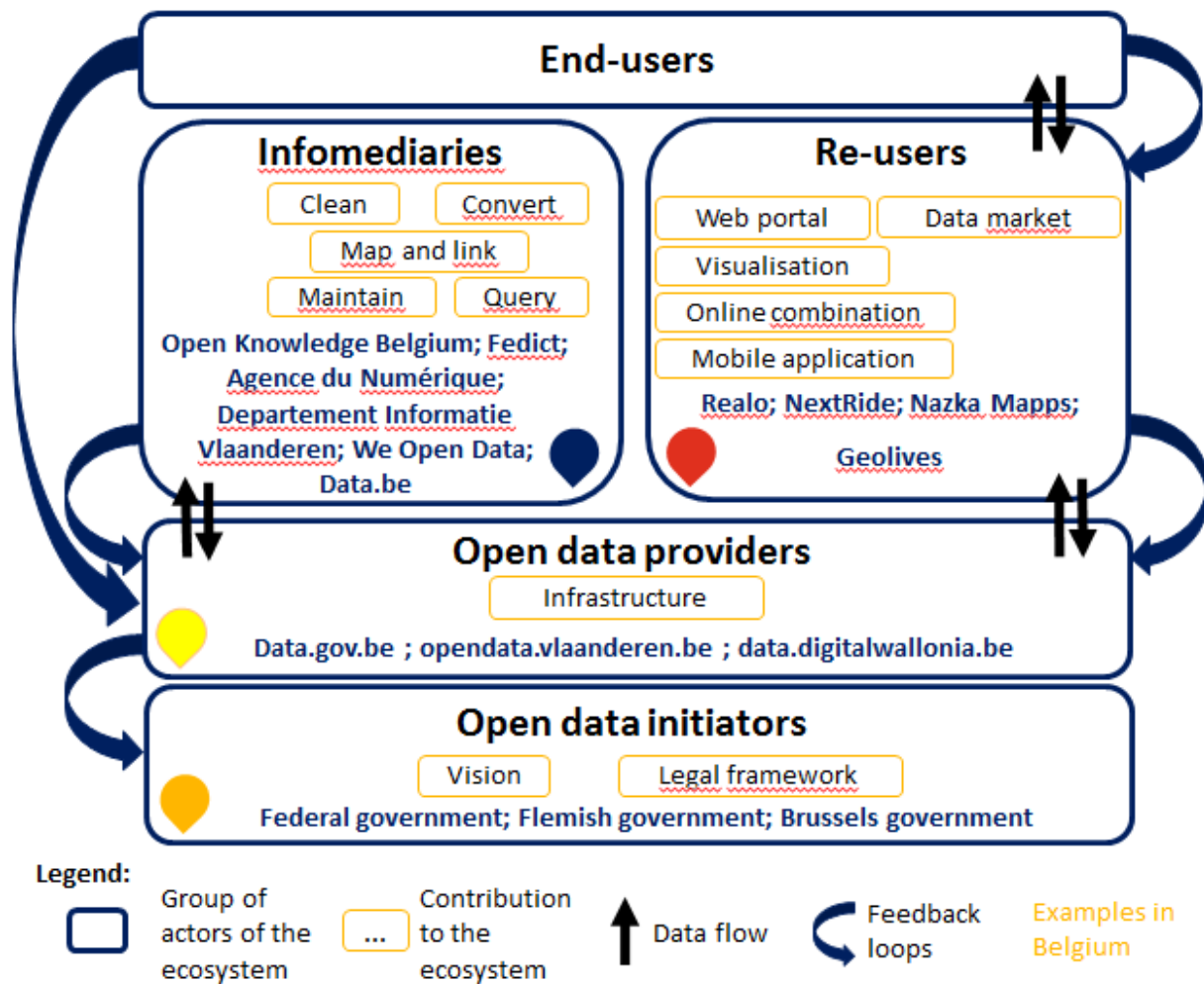
governments (Federal, regions and communities), in a similar way as many other countries, have been the main initiators of open data by building a legal framework for open data. Consequently, we can predominantly talk about a top-down approach adopted in a first period of time in Belgium even if some parts of Belgium are right now ahead of this first purely top-down move. One important actor has also been active in the creation of open data awareness in all the different groups since 2011: 'Open Knowledge Belgium'.

Figure 3 - Map of Brussels area inside the Belgian open data ecosystem



Based on existing representations of open data ecosystems in the literature, the different groups of actors with their main roles have been synthesised in Figure 4, with the data flows between these groups, and concrete examples among Belgian actors. It should be noticed that the group 'infomediaries' (or 'enablers') consists at the same time of public agencies and private parties. The first ones are responsible for opening government data but both types enable re-uses of open data and the establishment of an ecosystem in Belgium.

Figure 4 - Representation of the structure of the Belgian open data ecosystem



Finally, we created a table with the repartition of the actors per group and per region of Belgium (Table 1). From this regional perspective, most of the re-users are situated in Brussels-Capital and Flanders. Flanders was the earliest adopter in Belgium and Brussels-Capital gathers the most significant proportion of actors even if some of them are active at other regional levels as well. From a national perspective, open data providers represent the first category, followed by infomediaries and open data re-users. It seems to confirm the top-down approach adopted in Belgium for open data. Open data initiators stimulated public administrations and organisations to open-up their data, justifying the relatively big number of open data providers compared to re-users. In the future, this repartition should be rebalanced, the re-users becoming a more important group of actors. To confirm this forecast, it could be relevant to analyse the outcomes (to be published) of the survey conducted by the 'European Data Portal' (Glickman, 2016) on open data re-users in European Union

countries. By comparing the different repartition of groups per country and the open data maturity of the countries, one should be able to build an evolution model.

Table 1 - Repartition of the Belgian open data actors per region

Belgium	Actors	99	100%
	Open data initiator	4	4%
	Open data providers	60	61%
	Infomediaries	19	19%
	Re-users	16	16%
Flanders	Actors	38	38%
	Open data initiator	1	1%
	Open data providers	21	21%
	Infomediaries	8	8%
	Re-users	8	8%
Wallonia	Actors	20	20%
	Open data initiator	1	1%
	Open data providers	13	13%
	Infomediaries	3	3%
	Re-users	3	3%
Brussels-Capital	Actors	41	41%
	Open data initiator ¹	2	2%
	Open data providers	26	26%
	Infomediaries	8	8%
	Re-users	5	5%

To the question whether we can already talk about an open data ecosystem or not in Belgium, the figures we have showed here, built from our researches, the surveys and the interviews, are demonstrating that, yes, an open data ecosystem exists and is being further developed at the moment. If metrics to evaluate empirically the maturity of an ecosystem do not really exist except for the data publication part (Harrison et al., 2012), the interviews enabled us to compare the different perspectives on the current state of the Belgian ecosystem per group of actors.

¹ Important remark: for Open data initiators and Open data providers, the different actors based in Brussels but active at the regional level like the different Flemish administrations etc. have been categorized by the region of activity.

3.1.1 Open data initiators, open data providers and open data public agencies

As shown by the map, in 2016 all the different administrative areas in Belgium are covered by at least one public agency responsible for opening up data. The federal level as well as the Flemish region and the German speaking community (Gemeinschaft, 2015) have transposed the revised PSI directive from the European law into their own law level and the Walloon and Brussels regions are waiting for the vote in their respective Parliament (Colpaert & Vanagt, 2016).

In this context, the Federal level (government, 2015), Brussels (cirb-cibg, 2015), Flemish (Vlaanderen, 2015) and Walloon (Numérique) regions all of them have launched open data portals. Concerning these portals, if we follow the model proposed by Colpaert, Joye, Mechant, Mannens, and Van de Walle (2013) to differentiate data portals in 5 different categories, we can situate them on the first two bottom levels, respectively “a dataset registry” and “a meta-data provider”. The next stage consists in developing a “co-creation platform” with ways to assure interactions about the datasets and to share re-use practices. This can be undertaken before the four-star level, “A data publishing platform” which implies the publication of datasets through the portal itself, resulting in Linked Open Data and URI maintained in a sustainable way. The last and fifth level, “A common data hub” enables data portals to act as a hub or a common resource for all actors (Colpaert et al., 2013, pp. 3-5). For Belgium, it means that a lot of work still needs to be done to reach the next levels. But as mentioned by N. Van Herreweghe (personal communication, May 23, 2016) in his interview for the case of Flanders but applicable to other parts of the country: “We can be proud of the ground covered by open data implementation.” For example, the ‘Open Data Manual’ that the ‘Departement voor Informatie’ has published, is currently translated in several languages has a guide to achieve open government (Flanders, 2015). We analyse below different perspectives about the state of open data for open data initiators, open data providers and open data public agencies.

The first perspective starts with data strategies adopted by governments at the moment, each representative interviewed set his two main objectives as followed: developing a transparent state and maximising the value behind open data. In order to achieve these, they put as first priority the publication of data. Priority which has been concretised by the development of open data portals. But we should pay attention to the fact that, as detailed further in the section ‘Open data re-users’, re-uses are not only dependent on the availability of data. Therefore, top management could need more interactions from time to time with external contributors to help them set up awareness

around open data and find how to support best open data re-uses. This lack of awareness in the Belgian ecosystem leads partly to two important consequences. First, resources allocated to Belgian initiatives did not always match the needs considering the result expected. When taking into account the complex situation in Belgium, obtaining a level of openness equivalent to neighbouring countries should require a bigger initial investment to connect the different initiatives existing, both at the strategical and technical levels. Second, without rapid results measured in terms of economic or social value, initiatives have to face uncertainties about their funding. Yet, for example in open science, Parsons et al. (2011) defend that successful initiatives are implying funding policies for data managers and data centres for a longer period than four years. It is therefore important to embedded long term vision and long term funding through the support of the top management. Carrara, Fischer, and van Steenberghe (2015) drew a similar conclusion in their report on the European open data landscape in 2015: top management should be implied in the success of open data initiatives. For them, open data strategies are often 5-year plans, and if Parsons et al. (2011) advocate for a longer-term vision which is probably in line with the needs of open data initiatives before concrete benefits of open data will be developed and precisely measurable, five-year plans represent schedules closer to the reality of the current political world.

As second point, now that all Belgian open data public agencies have launched their own data portal, they can start undertaking new initiatives to improve open data. The first priority for some of the public open data agencies is still to convince most of the public administrations and local governments to release their data. Once data are made available openly on a data portal, new initiatives can emerge, always with the final objective of maximising the social and economic value of open data. Examples of current projects are: finding some proof-of-concept regarding linked data or finding some business models on how to make re-use of open data. But the implementation of their projects is limited most of the time due to multiple reasons, the lack of interactions with the different public agencies being an important one. During the interviews, representatives of these public agencies admitted that if on the one hand, open data public agencies are independent one from another; on the other hand they lack interactions between each other. Right now, a good counterexample is the harvesting structure which has been implemented through the collaboration of the different public agencies active in open data. Local open data portals are harvested by regions or directly by the federal portal and the same is being done for regions. But except for a few examples, public actors of open data lack regular discussions with each other in order to construct a

shared vision, a common line of actions, and to communicate on their best practices. This has concrete impacts on re-users like the small scale of some datasets. For example, the city of Antwerp has published datasets related to localisation and tracks of public transports. However, since it is the only local authority releasing these datasets in the area, J. Vansteenlandt (personal communication, May 27, 2016) from 'We Open Data' states that these "data are really interesting but limited to Antwerp area it is really hard to re-use."

Another observation highlighted by the map and related to the previous example is that the number of local governments involved in open data initiatives is rather low. The cities of Ghent, Antwerp and Brussels are the three active ones at the moment, as well as the province of Namur which proposes one page of its website to search for open datasets. The cities of Bruges and Namur are willing to start while, according to N. Van Herreweghe (personal communication, May 23, 2016), the two cities of Ostend and Kortrijk have been quiet for a few months after a good start. Bart Hanssens from the federal open data initiative agrees with Noël Van Herreweghe when saying that many local governments and public agencies still need to be convinced to release open data (B. Hanssens, personal communication. 2016, May 24; N. Van Herreweghe, personal communication, May 24, 2016).

At local level, the leading example of the city of Ghent is demonstrative. The city was able to develop a real ecosystem around their open data agenda. For T. Thoeve (personal communication, May 25, 2016), the city of Ghent achieved this result for multiple reasons. First, a number of managers from different levels, including top management of the city, strongly agree that transparency is an important goal for a public administration. Second, the city of Ghent can be seen as more than just a data provider but also as a community builder: "For the public sector, about 25% of the total investment in an open data initiative is technical and 75% of the investment is on community building for events, workshops, mailing list etc." Third, the European level initiative has also provided a legal base for local governments to build open initiatives. Fourth, compared to other levels of governance, locals are the first ring of contact with private actors, which is important to collect voluntary feedbacks.

As we mentioned before, at a legislative level, Belgian entities have also either already developed or are catching up on the 'open by default' model. They follow the guidelines from previous works and directives like the Open Data Charter from the G8 or the revised PSI directive of the EU.

Nevertheless, developing an ‘open by default’ program for transparency raises data security and privacy considerations (Verhulst & Young, 2016). N. Van Herreweghe (personal communication, May 23, 2016) claims that culture is an essential factor concerning the level of privacy people attach to their life and therefore the transparency level they are ready to reach. “Openness is about culture, about being transparent. [...] In Flanders, everything is sensitive.” What he is saying here can be applied to the whole of Belgium. He gave us one example: in Flanders, data about fine dust are known and accessible to everybody. At a moment someone was willing to link it to topology data. From this interesting application, this person started facing comments of people telling him that the value of houses around Zaventem airport were going to drop due to his data mashup. This consequence was obvious, but it represents by essence what transparency means: giving the choice to people who would like to buy a house there and are sensitive to such information. On the other hand, in order to be useful, it needs to be accepted by the people and it is a fact that Belgians are very careful concerning privacy and security of their own data. That is also why T. Thoeve (personal communication, May 25, 2016) is assertive with the fact that actors in the open data ecosystem will invest in ways to anonymise data in the future. Statement to which T. Vanagt (personal communication, May 30, 2016), chairman of Open Knowledge Belgium and active in different initiatives and re-uses of open data, adds that technologies to ensure the protection of identity are available. He illustrates his statement with one example in the legal sphere:

“In Belgium very few law cases are released when in other countries it is different. For instance, in the Netherlands, every law decision made by the government is fully published in an anonymised format. It is the court itself that anonymised within a few days after the decision has been taken with the support of a software. The software takes as inputs all names, addresses and other identifiers elements and release as outputs a cleaned dataset. Then a human verify the cleaning up process of the software. There is not much work because the whole document is already cleaned out. In Belgium, we are laying behind because we are not giving the right IT tools to people and agencies.”

3.1.2 Open data enablers or infomediaries

In the last section we have introduced public agencies acting as enablers of open data. We will now describe private enablers of the Belgian ecosystem. They are determinant actors of an ecosystem supporting the development of open data applications and services by proposing intermediary services like “data catalogues or marketplaces” (Archer, Dekkers, Goedertier, & Loutas, 2013, p. 13) to assure access and availability of data. Open data enablers “also offer additional services, such as

advanced querying, data visualisation, exporters of data in different formats.” (Archer et al., 2013, p. 13)

Our research confirmed that different business models of enabler exist (Janssen & Zuiderwijk, 2014) like intermediate data platform or technical enabler (M. Janssen & Zuiderwijk, 2014). Examples of these models in Belgium are ‘data.be’ and ‘lex.be’ for the first model, ‘We Open Data’ for the second model. The first two enrich data from different sources to propose new data on a platform to re-users. In the case of ‘lex.be’, data from the ‘Belgian State Gazette’, the website ‘Juridat’ and the ‘Federal Public Service Justice’ are combined to propose enriched law data and a freemium service for every citizen or company (lex.be, 2015). We Open Data can be considered as a technical enabler since it offers tools as the ‘Datatank’ “which allows users to semantify, publish and annotate data.” (Data, 2011). Some consulting firms like ‘High Latitudes’ and ‘PwC Belgium’ can also be considered as enablers since they propose services similar to both models. For example, ‘PwC Belgium’ has a role as a technical enabler for the Flemish public open data projects and on the European level. Finally, the position of the Open Knowledge Belgium Foundation is mostly related to enabling open data in Belgium but it cannot be limited to that. In fact they are also active as advisor of open data initiators or with different working groups around multiple topics of open data: ‘iRail’, ‘Open Street Map’ or ‘Open Tourism’ (Belgium, 2012).

At first glance, the Belgian open data ecosystem seems only supported by a few enablers as pointed out by B. Hanssens (personal communication, May 24, 2016): many well-known enablers of the ecosystem are people with an expert profile and not companies. But when combining desk researches and results of the surveys and interviews, we were able to draw a different picture for Belgium. Enablers of the ecosystem are numerous and with various profiles as it has been shown. However, opportunities still exist for new enablers, for example in Wallonia where open data are likely to be consistently developed in the coming years.

3.1.3 Open data re-users

From the survey on Belgian re-users of open data, we were able to make some observations on the re-users. As highlighted by T. Vanagt (personal communication, May 30, 2016), there are many re-uses, “but it is not always clear from a side based upon datasets.” The difficulty is in the identification of businesses re-using open data. This section is based on the analysis of the 17 complete answers received from re-users of open data among the companies we have been in touch

with. By clicking on the interactive map (link in the previous footnote), the full answers for each open data re-user¹ can be seen. Our key findings are highlighted below (Annex 9.1.3.1).

First, concerning the profile of re-users, most of the businesses are profit-oriented (65%), they propose services in B2B only (35%) or in B2B and in another orientation like B2C, B2G or other (65%), they have between 0 and 5 employees (47%), they are active in the 'Information and communication' sector (53%) and started open data projects after 2011 (94%). Examples of services proposed are collecting, analysing and visualising private and open data. To achieve this, some re-users like 'Nazka' offer businesses the entire service and the opportunity to build interactive maps (Nazka, 2013) . Other re-users like 'Cumul.io' propose a software so that businesses can build up their own analysis and visualisation internally. But as in the case of 'Cumul.io', they also propose additional services like making a catalogue of open data available in the right format to be directly integrated into the software (Cumul.io, 2016).

To describe profiles of re-users in other words, a new market can be said to emerge. In fact 2011 corresponds to the launching of the first clear initiatives in Belgium at the federal level (first version of data.gov.be) and at the Flemish region level (decision for a strategic framework for open data policies) (Commission, 2015b). This new market was mainly composed, in the first years, by profit-oriented companies. Even if non-profit organisations have also been launched for a few years, these are still a minority in Belgium. The majority of re-users have activities directly related to ICT, with 9 re-users considering themselves as active in the 'Information and communication' sector and the 8 others offering at least online services linked to data. The most attractive business orientation concerned B2B services. This last statement raises questions:

1. Is the Belgian open data ecosystem mostly oriented towards B2B services or is it just a coincidence in the profiles of our respondents? In order to answer this, more researches are required to identify a bigger sample of open data re-users.
2. If the first hypothesis is correct, could we suppose that B2B is more attractive for open data re-users than B2C and B2G or do independent factors explain this difference?
3. Finally, in the case in which B2B is more attractive than B2C, could a lack of citizens' awareness for open data represent a key factor to justify this bigger attraction?

¹ For each respondent who agreed to share the information collected by the survey

We will focus more deeply now on a description of the profiles based on the businesses developed around data re-uses. Approximately the proportion of re-users who work with open data as their main revenue source (47%) is the same as re-users who work with open data without seeing it as their main revenue source (53%). This shows that businesses look at open data as an added value to their value proposition, even if their core activity is not centred on open data.

Table 2 - Profile of the respondents using mostly 'Transport data' (Annex 9.1.3.2.1 - 9.1.3.2.7)

Profile: 80% of the respondents using transport data consider open data as their main revenue source. The same proportion proposes their services through a 'Mobile application' and a 'Website'. Besides, 80% of this category of re-users use data from at least 'Local' and 'Regional' sources. This last observation may be quite specific to Belgian transport data re-users since most companies providing transport data are regional (ex: De Lijn, STIB, TEC) but it could be interesting to know who the local sources are.

Every transport data re-user is combining internally open data and their own data. All these re-users do not exploit 'Often' support tools provided and see 'Data access & availability' as the main challenge. Finally, they all have a minimum of '100-1000' end-users of their services, which is above the average.

The data domain used most is related to 'Transport' data (29%), followed by 'Government and public sector' data (18%). It is in line with the Commission (2015a) and their European Data Portal study 'Creating value through open data' proposing 'Public administration', 'Industry' and 'Trades and transport' as the three domain with the most potential value.

47% re-users have a 'subscription model', 29% have 'consulting fees' model and 29% have a revenue model different from the five

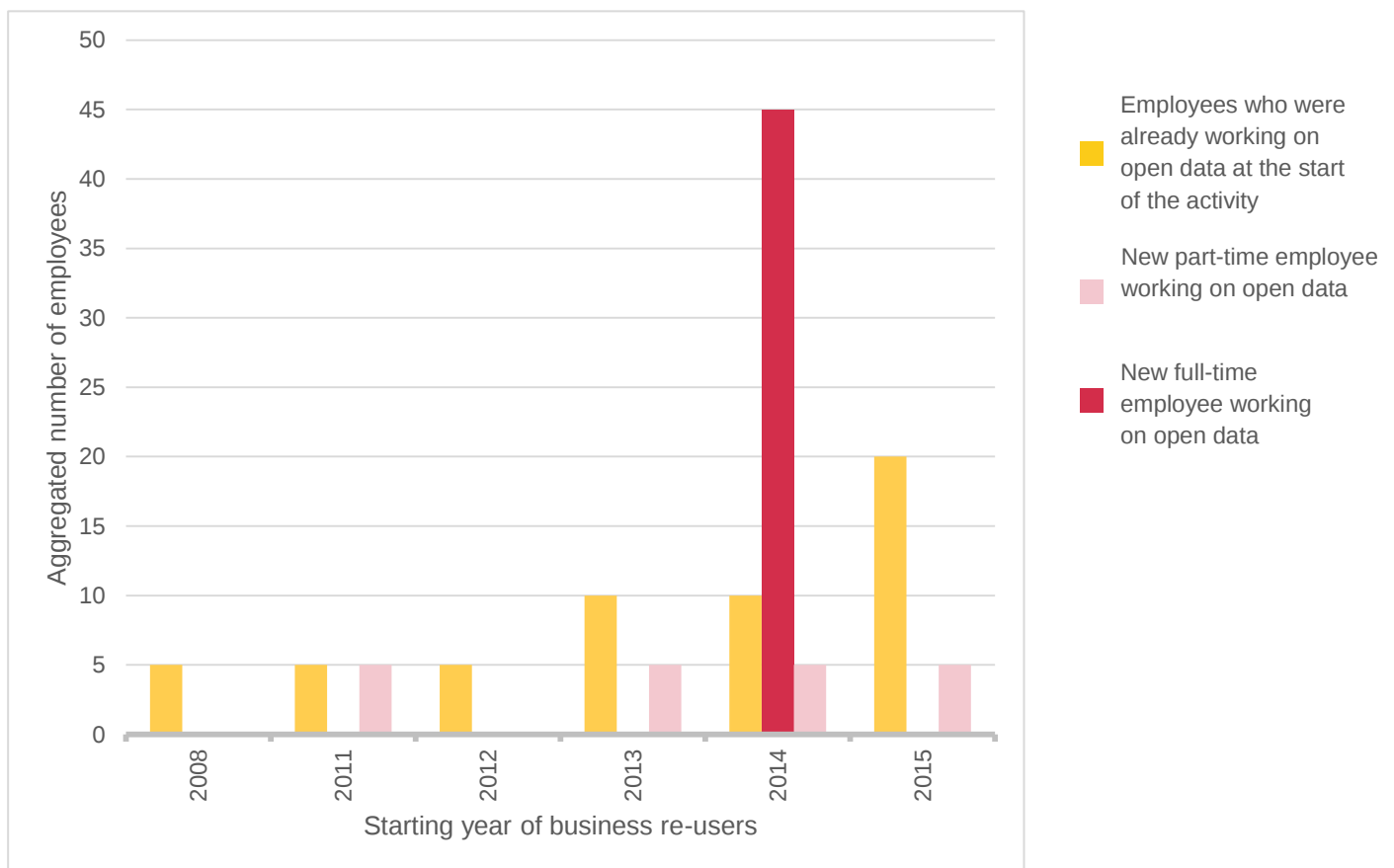
proposed¹, for instance subsidies or lump sum. From the interviews it can be added that 'consulting fees' model will probably keep increasing in the future in the Belgian open data ecosystem.

Most of the open data re-users have not hired any new employee related to an open data job since they got involved in open data projects (65%). N. Van Herreweghe (personal communication, May 23, 2016) specified in his interview that open data businesses on average need a long period before becoming economically viable. Therefore, it could be a factor justifying why companies are not hiring, this is especially true since 72% of the re-users in this case have also started open data

¹ Subscription model, Consulting fees, Analytics fees, Advertising, and Lead generation.

projects from 2013. In a future study it could be interesting to measure the average time required for an open data re-use to become viable and compare it to the time required by this re-use before hiring new open data employees. Figure 5 shows the aggregated number of employees working in the Belgian open data ecosystem in 2016. The employees are categorised by the starting year of the company in open data projects and if the company has hired new open data employee since then or not.

Figure 5 - Number of employees working in the Belgian open data ecosystem in 2016 related to the starting year of open data re-uses



In their open data business activities (Annex 9.1.3.2.8), 65% of the re-users are at least aggregating data while 47% are at least developing an activity on top of open data. For what concerns their distribution mechanisms, 76% of the re-users propose their services through at least a website and 59% through at least a mobile application with only one re-user proposing services through only a mobile application. Companies which do not run a website nor a mobile application use for example

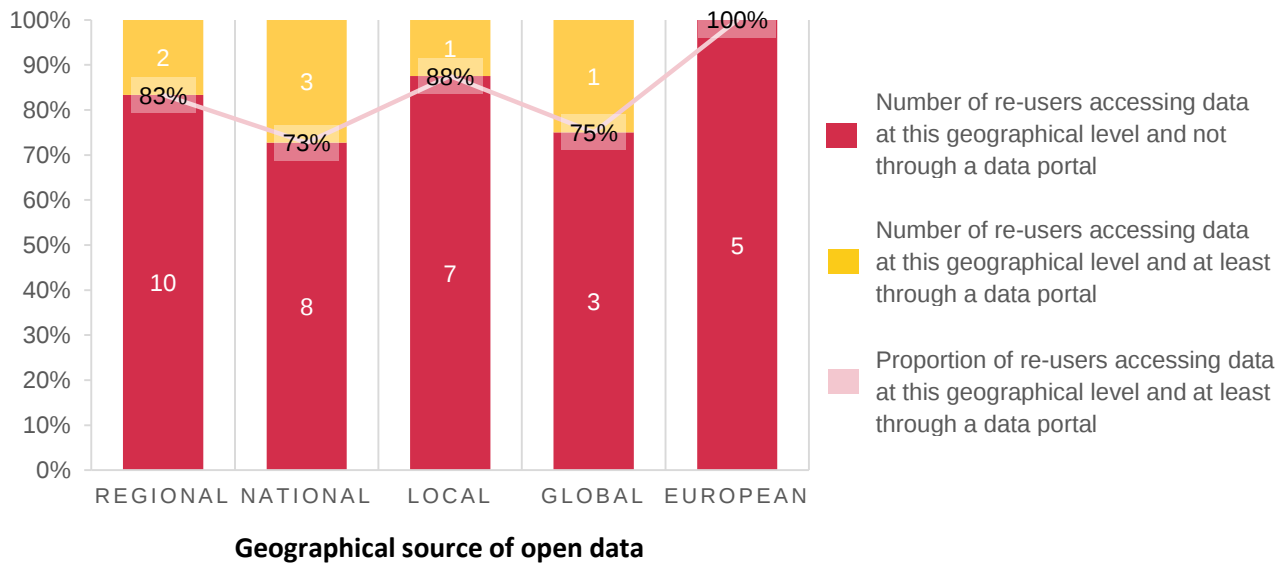
social media or data-as-a-service platform. In synthesis, aggregating activities, developing activities as well as a website and/or a mobile application seem to form the basis for many open data re-users. For aggregators, it could be interesting to notice that all re-users with consulting fees as revenue model are also aggregators of data (Annex 9.1.3.2.9). Further researches should evaluate what the correlation and causality between these two characteristics are.

Almost all re-users combine data internally (88%), joining the opinion of T. Vanagt (personal communication, May 30, 2016) about the value in internal combination of open data and private data.

The favourite sources of open data for Belgian re-users are for 71% at least regional or national, and for 94% at least from local, regional and national level (Annex 9.1.3.2.10). It means that re-users using at least Belgian data are twice as numerous as the 41% using at least European or global data sources. It matches the findings for the quality and quantity of services proposed at regional and national levels in Belgium. However, re-users choosing at least local sources (47%) show us that the demand side for local data providers exists. One challenge deduced from the development of local sources is to limit the number of interlocutors for a potential re-user. Interests for European and Global sources is real as well, which is confirmed by Y. Peeters (personal communication, May 24, 2016) from GeoLives, viewing in open data an opportunity to create “one integrated product around the world.”

For what concerns the technical type of data access (Open data portal, API, e-mail to the data provider, website of the data provider), each type was accessed by at least 8 re-users, meaning that at the moment no access can claim to make the unanimity (Annex 9.1.3.2.11). Two possibilities arise from this last point: either it shows that the Belgian open data ecosystem has the space to develop and progress; either it attests that a good density of interactions between actors and data providers has already taken place, meaning that re-users know where and how to find specific data. In any case, Figure 6 below shows the number of re-users per geographical source and accessing data through at least data portal (plus potential other access types like e-mail, API or website) or not through data portal but only by other ways.

Figure 6 - Number of re-users per geographical source of open data and accessing open data at least through a data portal or not



What we can see is that at the Belgian level, some open data re-users do not work at all with data portals. Concerning the portals this suggests either a lack of communication and promotion for potential users of their services, or a lack of interests from open data re-users for the services currently proposed by the data portals.

As we have already introduced before, the growing number of people and organisations building on top of open data demonstrates the attraction existing for this market. Moreover, the emergence in the last three years of non-profit oriented re-users attests that economic benefits are not the only motivations for potential re-users, supporting the findings of the literature.

During the interviews, we asked what the main motivators were that convinced the interviewee to start open data projects. Two respondents answered that the frustration coming from potential re-uses which were not developed because of ignorance or ill-will was the main motivation. Another interviewee put forward the transparency gain for the society, for instance with 'Open spending' project. Two others were attracted by open government data as a source of trustfulness in end-users' minds. As just explained, Y. Peeters (personal communication, May 24, 2016) highlighted the advantage of open data as a value source without boarder. And finally, the last one has seen in open data an opportunity to build services and to have real impacts on the development of a new market.

3.2 Interactions in the Belgian open data ecosystem

As the Open Knowledge Index assessed in 2016, Belgium has made a step forward in the open data ranking from the 53rd place in 2014 to the 35th in 2015 (Knowledge, 2014). In some key domains, Belgium is really catching up with the neighbouring countries as for ‘Government Budget’, ‘Election Results’ or ‘National Map’. ‘National Statistics’ did not improve between 2014 and 2015 but are already at a good level. Pieter Colpaert was one of the assessors for Belgium in 2015. He is summarising the state of Belgium in a presentation as changing positively but there is still room for improvement (Colpaert & Vanagt, 2016). Generally speaking, Belgium improved between 2014 and 2015 because all the different levels of the country started to be involved in the open data ecosystem, as open data initiator, provider, enabler or re-user, and public authorities maintained their efforts. This is a good point for Belgium since the country is characterised by multiple factors (languages, cultures, repartition of the public responsibilities, situation of Brussels etc.), which creates a complex situation on certain points compared to other neighbouring countries.

When asked what an open data ecosystem is, N. Van Herreweghe (personal communication, May 23, 2016) pointed out one of the most important characteristics: the interactions between actors of the ecosystem. “The concept of ‘ecosystem’ gives some insights about the way you look at it: actors, parties dependent on each other etc.” All types of interactions are organised through communication mechanisms. We clustered the different mechanisms that we encountered during our research in three categories: automated interactions¹, semi-automated interactions and human interactions. This is shown in Table 3, per objective pursued.

¹ Pay attention that for automated interactions, we understand automated as the automatic process for the data/service provider to collect the raw information from the interactions with other actors. In conclusion, a human input can be needed (for example to fill a web form) but no intermediary action is required to transfer the information to the right person.

Table 3 - Communication mechanism used by Belgian open data actors depending on the objective pursued

Objectives	Automated interactions	Semi-automated interactions	Human interactions
Data quality	Online comment (data.gov.be/en/node/10030)	Web form (data.gov.be/en/contact), Mail	Roundtable (Open Mobility Data Flanders) Events (e.g. Open Belgium, hackathons, Creative Ring Challenge, Trefdag , etc.) Working group (iRail , Open Tourisme) Debate
Data request		Web form, Mail	
Data benefits/value	Websites analytics (e.g. Google analytics)	Mail	
Support to re-use		Web form	
Other requests, Feedbacks		Web form, Mail, Forum (forumen.sitytrail.com)	
Strategy			
Methodology			

During most of the interviews, the respondents highlighted the fact that interactions between actors of the ecosystem are informal at the moment or even that no interaction is organised. In the case of Realo, it was impressive for S. De Meyere (personal communication, May 30, 2016) to see the difference of interactions, first between Realo and other private companies and second, between Realo and public administrations. A concrete example for Realo was when they were contacted by 'Cambio' using Twitter, asking whether it could be possible to see 'Cambio' sharing points on the interactive map proposed by Realo on their website. Four days later, the map had been updated. About the relations with governments, S. De Meyere (personal communication, May 30, 2016) is mainly citing the amount of interlocutors and the lack of clarity concerning the respective responsibilities of administrations among other barriers to establish two-way interactions. In a more general perspective, interactions can adopt different characteristics: formal or informal, structured or unstructured, concerning objects (for example data) or thoughts (for example policies) etc. An example of informal and unstructured feedback is the one being applied by 'lex.be' & 'data.be' on the data quality of the different official sources of information they use:

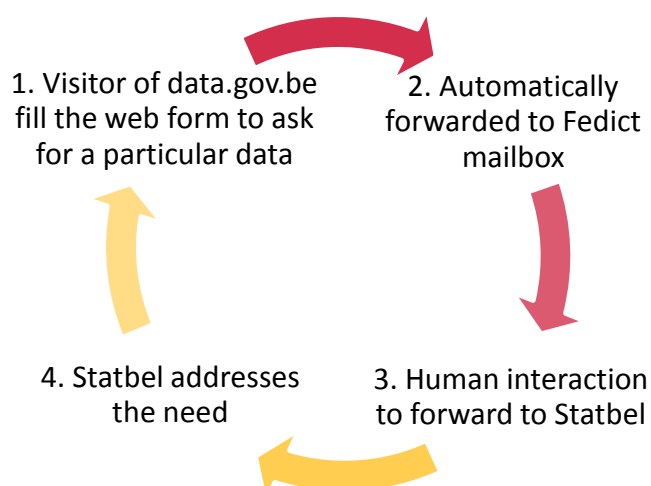
“When errors in the data sources are detected by Lex.be or end-users of Lex.be, Lex.be provides a feedback by e-mail to the right people in the data provider organisation. But this is only because they know personally those people, how to contact them and what they are responsible for.” (T. Vanagt, personal communication, May 30, 2016).

This example can be generalised to other objectives than only data quality: data request, other requests etc. In a normative way, feedbacks should be provided directly to the right data providers and not through the intermediary of the open data re-user as it is the case right now. For T. Vanagt (personal communication, May 30, 2016) a structured feedback should have two characteristics: an easy way for end-users and re-users to give feedback to the data collector with possibilities for the data collectors to assess quickly if it is relevant or not.

Some agencies inside the public sector have already implemented semi-automated open data feedback mechanisms. An existing case is the feedback loop represented in Figure 7 between visitors of ‘data.gov.be’, ‘Fedict’ (public agency responsible for federal open data projects) and ‘Statbel’ (federal public organisation managing the national statistics).

Figure 7 – Example of feedback mechanism existing for Statbel.

(B. Hanssens, personal communication, May 24, 2016)



When asked what were the main factors explaining the success of this semi-automated feedback mechanism, B. Hanssens (personal communication, May 24, 2016) claimed that on one hand, ‘Statbel’ is publishing data which can be easily combined, creating a high demand for their data and therefore a willingness to interact from the re-users. On the other hand, ‘Statbel’ has been active online in data collection and analytics since 1998, they work with skilled people and experience

about what a potential re-user wants to see on the data portal and what this potential re-user will be able to get from a specific data. B. Hanssens (personal communication, May 24, 2016) also remarked that the main problem in this example is not the system requiring a transfer by e-mails and more human actions but to address the re-user's need which often includes the opening up of the dataset for which the party providing the feedback is looking. All interviewees from public agencies agreed that convincing administrations to open up their high value datasets and their datasets in general was still a big challenge. S. Vince (personal communication, June 2, 2016), from the Walloon Numeric Agency, described the perception of many administrations towards government data as a perception of ownership. A cultural shift is needed inside public administrations for a simple reason, the network effect. P. Colpaert (personal communication, May 30, 2016) explained it clearly, the more people, companies and organisations will get involved in the movement for more openness, the more impact this movement will have— positive for active initiatives and negative for uninvolved entities. In this network effect, we are not only talking at the Belgian level. If Belgian entities do not invest time and resources to develop openness while more and more countries and regions of the world do, the confidence of international partners in Belgium and wellbeing of the Belgian population will be affected. In this debate with public administrations, S. Vince (personal communication, June 2, 2016) attested that the influence he can have on these administrations is insignificant due to the little top management support he gets. He hopes that the 'Open by default' decree will change the balance of bargaining powers.

Concerning human interactions, many of the different interviewees, and among them all the representatives of data providers, are in favour of the establishment of dialogues with the different actors present in the open data in Belgium. But some initiatives in this direction have already started: for the second year, the Flemish government is organising roundtables on specific topics of open data, gathering various types of actors. The first year they focused mainly on constructing a first discussion but this year they have concrete objectives such as knowing what the needs of the demand side of open data are (N. Van Herreweghe, personal communication, May 23, 2016). For example, end of May 2016, a roundtable around 'Mobility and Open Data' took place in Brussels, tackling various issues as open data strategies and methodologies. Some suggestions were proposed on how to aggregate and leverage initiatives by local governments or how to keep data up to date. Another discussion mechanism already implemented in Belgium is represented by the different events organised by various actors and with several purposes. 'Open Belgium' and hackathons are

the most famous of them. Open Belgium is an annual one-day event, organised by the 'Open Knowledge Belgium foundation', with the aim of gathering all the different actors of the open data ecosystem in a one-day workshops and discussions event (Belgium, 2014, 2015). Hackathons are more specifically aimed at opening data re-uses and to create open data awareness. They usually consist of a one or two-day event where teams with various skills and backgrounds try to build beta applications, for mobile or not, from the datasets released by the organiser. The major limit of hackathons is the low rate of viable businesses created from the ideas or applications proposed during the events. A third type of event which appeared this year in Belgium consists of a more long-term challenge in the perspective of creating real businesses around open data and supporting them to become viable. One example of the latter type is the 'Creative Ring Challenge' in Ghent (Gent, Ghent, & Eskidoos, 2016). An interesting analysis of the factors enabling these long-term challenges to be more successful in helping re-users to start projects was made by PwC (2015) on a similar challenge in the UK. Appropriate repartition of resources and time to participants, adequate support and expertise were the main factors (PwC, 2015). An interesting point outlined by T. Vanagt (personal communication, May 30, 2016) is that events supporting re-uses of open data in general are good opportunities to push potential re-users to really start a project in specific deadlines. But at the same time, it is sometimes difficult to have data with quick beneficial value for re-users willing to participate as J. Vansteenlandt (personal communication, May 25, 2016) outlined. S. De Meyere (personal communication, May 30, 2016) also found that these events should better describe and communicate about their follow-up and objectives. In conclusion, if events supporting re-uses of open data are not perfect, they are clearly contributing to the global advancement of open data to achieve the objectives of a better society. One cannot just view those events for what they are not pursuing. The most important is to define consciously the objectives and to communicate in an appropriate way about it to participants in order to achieve the objectives.

Finally, P. Colpaert, N. Van Herreweghe and T. Vanagt (personal communication, May 30-24-30, 2016), also remarked that the problem of informal and unstructured feedbacks and the lack of discussion mechanisms is not a national problem. As far as they know, no country at the moment has been managed to develop a high density of interactions inside their open data ecosystem with structured, automated feedbacks and regular discussion mechanisms.

4 Innovation and open data in Belgium

Up to now, we have centred our talk on the Belgian open data ecosystem and the actors composing it. More specifically for re-users, we described some observations that we were able to find about their profiles and motivators to start open data projects. But we have not yet detailed the relation between innovation and open data. Or in a narrower perspective, how is innovation made possible thanks to open data and open data ecosystems?

There are “strong claims [...] about the potential of opening government data to drive service innovation.” (Kuk & Davies, 2011, p. 1) But so far, insignificant amount of knowledge has been gathered about how services are created from datasets, and what are the factors encouraging the context in which data release are transformed into service innovation. (Kuk & Davies, 2011) Or more generally, key metrics to assess open data innovation are difficult to define due to several reasons like “the wide variability in outcomes” (Verhulst & Young, 2016, p. 16), the focus being on different areas than innovation itself or the difficulty of tracing the value-chain for data which, by its nature, required no formal relationship between supplier and user (Kuk & Davies, 2011). Concretely, in the purpose of gaining some insights about the link (correlation and causality) between innovation and open data initiatives and, at a later stage, to be able to “scale the impact of the individual case studies [...], we need a better – and more granular – understanding of the enabling conditions that lead to success.” (Verhulst & Young, 2016, p. 16) In *‘Innovation and Big Data in Smart Service System’*, Maglio and Lim (2016) defend the creation of a model that can describe service and data together from the data perspective. In more concrete terms, a service would be detailed as a set of variables that could be measured. It would be helpful to design and evaluate service. Up to now in the facts, compared to other sectors, little is known about data-driven services.

In our interviews, S. De Meyere (personal communication, May 30, 2016) from Realo described briefly how they were building metrics internally to evaluate the impacts of their open data services. Further investments are planned to achieve this process. For example, by collaborating more closely with real estates agencies to know the proportion of clients who were convinced by their services. At the same time, they are also developing social metrics to construct a neighbourhood rating. Internal evaluation is a reality in every type of business. P. Colpaert (personal communication, May 30, 2016) compared open data today to websites 10-15 years ago: at the time, no tools existed to assess the impact that a website had on your business activity and results. Now, services like Google Analytics

propose intuitive tools to assess interactions with your website etc. It is not perfect, but it is an improvement towards finding information and constructing criteria.

In order to identify some factors specific to open data innovation in Belgium, our approach combines the findings of previous researches and applies it to the Belgian open data ecosystem thanks to the different interviews and surveys we conducted. Articles on innovation are numerous, with a general focus on companies or a more restrained focus on re-users of data or open data. Therefore, some features extracted are generic for all types of companies while others concern only business building data services. With this funnel technique, we finally kept two different articles, one by Ceschin (2013) on *'Critical Factors for Implementing and Diffusing Sustainable Product-Service Systems'* and another by Jetzek et al. (2014) on *'Data-driven innovation through open government data'*. We present the combination of the most relevant factors of each article for innovation in the Belgian open data ecosystem in Table 4. We added different inputs, either from the literature or from our researches. The first part concerns government's influence on factors.

Table 4 - Description and examples of the main innovative factors observed in the Belgian open data ecosystem

Government	
Assure openness	Opening up datasets is the first condition and if it "might be seen as the necessary condition in this context, it is insufficient on its own" (Jetzek et al., 2014, p. 117) to sustain concrete re-uses.
Support to re-use: Data quality, Sustainability, Interoperability, Re-user's contribution, Protected environment.	Usability and re-usability could be supported by a better quality and sustainability of data. For P. Colpaert (personal communication, May 30, 2016), it goes first through a less costly interoperability for potential re-users. "We need to bring interoperability to such a high level that re-users can just start re-using it without having to pay a lot to bring different kinds of datasets together." He is summarising interoperability with 5 different categories: 1. Legal: Are we allowed to bring together different datasets? In Belgium, we need to educate people to put a license on their data. 2. Technical: How can we bring together the data physically? We have the

internet.

3. Syntactic: Do we use XML or...? It is not really a problem anymore, because programming languages can convert one to another.
4. Semantic: using the same words or terms for talking about the same things. We are working [at international level] towards that.
5. Querying: How can we ask questions about datasets? Because a question does not simply stop at a machine. And that is where we need to perform further work: building blocks on the Web which provide certain functionalities on top of a certain datasets.

For S. De Meyere (personal communication, May 30, 2016), a **bottom-up approach** should be more **integrated into current services**. It should be easier for re-users to contribute to datasets released, for instance by helping to update the data. He adds that current changes in our society also imply changes in the content of the jobs and open data should be one of those changes. For example: civil servants but also citizens in everyday life could help to update everyday life data as geolocation of bus stops or for citizens, any data directly related to them personally.

A **protected innovative environment** could be indicated to assure the right context for 'socio-technical experiments' or open data re-uses to emerge. It could take the form, for instance, of less fiscal charges to assure the right context to take shape. This protection should be softened once businesses re-uses are making profit and scaling up.

Business

Capabilities:
Analytic and
IT skills,
Absorptive
capacity.

Potential re-users must have the right **capabilities**, including **analytics and IT skills**. It should also include a mindset encouraging learning processes or an **absorptive capacity**. An absorptive capacity is defined as a capacity of a company to identify, assimilate and apply an external knowledge to a new case (Jetzek et al., 2014). For an open data re-user, it could be the identification, potentially from data analytics,

	of an opportunity in the behaviour of clients for instance and the consecutive knowledge necessary to catch it. For Realo, the construction of the 'neighbourhood rating' is probably based on an identified opportunity and the knowledge related to it (S. De Meyere, personal communication, May 30, 2016).
Shared project vision	A long term consensus from the actors involved is a necessity. With the protection of the environment, the right capabilities and a shared project vision , re-users should be able to reach a critical mass. A point from which their activity is theoretically viable.
Develop a broad network of actors	"The establishment and development of a [...] socio-economic network is recognised a crucial process to protect, support and foster radical innovations. [...] It should be a heterogeneous network characterized by scientific, social, economic, politic and cultural linkages. The network should be managed dynamically because different stages of a societal embedding process require different network compositions." (Ceschin, 2013, p. 8) Governments are important actors in the open data ecosystem but the limit to which they should stop interfere is not always clear. We discuss this point in the 'Findings and recommendations' part. What is certain is the support an open data ecosystem can bring to innovation as a dynamic network of actors interacting.
Access to technical platforms – partnership	New technologies allow re-users of open data to "analyse, mash up and make sense of different types of data" as well as to develop "the networking capabilities to access and link data from various sources." T. Vanagt (personal communication, May 30, 2016) sees the mash up between opened and closed data as the present and future big source of value for open data businesses. N. Van Herreweghe (personal communication, May 23, 2016) adds that the combination between different domains like open data and IoT can bear opportunities as well. According to Jetzek et al. (2014), the three main components of these technical platforms are: 1. "The infrastructure that facilitates data exchange between government agencies, private sector firms and the public

2. The dissemination of software, including data organization management software, as well as analytics and discovery software
3. The access via multiple platforms, such as mobile and web-based platforms.” (Jetzek et al., 2014, p. 113)

From our findings, we can add up a fourth component to these three, which is the integration of interactions mechanisms.

In the case of Belgium, the role of governments and infomediaries are determinant to enable potential re-users to start open data projects. Together, they propose multiple open data platforms like open data portals or platform similar to the one proposed by ‘lex.be’ as well as technical services and interactions opportunities. Verhulst and Young (2016) are viewing this as the construction of **partnerships**. For them “two forms of collaboration [...] are particularly important: partnerships with civil society groups” (Verhulst & Young, 2016, p. 17) (mobilizing and educating citizens) and “partnerships with the media” (Verhulst & Young, 2016, p. 17) (inform citizens, analyse and find meaning in raw open data).

5 Findings and recommendations

5.1 Benefits of open data in Belgium

In the current state of open data in Belgium, precise benefits from open data have been more assessed at individual and company levels. Measuring national and global benefits is still debated in order to know who should conduct it. Moreover, measuring tools, adopted by a large set of actors, have not been developed yet. On a small scale, the answers we got from the interviewees were as diversified as open data itself is interdisciplinary.

- Economic benefits:

Starting with the economic impacts, many studies have already been published about the economic impact of open data but no consensus has been reached yet about one measurement model. Criticisms about current measurement models are mainly related to the lack of concrete results which should take into account the biggest part of the open data impacts. For example, in a growing open data business with a sustainable competitive advantage, what is the contribution from open

data? As a consequence, new measurement models or the convergence between multiple existing models is required to help companies and governments to better estimate the value of open data.

What we can say about Belgium is the following: some re-users have constructed a viable business around open data, counting fully or partially on open data as revenue source. 41% of the re-users who completed our survey had assessed an average improvement above 5% in increase of revenues or decrease of costs.

We have also highlighted before that the Belgian open data ecosystem has been attractive for a few years for some organisations to start a project.

A third observation from the surveys is that 35% of re-users have hired one full or part-time open data employee since they have started open data projects. Consequently, jobs are created first when the company is founded and second for the growing companies when they hire new employees. In concrete numbers as showed by Figure 5 (p. 21), the 17 re-users contacted represent between 37 and 120 jobs related to open data (since they had answered for categories, e.g. 0-5 employees). More than 80% of the jobs were created from 2013 to 2015 and all the respondents who said that they had hired new open data employees after having started open data projects did so in those years.

For what concerns costs saved by public administrations thanks to open data, we found no evidence during our researches. Beside this, as demonstrated at the federal level, open data agencies work in collaboration with agencies for the administrative simplification. Meaning that open data is seen as a way to simplify and improve the efficiency of public administrations.

- Open data as an added-value for businesses:

Another benefit from open data was the large scale of information accessible. It creates new business opportunities for re-users of open data to start an open data project. For example with open data, a start-up can access and mine datasets from all over the world in order to get information and develop a competitive advantage in their services. To fully achieve this, data portals must reach the fifth star of open data and become a kind of “global decentralised knowledge accessible by anybody” as described by P. Colpaert (personal communication, May 30, 2016) but in

which some companies and entrepreneurs will identify opportunities than others cannot see. At this point, open data would represent an added-value for businesses to scale up rapidly.

The development of new combinations of open data and other information also shows the added-value of open data to a business. For consulting and IT services companies, it represents a potential market: companies possessing data and willing to combine it with open data available but without experience or skills to do it internally will ask external services to manage it.

- Societal changes:

We can observe many changes from a societal point of view, like for example 'Google Maps' and re-users of open data proposing better routing path for public transport thanks to the data released for instance by the Belgian railway company 'NMBS/SNCB'. Keeping track of all these changes is impossible at the moment. A second effect complementary to the concrete societal changes we can assess is the network effect about which we have already talked. The bigger the number of actors inside the ecosystem is, the bigger the impacts in the form of positive and negative spill overs are, depending on whether the individual or business is active or not in the open data.

Individuals can also profit from social benefits in terms of auto satisfaction or personal interests for the job's content for example.

The transparency effect was one of the first gains outlined by open data advocates, even if publishing data in itself is not a gage of a real gain in transparency as explained in our introductory part 'Why open data'. For Jetzek et al. (2014, p. 106), "transparency mechanisms can enable value generation if they reduce information asymmetry." One example could be 'lex.be' which decreases the information asymmetry between citizens and large organisations by enriching Belgian law data and making it easily accessible online.

5.2 Organising the Belgian open data ecosystem and supporting open data re-uses

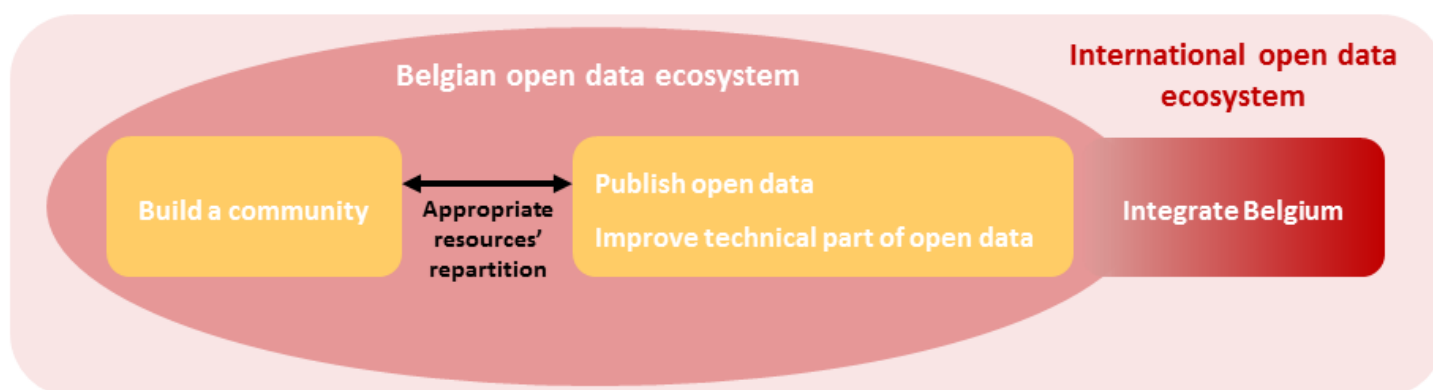
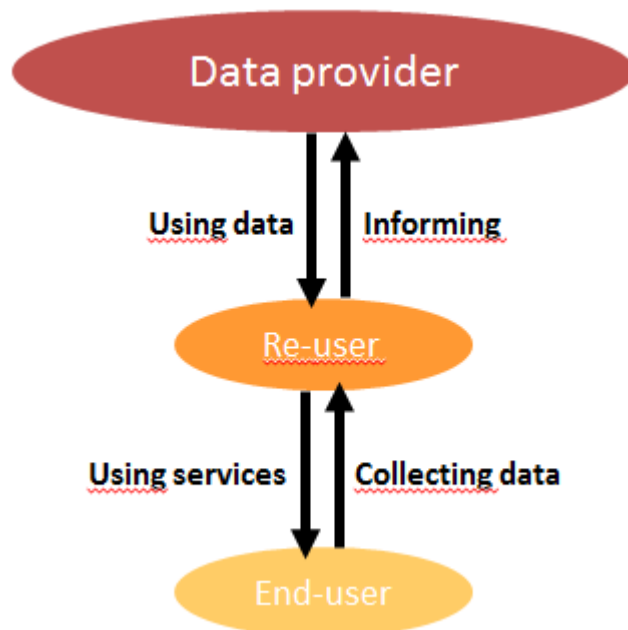


Figure 8 - Synthesis of our recommendations for the development of the Belgian open data ecosystem

The recommendations we have set up from our researches can be synthesised by this graph (Figure 8). Two pillars form the main blocks on which the Belgian open data ecosystem would be able to develop further and achieve the economic and social benefits aimed. The two pillars are 'Build a community' and 'Publish open data – improve technical part of open data'. We will start with the first pillar and after explaining the second one, we will discuss briefly our recommendations regarding the other elements of this graph: the resources' allocation between the two pillars and the integration of the Belgian open data ecosystem into the international community.

What we have described for economic benefits of open data in the previous part also applies for social benefits: actors in Belgium and internationally are still struggling to measure the social impacts of open data. Nevertheless, the need for measurement models is real. For data providers to justify the investments allocated and for re-users to be attracted by the opportunities open data represents. In this process, the role of public agencies is not isolated. **Measuring the benefits of open data is a necessity, but not under any conditions. Guaranteeing the free accessibility of open data consists in one of those conditions.** If open data portals could ask a simple registration proposing advantages for re-users like notifications of data updates or new datasets available related to their data activities, it could enable the portals to assess different information and at least obtain a number of actors in the ecosystem. But this **registration should be done on a voluntary basis**, to keep secure the access for re-users who would like to remain anonymous. Another condition concerns the questions asked to re-users to assess the benefits of open data. Those questions should not cover the business model and value proposition of re-users to ensure the real openness of data released.

Figure 9 – General interactions on the main axe of the data chain to measure the benefits of open data



As described more in details with the graph (Figure 9), **measuring the benefits of open data is not only related to public agencies but it is a community responsibility**. All different actors of the ecosystem will be involved, like re-users in developing new tools and models to measure benefits or to provide voluntary feedbacks. We will develop later on (in Table 5) some ideas of interactions mechanisms which could help assess re-uses of open data. In fact, measurement models will go along with the development of interactions mechanisms. For the simple reason that the better interactions mechanisms will be implemented, the easier it will be to collect data and information about the different ways datasets are re-used or why a data domain is more attractive than the other. In this process, consulting companies can play an important role to help re-users developing the new models required and integrating them into their existing systems.

Starting from the bottom of the graph, by using the services proposed by re-users, end-users help them to collect information and complete the measurements models built to estimate the benefits generated by open data. From there, re-users are able to propose different aggregate information to the data providers in order to help them know what is the value behind the open data they publish. The type of information given by re-users could be decided between the actors in bilateral or multilateral discussions.

Some conditions are required to enable this information system to appear. First, **new measurement models should be developed, mainly by re-users**, taking into account multiple levels and impacts of their services, for example, the success rate of their service offered or the importance of open data into this service and success etc. The second enabling condition should aim at **establishing a discussion mechanism between the actors at multiple levels**. In a first move, large events could be organised to determine who are the actors interested in participating to a dynamic open data ecosystem and what their specific needs are. If some specific needs are identified, in a second move, smaller meetings should be organised, formal or informal, between the specific actors willing to implement close interaction mechanisms. In general, the interacting actors inside the Belgian open data ecosystem have to choose an **independent organisation able to federate the different events and interactions existing at the moment**. It could be a way to officialise every important event and be sure that each event is differentiated from the others. In this way, actors would be sure that new joiners will not plan new events related to the same topics as the events already federated. But this **federation should also encourage new measurement tools and models to be spread inside the ecosystem** as well as be in contact with international organisations to know what is made elsewhere.

The role of the public agencies in this process towards an information system is twofold. On the one hand, **implement different interactions tools in their services in order to adapt to the new measurement models developed**. And on the other hand, **support the establishment of a dialogue and of real interactions between the actors of the ecosystem without trampling on the competencies of the federation**. The two entities must act in complement of each other.

Concretely, we summarise in the Table 5 different interaction mechanisms already existing but not yet implemented in the Belgian open data ecosystem:

Table 5 - Communication mechanism which could be implemented by Belgian open data actors depending on the objective pursued

Objectives	Automated interactions	Semi-automated interactions	Human interactions
Data quality	Direct contribution (allowing direct update or direct uploading if providing authentic source (ex: eID) or types of data subject to data collaboration (peer review)) Web form per category (sent directly to the right responsible: engagedata.eu) Online dataset rating (engagedata.eu/dataset/87725/egri-the-e-government-reference-library-v115-cf) Automated online survey (Pop-up like Skype etc.)		24 hours of open data
Data request	Web form per category (sent directly to the right responsible)		
Data benefits/value	Automated online survey		
Support to re-use	Web form per dataset (type of item re-using the specific dataset, URL of the item)		
Other requests, Feedbacks	Web form per category	Forum	
Strategy			
Methodology			

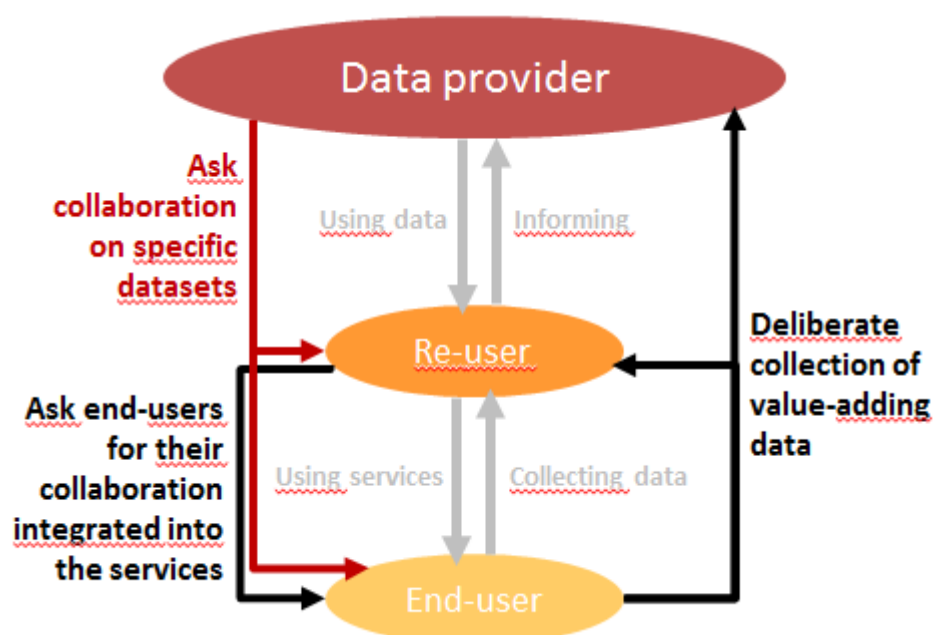
From the table, three of the different mechanisms highlighted are now detailed.

More **direct contribution could help data providers to collect new data without huge investments as well as keeping up-to-date some datasets**. For end-users and re-users, it would represent a way to participate directly to the releasing process and get visibility from data providers. Concretely, we adapted Figure 9 to consider the direct contribution mechanism proposed in Figure 10.

First action should be taken by data providers in collaboration with re-users. They have to decide for which datasets it could be interesting and possible to have a direct contribution of re-users but also from end-users. A challenge to take into account would be obviously how to control the quality of

data. One possibility could be to ask for authentication before changing or adding data, another

Figure 10 – New interaction mechanisms on the main axe of the data chain to measure the benefits of open data



possibility could be to ask for the collaboration of civil servants as part of their jobs' content. Data providers should also implement accessible online tools to allow contributors to bring and change data in a user-friendly way. Examples of data for which it could be considered to allow direct contribution are data related to everyday life (e.g. geolocation of different services) subject to the condition of respecting privacy and security, disasters, administrative procedures, etc. Re-users could help data providers by integrating direct data collection procedures into their main services. It would consist of automated collection of data or deliberated collection, both depending on end-user's agreement. An enabling condition would be to educate end-users how to contribute appropriately to data collection.

A second mechanism would be implemented in data portals for instance. Automated online survey could allow data providers to collect information about:

- The reasons why re-users of open data use some data domains,
- The characteristics of datasets that are the most important for re-users and consequently the characteristics that currently lack,
- Some basic benefits that re-users are able to assess from the utilisation of the dataset.

In practice, the purpose would be to **propose to users of open data to answer a one-window survey** with preferably only one question and a multiple selection. The opportunity not to answer should be provided. The survey could appear as a pop-up window. Beforehand, clearly explaining the reasons why answers from users of open data are helpful for data providers to justify the public investments and improve their services should encourage a better rate of responses.

The last mechanism we highlight is the 24 hours of open data. Right now, only one event gathers all the actors of open data in Belgium on a yearly basis, 'Open Belgium'. The idea of a second national event would be first, to **decrease the time between two national meetings** to half a year and second, to **focus more on attracting potential actors of the open data ecosystem as well as developing awareness and a culture of openness in Belgium at the end-user's level**. Practically, the 24hours could be seen as an opportunity for actors of the open data ecosystem to meet end-users in person. The event could be organised by the federation we have talked about before, supported by different actors.

The second pillar of our recommendations willing to develop an open data ecosystem and support re-uses concerns the datasets released and the technical part of open data.

Many administrations and data collectors in Belgium still need to be convinced of the necessity to open up their datasets in order to support more openness in the country. Flanders has already reached a good level with more than 4,000 datasets published and most of the re-users present in this region. But some work still needs to be done. For example, some high-value datasets are not yet available or only on a too small scale. To reach a level in which at least the high-value datasets are released, **top management support is required** to guarantee a proper funding plan in the long-term and to give more weight to the agency responsible for the open data project regarding the other administrations. A shared vision about the objectives and strategies to reach those is also needed. Next to top management support, a **data management should be instilled in the different data collectors and data providers services** (e.g. administrations and public agencies but also in private actors). The purpose must be the comprehension from all actors of the basic requirements for data providers to publish good quality and standardised data. The essential condition of the legal framework as it has already been mentioned is on the way of being achieved with on 'open by default' model.

The second point concerns the technical side of open data. We have highlighted before the need for more interoperable data, especially for the querying part of interoperability. Beyond the two main objectives of most of the open data projects -namely opening up data and supporting re-uses on top of it- the main goal pursued is to transform our society and the people or companies composing it towards an easier and better existence. In other words, the goal is to make smarter clients or end-users of open data to enable them to take better decisions regarding the problems they are facing. In this process, querying has an important role since it allows open data users to ask for particular subset of data and obtain a better and faster accessibility (Fragments, 2013). Therefore, **public agencies and private actors should support querying their data to make more intelligent clients by decreasing the time spent looking for data and improving the exactitude of the data found for a particular need.** Moreover, improving data query also improves automated interactions for a simple reason. The smarter is a data re-user in his/her query, the more he/she is able to find what he/she is looking for and the less likely he/she would be to make data request to the data provider. In this way, interactions are improved since data providers will be allowed to focus more on data request really missing.

Besides these two pillars, a public agency or other actor investing resources in it has to decide about the repartition of the resources between these two pillars. The repartition depends on every specific context with different characteristics: support from the top management, motivation of the different administrations, awareness in the different groups of actors, citizen's support etc. According to the different degree of achievement of every characteristic, a public agency should invest more in one pillar or the other.

We have introduced in the second part of this report the fact that finding solutions to automated and structured interactions mechanisms were international problematics. Solutions and tools to support it are expected to emerge in the coming years, for instance through a decentralised infrastructure proposing interactions mechanisms for which every user will only keep track of his own interactions. Dat (2014) is a recent project with interesting services related to decentralised infrastructure as "distributed synchronisation" or "intelligent storage".

In a more general perspective, what is important for open data public agencies is:

1. To **stay aware of the approaches adopted inside the international community to build an infrastructure and an open data ecosystem;**
2. To **construct an infrastructure that could be integrated with other ones in later development stages.**

Finally, concerning the public competencies to propose services on top of open data and thus interfere with the private sector, the limit is sometimes hard to define.

First of all, this debate has its roots in the ideology about what should be the competencies of the public sector in a normative way. No consensus exists in Belgium about this issue between the different governments and public administrations.

From our researches, we would state that **governments should not interfere with services which could be offered by private parties.** One exception can be observed on issues for which the data are available but no interest has been showed by private entities to catch it. A concrete example is the development of 'ParkAdvisor' application by the government for people with a disability. Nevertheless, the government is willing to find private partners to launch the application once a 'beta' phase has been developed. In a more general perspective, the limit should be put around the co-creation between public and private parties with always a private finality.

To summarise the different points we have tackled in this report in the perspective of the position adopted by governments, governments should keep both objectives, such as publishing data and supporting re-uses. They have to be certain that open data and data management are in all respective agendas of the public sector. Issues such as how to manage data or how to keep track of metadata and updates must be fixed. But internal awareness is not the only important thing, if awareness is already at a good level in some areas for potential re-users of open data, end-users could be more aware of open data. For what concerns the co-creation limit, it will settle down in the coming months between a direct and active participation from the government in partnerships and an indirect participation in research labs such as iMinds in Flanders or Digital Wallonia Hub in Wallonia. Finally, governments currently represent important actors of the Belgian open data ecosystem. As a result, their participation to create an efficient and dense ecosystem is required, as co-organiser of events or organiser of hackathons.

6 Conclusion

Considering the current state of the literature around open data and open government, there are few metrics and measurement models to assess the generated economic and social value. Those metrics are incomplete and do not fit properly or partially to the reality they should represent. Even if open data is recognized as a source of innovation, the knowledge about the factors that enable the innovation through open data is limited. Our study is focusing on those two points in the context of Belgium. It contributes to evaluate the situation and the impacts of the Belgian open data ecosystem on the innovation and value of open data. What are the best functionalities of infrastructures supporting the ecosystem or how interoperability and data quality can support new business opportunities are underlying questions.

Given the two main topics of research and related questions arising, three objectives have been defined.

First, develop awareness and basic knowledge around the concept of open data ecosystem and how this ecosystem is applied to the Belgian case. The identification of the factors that enhance innovation through open data is part of this objective.

A second objective is to develop a better understanding of these factors to be able to give concrete recommendations on how to support the development of the Belgian open data ecosystem and innovation for all types of open data actors in Belgium.

The last objective consists of highlighting guidelines for further researches on open data innovation.

The following paragraphs synthesise for each objective the approach, the concerns and issues, and the main results.

6.1 Awareness

We exposed throughout our work the different information and descriptions from the review of the literature or the surveys and interviews conducted. We synthesised first the current positions of the **literature about open data ecosystem**, characterised by autonomy, interaction and interdependency of the actors composing it.

In the second part, we sketched the **current picture of the Belgian open data ecosystem**. Concretely, it is the first time an attempt has been made to map the Belgian open data ecosystem

and, if our methodology was not exhaustive, we were able to identify actors from all the different groups commonly identified in open data ecosystems, respectively open data initiators, providers, infomediaries, re-users and end-users. In Belgium like in many countries, **open data initiatives were mainly top-down driven**, with the **focal point being the role of governments as data providers**. It was concretised in the last years by the development of an environment favourable to open data, most of all in the legal and technical spheres with the 'open by default' decrees and the data portals. An important point in the **Belgian ecosystem is the different stages of development between the regions**: there is at least one data portal in every place in Belgium today but with different capabilities in terms of quantity of datasets published on it, and the support from top public management fluctuates.

Despite this, the Belgian open data ecosystem is now moving towards a second phase of development, involving more the actors composing it, to achieve a new level encouraging re-uses. This appearing ecosystem will support more and improved interactions mechanisms, viewed today as too informal, unorganised and unstructured. A remark on this from interviewees pointed out that the construction of ecosystems with structured interaction mechanisms is an international issue.

In combination with a first **map of the Belgian open data ecosystem**, this study also represents a first attempt to identify and study the **profiles of open data re-users in Belgium**. Regarding the profiles, we observed a majority of profit-oriented companies active in the B2B sector. The data domains mostly used are similar to those assessed at the European level and open data is considered by all re-users as an added-value to their business activities. In Belgium, we observed a growing number of re-users since 2011 which corresponds to the real kick-off of the Belgian open data movement. Non-profit organisations also increased their presence for three years showing that open data is attractive for multiple reasons, which are not only related to economic benefits.

The literature has highlighted the lack of metrics and models to **measure the social and economic value of open data**, saying that a more 'granular understanding' was needed. By associating the findings of the literature and the evidence from the interviews and survey, we proposed **several factors for governments and businesses to sustain innovation through open data**. For the government-side, assure openness, quality, sustainability, and interoperability of data as well as make decentralised contribution possible and create a protected environment. For the business-side, capabilities of the people in IT and analysis but also an absorptive capacity, a shared vision, a broad

network of actors and finally a guaranteed access to technical platforms and appropriate partnerships are factors supporting the emergence and viability of re-users.

The **creation of new jobs** and **economically viable businesses** were two benefits attested by respondents to our survey and interviewees even though these benefits are still limited. For example, no quantitative evidence on a large sample have been assessed yet at the public administrations level even if the public sector is waiting for cost reductions as demonstrated by their will to link open data initiatives to administrative simplification. Next to these benefits, open data also represents new business opportunities and a gain in transparency through re-users of data released by public administrations.

6.2 Recommendations

Our **recommendations** are synthesised in two axes: build a community in a structured ecosystem and publish as well as improve the technical part of open data.

Interests to structure an ecosystem are multiple but one of the most important is the support it can bring to **encourage re-uses of open data**. Nowadays, we face the reality: **measuring the social and economic value of open data has become a necessity, but under some conditions**. However, we defend the fact that **measuring the value is a community responsibility** since only one actor will not be able to build a complete model. This community responsibility will go through the **improvement of interaction mechanisms** inside the ecosystem. We briefly explained three mechanisms: direct contribution, automated online survey, and a second nationwide open data event. Another initiative to organise the ecosystem could be to **appoint an independent organisation as federator of the different events and interactions inside the ecosystem**. The roles of the ecosystem could be to **spread awareness** in the sectors for which the level is not high enough, to **start the cultural shift towards openness**, and of course to help in the achievement of a measurement model. In our interview with Bart Hanssens (personal communication, May 24, 2016), the factors supporting successful structured feedback mechanisms were mostly related to the human skills and experience: to release data for which publishers know that re-users will be interested in interacting afterwards and to have the appropriate skills to answer the demands.

In order to reach an adequate level of openness in the whole ecosystem, in terms of quantity of datasets published or in terms of quality of the services proposed, **top management involvement is**

required. Their involvement could lead to long-term vision defined and respected, to bargaining power for the open data public agencies and to sustainability of the funding plans. Furthermore, public administrations should also **ensure minimum data management skills** in their services. Finally, **querying** is seen as an important **point of improvement for interoperability**, completing the publication of data by an increased capability of open data re-user to find the data they are searching.

The **right repartition of resources** is required between the two main axes of our recommendations (Building a community and Publishing open data as well as improving the technical part of open data). This repartition will depend on the specific characteristics of each organisation. This should not make us forget that the Belgian open data ecosystem is part of an international community dynamically evolving at the same time. It is therefore important to stay aware of the different approaches adopted to build an infrastructure and ecosystem supporting the creation of value through re-uses.

6.3 Limits and critical perspective

We differentiated the approaches to validate the observations by crosschecking but different limits arose from our methodology and the results we obtained. For what concerns the surveys, the difficulty to identify an appropriate number of re-users led to a relatively low number of answers. From there, no predictive analysis nor statistical significance of the descriptive analysis such as correlations and causalities were possible. The representativeness of the re-users in the surveys is uncertain when taking into account the low number of answers compared to the numerous characteristics (e.g. the different business models) and the uncertainty about the real number of re-users in Belgium for each characteristic. As a result, the observations we were able to detail can be used as indicators of the current status of open data in Belgium or potential factors and paths of development. But these observations will require further confirmation once the market will be more developed and stable.

Nonetheless, qualitative interviews implied different drawbacks as well: a limited sample or a choice among the potential profiles to be interviewed represented two limits that we faced. For the first one, we noticed the fact that all interviewees were experts of open data and consequently they were not representative of the diversity (neither the average) of opinions on certain open data topics. For example, when asking if the open data awareness was at a sufficient level in Belgium today, experts

answered typically yes. And obviously it represents a normal view of the Belgian ecosystem for people used to work constantly with others also active in the open data. For the second limit, we chose to interview amongst others the biggest regional public open data providers (Walloon and Flemish regions) and the most advanced local open data provider (Ghent). An exhaustive overview would have required to contact as well the Brussels region, the German-speaking community and other local communities active in open data. But as we already said, all the different categories of actors of a typical open data ecosystem were interviewed except for end-users of open data. It could be interesting in the future to conduct a research combining different views including end-user's ones.

From a critical point of view, two topics mainly drew our attention. If our study does not revolutionise quantitative research, for instance by providing significant evidence of the relation between social and economic benefits and open data, it definitely contributes to the expansion of the Belgian open data ecosystem and indirectly to a future ecosystem of a proper size and density to be the object of quantitative researches. It could also be used by quantitative research at a larger scale (European or global).

Expecting the low number of answers from re-users, we preferred a detailed and time-consuming questionnaire to a small and less time-consuming one. If we had to restart the same analysis, it could be a good idea to propose less multiple choices which complicated the analysis (less significant).

6.4 Guidelines for further research

The third and last objective of this work supports further research by describing different situations for which research questions arose.

Assuming that interactions mechanisms will improve and increase in density in the coming years, a further study should be able to collect enough quantitative data on the re-users' profiles, with the modelling of the link between innovation and open data as an objective. Examples of questions for a quantitative study could be:

- Why is B2B more attractive for open data re-users at the moment?
- What is the average time required for a company re-using open data to become economically viable? And from which stage does it start to hire new employees?

- What is the relation between re-users aggregating data and using consulting fees as business model?

Another axis of study could generalize the factors developed in this study about how to organise structured feedback mechanisms: what are the factors required to implement successful interaction between open data re-users and providers?

The third axis of study could focus on the relations between the effective development of the open data ecosystem in a country and the nations' maturity reports published by independent organisations.

7 Tables of figures and tables

Figure 1 - Assemblage of open data complementarities (Kuk & Davies, 2011)	7
Figure 2 - Map of the Belgian open data ecosystem*	12
Figure 3 - Map of Brussels area inside the Belgian open data ecosystem	13
Figure 4 - Representation of the structure of the Belgian open data ecosystem	14
Figure 5 - Number of employees working in the Belgian open data ecosystem in 2016 related to the starting year of open data re-uses	23
Figure 6 - Number of re-users per geographical source of open data and accessing open data at least through a data portal or not.....	25
Figure 7 – Example of feedback mechanism existing for Statbel.....	28
Figure 8 - Synthesis of our recommendations for the development of the Belgian open data ecosystem	38
Figure 9 – General interactions on the main axe of the data chain to measure the benefits of open data.....	39
Figure 10 – New interaction mechanisms on the main axe of the data chain to measure the benefits of open data.....	42
Table 1 - Repartition of the Belgian open data actors per region.....	15
Table 2 - Profile of the respondents using mostly 'Transport data' (Annex 9.1.3.2.1 - 9.1.3.2.7)	22
Table 3 - Communication mechanism used by Belgian open data actors depending on the objective pursued	27
Table 4 - Description and examples of the main innovative factors observed in the Belgian open data ecosystem	32
Table 5 - Communication mechanism which could be implemented by Belgian open data actors depending on the objective pursued	41

8 References

- Archer, P., Bargiotti, L., De Keyzer, M., Goedertier, S., Loutas, N., & Van Geel, F. (2014). *Report on high-value datasets from EU institutions*. Retrieved from http://ec.europa.eu/isa/documents/publications/report-on-high-value-datasets-from-eu-institutions_en.pdf
- Archer, P., Dekkers, M., Goedertier, S., & Loutas, N. (2013). *Study on business models for Linked Open Government Data*. Retrieved from http://ec.europa.eu/isa/documents/study-on-business-models-open-government_en.pdf
- Arrassi, Z. A., Deleu, W., Galasso, G., Goedertier, S., Lauk, M., & Vermeyleen, A. (2016). *Report on the current landscape of innovative, open and collaborative government in Europe and recommendations on an effective policy mix*. Retrieved from available in the coming months
- Belgium, O. K. (2012). "A world where knowledge creates power for the many, not the few.". Retrieved from <http://www.openknowledge.be/>
- Belgium, O. K. (2014). Open Belgium 2015. Retrieved from <http://2015.openbelgium.be/>
- Belgium, O. K. (2015). Open Belgium 2016. Retrieved from <http://2016.openbelgium.be/>
- Bertot, J. C., Gorham, U., Jaeger, P. T., Sarin, L. C., & Choi, H. (2014). Big data, open government and e-government: Issues, policies and recommendations. *Information Polity: The International Journal of Government & Democracy in the Information Age*, 19(1/2), 5-16. doi:10.3233/IP-140328
- Bertot, J. C., Jaeger, P. T., & Grimes, J. M. (2010). Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. *Government Information Quarterly*, 27(3), 264-271. doi:<http://dx.doi.org/10.1016/j.giq.2010.03.001>
- Carrara, W., Fischer, S., Oudkerk, F., van Steenbergen, E., & Tinholt, D. (2015). *Analytical Report 1: Digital Transformation and Open Data*. Retrieved from <http://www.europeandataportal.eu/fr/content/read-report-open-data-digital-transformation>
- Carrara, W., Fischer, S., & van Steenbergen, E. (2015). *Open Data Maturity in Europe 2015 Insights into the European state of play*. Retrieved from <https://www.fr.capgemini-consulting.com/publications/open-data-maturity-in-europe-2015>
- Ceschin, F. (2013). Critical factors for implementing and diffusing sustainable product-Service systems: insights from innovation studies and companies' experiences. *Journal of Cleaner Production*, 45, 74-88.
- Choenni, S., van Dijk, J., & Leeuw, F. (2010). Preserving privacy whilst integrating data: Applied to criminal justice. *Information Polity*, 15(1, 2), 125-138.
- cirb-cibg. (2015). Open data: Brussels open datastore. Retrieved from <http://opendatastore.brussels>
- Colpaert, P., Joye, S., Mechant, P., Mannens, E., & Van de Walle, R. (2013). *The 5 stars of open data portals*. Paper presented at the Proceedings of the 7th International Conference on Methodologies, Technologies and Tools Enabling e-Government.

- Colpaert, P., & Vanagt, T. (Producer). (2016). The state of Open Data in Belgium. [Introduction of Open Belgium 2016] Retrieved from <http://fr.slideshare.net/OpenKnowledgeBE/the-state-of-open-data-in-belgium>
- Commission, E. (2015a). *Creating Value through Open Data Study on the Impact of Re-use of Public Data Resources*. Luxembourg Retrieved from https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf.
- Commission, E. (2015b). *eGovernment in Belgium*. ISA Editorial Team Retrieved from https://joinup.ec.europa.eu/sites/default/files/egov_in_belgium_-_january_2015_-_v_17_00_final.pdf.
- Commission, E. (2015c). European legislation on reuse of public sector information. Retrieved from <https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>
- Cumul.io. (2016). Data Analysis for everyone. Retrieved from <https://cumul.io/>
- Dat. (2014). The Dat Project. Retrieved from <http://dat-data.com/>
- Data, W. O. (2011). Supporting you by structuring, enriching and publishing your data. Retrieved from <http://weopendata.com/>
- Davies, T. (2011). Open data: Infrastructures and ecosystems. *Open Data Research*.
- Dawes, S. S., & Helbig, N. (2010). Information strategies for open government: Challenges and prospects for deriving public value from government transparency *Electronic government* (pp. 50-60): Springer.
- Eurostat. (2015). Statistiques sur la recherche et l'innovation au niveau régional. *Annuaire régional d'Eurostat*. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Research_and_innovation_statistics_at_regional_level/fr#Emploi_dans_les_secteurs_de_haute_technologie
- Ferro, E., & Osella, M. (2013). *Eight Business Model Archetypes for PSI Re-Use*. London Retrieved from http://www.w3.org/2013/04/odw/odw13_submission_27.pdf.
- Flanders, G. o. (2015). *Open Data Manual Practice-oriented manual for the publication and management of Open Data using the Flemish Open Data Platform*. Retrieved from <http://www.vlaanderen.be/opendata>.
- Foundation, W. W. W., & development, O. f. (2013). Open Data Barometer. Retrieved from <http://opendatabarometer.org/>
- Fragments, L. D. (2013). Linked Data Fragments: Query the Web of data on Web-scale by moving intelligence from servers to clients. Retrieved from <http://linkeddatafragments.org/>
- Dekret zur Abänderung des Dekrets vom 18. Dezember 2006 über die Weiterverwendung öffentlicher Dokumente, (2015).
- Gent, S., Ghent, A. f., & Eskidoos. (2016). Creative Ring Challenge. Retrieved from <http://appsforghent.be/creativering>

- Glickman, Y. (Producer). (2016, 28/07/2016). European Data Portal runs Survey on the benefits of re-using Open Data. Retrieved from <https://www.linkedin.com/pulse/european-data-portal-runs-survey-benefits-re-using-open-yuri-glickman>
- government, F. (2015). Data.gov.be. Retrieved from <http://data.gov.be/>
- Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating Open Government Ecosystems: A Research and Development Agenda. *Future Internet*, 4(4), 900-928. doi:<http://dx.doi.org/10.3390/fi4040900>
- Hucq, B. (2016). *Digital strategy for Wallonia*. Paper presented at the Open Belgium 2016 - The state of Open Data in Belgium, Antwerp.
- Jaeger, P. T., & Bertot, J. C. (2010). Transparency and technological change: Ensuring equal and sustained public access to government information. *Government Information Quarterly*, 27(4), 371-376. doi:<http://dx.doi.org/10.1016/j.giq.2010.05.003>
- Janssen, K. (2011). *PSI in Belgium: a slow journey towards open data?* (2011/1).
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, 29, 258-268. doi: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. doi:10.1177/0894439314525902
- 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.
- Katsonis, M., & Botros, A. (2015). Digital Government: A Primer and Professional Perspectives. *Australian Journal of Public Administration*, 74(1), 42-52. doi:10.1111/1467-8500.12144
- Knowledge, O. (2014). Tracking the state of government open data. Retrieved from <http://index.okfn.org/>
- Kuk, G., & Davies, T. (2011). *The Roles of Agency and Artifacts in Assembling Open Data Complementarities*. Paper presented at the Thirty Second International Conference on Information Systems, Shanghai.
- Lassinantti, J., Bergvall-Kåreborn, B., & Ståhlbröst, A. (2014). Shaping Local Open Data Initiatives: Politics and Implications. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 17-33.
- Lejeune, C. (2013). *Analyse qualitative*. Paper presented at the LLSMD 2094 Séminaire d'accompagnement du mémoire (en sciences de gestion), Louvain-la-Neuve.
- lex.be. (2015). Search the whole Belgian legislation & jurisprudence. Retrieved from <https://lex.be/>
- Lindman, J. (2014). Similarities of Open Data and Open Source: Impacts on Business. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 59-70.
- Maglio, P. P., & Lim, C.-H. (2016). Innovation and Big Data in Smart Service Systems. *Journal of Innovation Management*, 4(1), 11-21.
- Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Almasi Doshi, E. (2013). Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute. *McKinsey Center for Government and McKinsey Business Technology Office*.

- Meijer, R., Conradie, P., & Choenni, S. (2014). Reconciling contradictions of open data regarding transparency, privacy, security and trust. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 32-44.
- Nazka. (2013). Wat doen we? Retrieved from http://nazka.be/skills_nl.html
- Numérique, A. d. Open data pour la Wallonie. Retrieved from <http://data.digitalwallonia.be/>
- Parsons, M. A., Godøy, Ø., LeDrew, E., De Bruin, T. F., Danis, B., Tomlinson, S., & Carlson, D. (2011). A conceptual framework for managing very diverse data for complex, interdisciplinary science. *Journal of Information Science*, 37(6), 555-569.
- Parycek, P., Höchtl, J., & Ginner, M. (2014). Open Government Data Implementation Evaluation. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 80-99.
- Petychakis, M., Vasileiou, O., Georgis, C., Mouzakis, S., & Psarras, J. (2013). *A State-of-the-Art Analysis of the Current Public Data Landscape from a Functional, Semantic and Technical Perspective*. Journal of Theoretical and Applied Electronic Commerce Research.
- Pollock, R. (2011). Building the (Open) Data Ecosystem. Retrieved from <https://blog.okfn.org/2011/03/31/building-the-open-data-ecosystem/>
- Pop, D., Zeleti, F. A., Beer, E., Canova, L., Dekkers, M., Harding, C., . . . Winstanley, P. (2016). *Standards for Open Data and Public Sector Information*. Retrieved from available in the coming months
- PwC. (2015). *Open Data Challenges Series*. Retrieved from <https://www.pwc.co.uk/assets/pdf/nesta-and-the-open-data-institute-pwc-report-october-2015.pdf>
- Rohunen, A., Markkula, J., Heikkilä, M., & Heikkilä, J. (2014). Open traffic data for future service innovation: Addressing the privacy challenges of driving data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 71-89.
- Sandoval, R. (2011). The two door perspective: An assessment framework for open government. *JeDEM- eJournal of eDemocracy and Open Government*, 3(2), 166-181.
- Scassa, T. (2014). Privacy and Open Government. *Future Internet*, 6(2), 397-413. doi:<http://dx.doi.org/10.3390/fi6020397>
- Ubaldi, B. (2013). Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives. *OECD Working Papers on Public Governance*, 22.
- Vanlshout, Z. (2015). [BeST-Address].
- Verhulst, S., & Young, A. (2016). Open Data Impact: When demand and supply meet: GovLab and Omidyar Network.
- Vlaanderen, D. I. (2015). Open data in Vlaanderen. Retrieved from <http://opendata.vlaanderen.be/>
- Zuiderwijk, A., Gascó, M., Parycek, P., & Janssen, M. (2014). Special Issue on Transparency and Open Data Policies: Guest Editors' Introduction. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 9-I,II,III,IV,V,VI,VII,VIII,IX.
- 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.

- Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. *Information Polity: The International Journal of Government & Democracy in the Information Age*, 19(1/2), 17-33. doi:10.3233/IP-140329
- Zuiderwijk, A., Janssen, M., Meijer, R., Choenni, S., Charalabidis, Y., & Jeffery, K. (2012). Issues and Guiding Principles for Opening Governmental Judicial Research Data. In H. Scholl, M. Janssen, M. Wimmer, C. Moe, & L. Flak (Eds.), *Electronic Government* (Vol. 7443, pp. 90-101): Springer Berlin Heidelberg.