

## Annex 2

### 1 Interviews

#### 1.1 Interview guide

- A. Project framework
  
- B. Objectives of the interviews:
  - a. Identify the actors of the Belgian open data ecosystem
  
  - b. Characterise the interactions between actors/infrastructure/ecosystem and how it works
  
  - c. Precise the relation between innovation and open data

##### 1.1.1 First part: interviewee's perspective without influence of our previous work.

- A. Definition of open data ecosystem
  
- B. Structure of an open data ecosystem
  
- C. What is the general state of the Belgian open data ecosystem from your perspective?
  
- D. Position/role of the interviewee's organisation in the Belgian open data ecosystem
  
- E. Internal structure of the open data section in the organisation?
  
- F. What is your view over the maturity of the Belgian open data ecosystem?
  
- G. What are the main challenges faced by your organisation? (resources, harvesting, providing data quality, standardisation, interoperability, ...)
  
- H. What are the main benefits you assessed thanks to open data initiatives?

##### 1.1.2 Second part: interviewee's perspective and confirmation of observation

- A. Current state of our research

In your opinion, what are the potential benefits that an open data ecosystem (and open data in general) can have on innovation?

- B. In the literature, potential benefits from the ecosystem to support re-uses of open data and innovation

Do you see those benefits in Belgium?

- C. Is there an adequacy between the objectives pursued by Belgian open data initiators and the actions undertaken?

## **1.2 List and profile of persons interviewed**

### **1.2.1 Noël Van Herreweghe**

Noël Van Herreweghe works at the 'departement of Information' depending on the Flemish Government. He is the manager of the open data initiative in the Flemish region.

### **1.2.2 Bart Hanssens**

Bart Hanssens is an open standards and open data expert who works at Fedict, the Belgian Federal Public Service for ICT.

### **1.2.3 Yves Peeters**

Yves Peeters has been the general manager of GeoLives for eight years. "GeoLives offers multiple solutions for improving the distribution of topographic maps, digital terrain models, aerial imagery and city maps, to the general public for outdoor and touristic activities. Our mobile solutions help outdoor enthusiasts to determine their position via GPS on different devices, and offer relevant Points Of Interest linked to cultural information, photos and even audio." (Peeters)

[www.citytrail.com](http://www.citytrail.com)

[www.cityzen.net](http://www.cityzen.net)

[www.citytour.com](http://www.citytour.com)

#### **1.2.4 Thimo Thoeye**

“Thimo Thoeye studied software engineering at Ghent University and now works for the City of Ghent, which allows him to combine his love for his city with his passion for data and information management. He is a strong open data advocate both within the local government and beyond.” (Belgium, 2014)

#### **1.2.5 Jan Vansteenlandt**

Jan Vansteenlandt is the co-founder of We Open Data, a company that works on open data projects and supports open source data tools (Data, 2011). He is one of the main contributors of the open data platform ‘The Datatank’. He is a member of Open Knowledge Belgium Foundation. (Belgium, 2015).

[www.weopendata.com](http://www.weopendata.com)

#### **1.2.6 Ides Bauwens**

Ides Bauwens is the co-founder of ‘nazka mapps’ and ‘nostalgeo’. ‘nazka’ makes smart mapps and location-based API's. They distinguish themselves by developing intelligent, interactive & user friendly maps and services that convey sustainability. (Nazka, 2013). Ides Bauwens focus for his part on the marketing evolution on short and long-term as well as on the financial part of the company (Bauwens).

[www.nazka.be](http://www.nazka.be)

#### **1.2.7 Toon Vanagt**

“Toon Vanagt is an entrepreneur with broad experience in all aspects of setting up and managing ventures in the Benelux. He has hands-on experience with e-business, consulting, big data and ICT projects.” (Belgium, 2015).

#### **1.2.8 Pieter Colpaert**

“Currently Pieter Colpaert is working on a PhD on linked open transport data at the MultiMedia Lab at iMinds. To put it simply, he is researching how you can be better informed about your travel and daily commuting experience.

Before that, he has co-founded an umbrella organisation for open knowledge: a chapter of Open Knowledge called Open Knowledge Belgium. Open knowledge gathers terms like opengov, open science, open access, open source, and so on. While the philosophy has been around for a while, they believe it can be useful in a lot of other areas as well, leading to more innovation and better products in a world where everyone and everything is connected. He is chairman of the board of Open Knowledge Belgium and is leading the Open Transport working group.

In February 2012, he started FlatTurtle together with Yeri Tiete and Christophe Petitjean. With FlatTurtle they are building a product that will redefine digital signage and digital pinboards as we know it today by visualizing and interlinking public and internal data.” (Belgium, 2014).

### **1.2.9 Sven De Meyere**

Sven De Meyere works at Realo. “Realo is an online real estate marketplace [with as their] mission to help people to buy, sell or rent a home by providing them insights about houses, neighbourhoods and service providers such as real estate agents.” (De Meyere, 2016).

[www.realo.be](http://www.realo.be)

### **1.2.10 Stéphane Vince**

Stéphane Vince works for the Numerical Agency of the Walloon government (AND) as information & technology manager. (Doucet, 2016)

## **1.3 Synthesis of interviews**

### **1.3.1 Noël Van Herreweghe – Informatie voor Vlaanderen, Flemish Government – 23/05/2016**

#### **1.3.1.1 First part: interviewee’s perspective without influence of our previous work.**

##### *A. Definition of open data ecosystem*

*“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)*

*When applying the concept of ecosystem to the open data field, the literature describes it like:*

*“A multi-level and multidimensional entity where raw material, as far as distribution and developing are concerned, is the target of cooperation.”(Anneke Zuiderwijk, Marijn Janssen, et al., 2014)*

The concept of ‘ecosystem’ gives some insights about the way you look at it: actors, parties dependent on each other.

*B. What is the general state of the Belgian open data ecosystem from your perspective?*

Flemish government: different ways for open data:

- strategic level, open data is the norm, they have support of many parties, short-term and long-term planning,
- legislative level (new decree: ‘everybody has to open-up data for re-use’), licenses,
- an open data platform (about 4000 datasets), a knowledge platform, created a handbook which is translated in seven languages in different countries in EU.

Still a lot to do: more datasets, knowing which dataset should be open or not, what is the planning for opening nationally etc. They have currently a poc on LOGD for example with CRAB (Centralized Registry of Address in Belgium).

Federal government is catching up, they were really active for the last months. They will have a legislation in the coming months normally, not sure about the licenses at the federal level.

Local governments are quite under-developed in term of number involved in open data. Two cities are part of the ecosystem: Gent and Antwerp. Bruges is starting, Ostend has put some efforts but has been quiet for the last few months, Kortrijk as well.

New hype touching the open data field: Big data and IoT.

There is still this deception that some potential data providers want to keep the data into one big dataset behind walls of protection.

*C. Position/role of Flemish government in the Belgian open data ecosystem?*

That is why the Flemish government is trying to establish for 2017 'open' data as a strategic instrument for the Flemish and local governments.

But there is incertitude regarding if open data will not take a back seat in the next coming years due to costs savings from the public sector. In any case, the budget is going to be limited even if it is the government agreement (decree etc).

*D. Internal structure of the open data section: organisational structure?*

Two people. If compared with the UK or US where they have 300-400 people with independence from the political world (money from government and private institutions)

*E. What is your view over the maturity of the Belgian open data ecosystem?*

Cultural reasons play a big part to explain the difference of evolution between open data ecosystems in English speaking countries and others. Noël Van Herreweghe defends that we will never reach the same proportion or level of ecosystem for this reason. Openness is about culture, about being transparent. This is also the reason why the Flemish government opens data without willingness to measure what is made with it. There is no registry, and it is not really in the role of the government to do this.

We could think of a feedback loop from the demand-side but on a voluntary basis, not from something constructed by the public sector. And this sector will not tell potential re-users what they could do with open data.

*F. Vision on short and long-term (5-10 years) over the evolution of the Belgian ecosystem and the position/role of the Flemish government?*

*G. What are the main challenges faced by your organisation? (resources, harvesting, providing data quality, standardisation, interoperability, ...)*

Challenges are a bit spread in the other answers. And these challenges justify why an ecosystem could be useful to solve some of them. With ecosystems, we can go further than just saying that it is a way to support the creation of an informal dialog between actors. We can say that it is “the only way we are going to create an open data ecosystem with all actors aware of the value for the whole community and for themselves of the data.”

*H. What are the main benefits you assessed thanks to open data initiatives for different levels (politicians, management, you, help potential re-users)? (costs reduction, social value, economic value)*

He started personally with open data projects for the transparency effect. Example of last ‘hackathon’ for which he has fought to have a funding from the Flemish government of twice 20,000€. And at the moment the event has only one start-up registered to the event. He should have never organised this. Because the consequence is of course that the management above him is waiting for results and this is a perfect justification for them to stop financing.

Transparency is not part of the culture in Flanders. Everything is sensitive. One example: data about fine dust in Flanders are known and available for everybody. But there was a guy who was willing to link it to topology data. This person reported to Noël that by doing this, people told him that the value of houses around Zaventem airport was going to drop which was obvious but this by essence what means transparency: give the choice to people who would like to buy a house there and are sensitive to such information.

Why open data is not used more or just more accepted? It is a combination of different reasons: it takes a long time for many re-users to launch and make their initiative viable. As a result of insufficient resources (lack of people, initial and long-term investments etc.), awareness is not spread sufficiently around potential actors. Open data is a young concept as a worldwide initiative. We don’t know if it is economically viable. “Show me a business plan or a couple of them like Realo or data.be which are viable. We have 4000 and more datasets, but most of it

are 'no data' in terms of statistics behind. 'Do data' and real-time data are very little." The last reason is a political one: if we are not able to propose integrated data in a bigger scale. "If I want to build an app, there is nothing I can do with Flemish data only."

The solution is at the management level. We need cooperation from the middle management of each entity to propose an integrated solution for open data inside the Belgian ecosystem. The problem in Belgium is the political impact on every high management levels which is part of the culture of the country. And for political parties, the main question for supporting an action or not is whether the party can shine or not in the media. We need to put them around the table, in every kind of event where discussion can take place and on every important subject like: open data concept itself, number of portals, support to each other, etc.

Therefore, two factors are important to take into account:

1. Top management must not interfere
2. Actors of the ecosystem must organise round tables and opportunities to discuss and establish a dialog

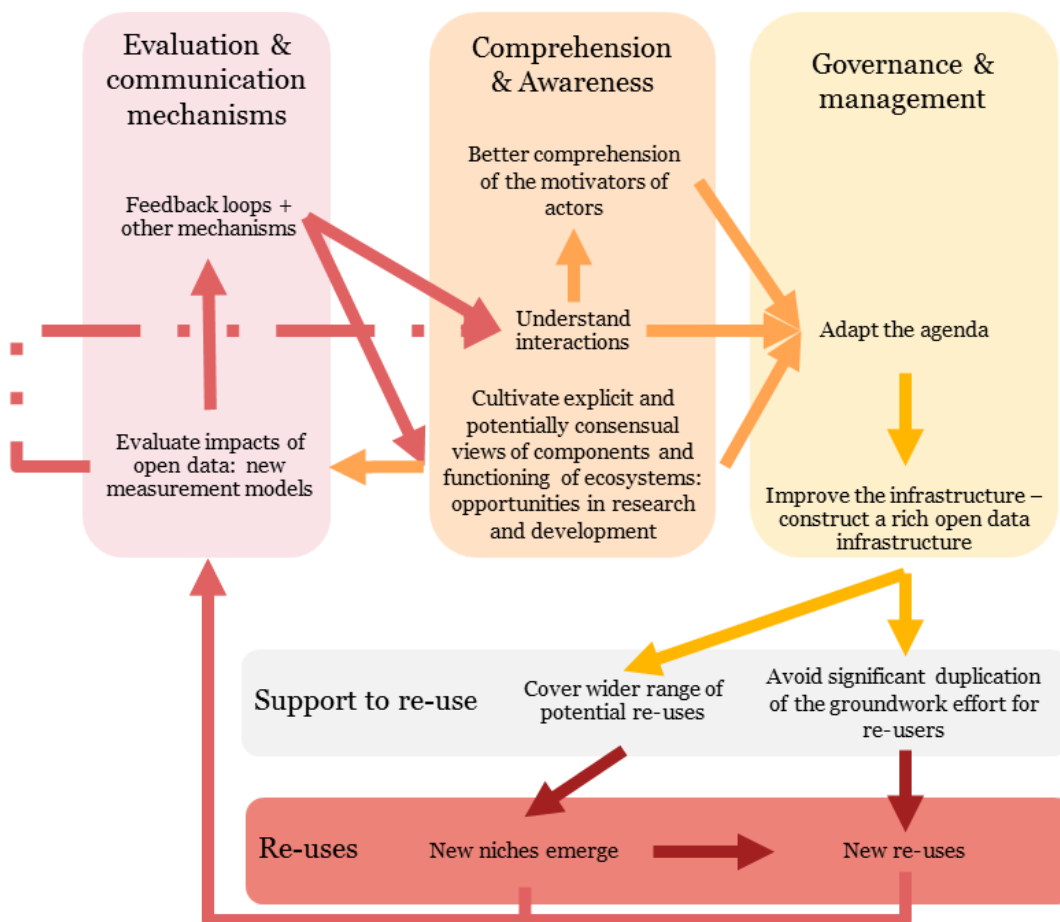
The Flemish government is supporting the second point at the moment with 5 round tables this year. They started in 2015 with 5 other, just to construct a first discussion. This year it is different. They would like to know what are the priorities from the demand side. Because data providers think they are doing the right thing but are they really? And the demand side needs to realise what are the difficulties faced by data providers. Feedback mechanisms are a necessity but the Flemish government won't take the initiative by constraining re-users. More awareness is needed in the demand side that data providers require some informal feedbacks to keep going further in the initiatives (simply in terms of funding). And in complement of this, potential re-users need data collectors (private or public companies) to share their data. Example of a potential re-user who wanted to build services for companies based on the average energy consumption. But without access to data resources from energy companies, it is impossible.

During a conference in the Royal Academy of Sciences in Brussels, Noël Van Herreweghe was asked by a member of the Nairobi university if data related to particular diseases were released in Belgium. The answer was no and the reaction from the requester was to say that Belgium was committing murders. Because data on these diseases could be really useful to fight and prevent their apparitions in other countries. But institutions in Belgium are not opening it because they are expecting to make money from it.

About access to data, networks around universities are also a good example: between 25 and 30 million euro are paid yearly towards investment companies. Why is that so? First, publishing articles is a necessity for any researcher. In the case of universities, the institutions have to pay sometimes up to 12,000€ for one article to be published in an important editor like Springer, Elsevier etc. which are investment companies, not publishing companies. Even worse, we have to pay as re-users for articles that our own universities have created. And the big majority of actors are not saying a word against this system. “Bernard Rentier is a shining example for the rest of universities, but do we see any other university following his movement? No.”

#### **1.3.1.2 Second part: interviewee’s perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



*1. Have you implemented or plan to implement those mechanisms?*

It is happening (in the digital agenda of De Croo for example) but it has not been implemented at a large scale. The framework to start it is there: in terms of tools installed or actors existing in the ecosystem.

If there is a feedback mechanism from re-users, it should be voluntary. Putting the emphasis on feedback is detrimental for the open data movement in general.

But feedback in the sense of a dialog should be initiated in order to know and understand the needs of actors. Although it is really hard to get it in a semi-automatic way, for example by constructing awareness first. People don't take the time and don't know what their needs are if they have to say it from their own on a computer. At the moment, the Flemish government is

organising round tables on an invitation basis. And for this year version, they asked last year participants to give them propositions of actors who are interested in participating. In conclusion, the dialog is increasing at the rhythm of a word of mouth. Why not organising in Belgium the 24hours of open data with discussions between actors? But government is probably not going to make the initiative of this ecosystem.

The number of data portals is not a problem until standardisation is achieved between the portals. But it goes further than this, to discuss the methodology and strategy

*J. Could the new hypes be of some help to build services on top of open data?*

“Yes, I’m thinking about mobility, tourism, health, transport. If we have smart ships telling at what speed, where and by which direction they are going etc. we are going to be able to do much better regarding the whole mobility problem. But as a citizen I have to agree to give my data. And there you have to differentiate people who does not want to see even a photography on the internet but at the same time, they will share until their intimacy during a trip.”

The role of Europe is also ambiguous since privacy and open data are both supported and sometimes clashing.

*K. How to support adequately innovation?*

“Hackathons are cool, students are interested in participating and so on, but nothing has never came up from hackathons. It stays with the concept.” And participants are not doing it with the idea of developing a business plan.

Governments should not interfere with the roles of private companies by proposing services from the data because it would be making tax-payers support again. This is what is done in the academic system:

1. Tax-payers support first the costs of education itself
2. They are indirectly paying for the publications of universities

3. They are also indirectly paying to have an access on the publications
4. They pay as a consumer when using services of spin-offs based on data and researches.

Data providers should open immediately data-as-a-service but they can't for various reasons: data privacy or risky data (defence of the country, terrorism) but it represents a minority. Most of the data should be made open. But do data providers have the resources to accomplish it? No, they need more time. Moreover, top management is waiting for proves of some results with open data. And this is the knot of the problem: data providers should be opening up their data for a transparency purpose. But in these times of economic savings, politicians have only one idea in mind: 'what is in it for me?'. This way of thinking is very difficult to change.

### **1.3.2 Bart Hanssens – Fedict – 24/05/2016**

#### **1.3.2.1 First part: interviewee's perspective without influence of our previous work.**

##### *A. Definition of open data ecosystem*

*"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)*

*When applying the concept of ecosystem to the open data field, the literature describes it like:*

*"A multi-level and multidimensional entity where raw material, as far as distribution and developing are concerned, is the target of cooperation." (Anneke Zuiderwijk, Marijn Janssen, et al., 2014)*

Open data: Machine readable, open licensed, free for re-use included for commercial re-use (specified for data providers who thinks opening can be limited to non-commercial use only).

Open data ecosystem: more than technology, about people, about getting a community working, getting people from different domains working together, need for some

standardisation, harmonisation in the way data are published or made available. Something that is probably lacking at the moment.

*B. What is the general state of the Belgian open data ecosystem from your perspective?*

Still at the beginning. There is data available in quantity. But it is difficult for data providers to see much re-uses of it. Sometimes, providers get responses: 'well the data made available at the moment don't correspond to the one we would like to have'. But at the same time it is difficult to get feedback from the community itself to say what they would like to see, even in Flanders probably. What is done by the community, how providers could improve. It remains still vague.

Concerning enablers, the main actors are Jan Vansteenlandt with the DataTank and WeOpenData, data.be from Toon Vanagt, Paul Hermans from ProXML, Noël Van Herreweghe from the Flemish government, Thimo Thoeve from the City of Gent. So it is more about key people than key companies in Belgium.

Sometimes difficult to integrate the data provided in the Belgian ecosystem between the different providers (some are more advanced than others). Consequently, it must be hard for re-users to use it easily. And we are only talking about the national level and not cross border.

*C. Position/role of Fedict in the Belgian open data ecosystem*

Fedict, for what concerns open data, is quite ok. The development of the portal is over even if it could be of course still improved. Right now, they are trying to get some proof of concept out of the door. Get some administrations and start-ups (smaller or not) to work with them to make some kind of concept to show that it is not too difficult to make something out of open data. Difficulties to convince people to release their data and to find the right way how to make the poc. In terms of finding the right people and in terms of the procedures they have to follow (tender procedure).

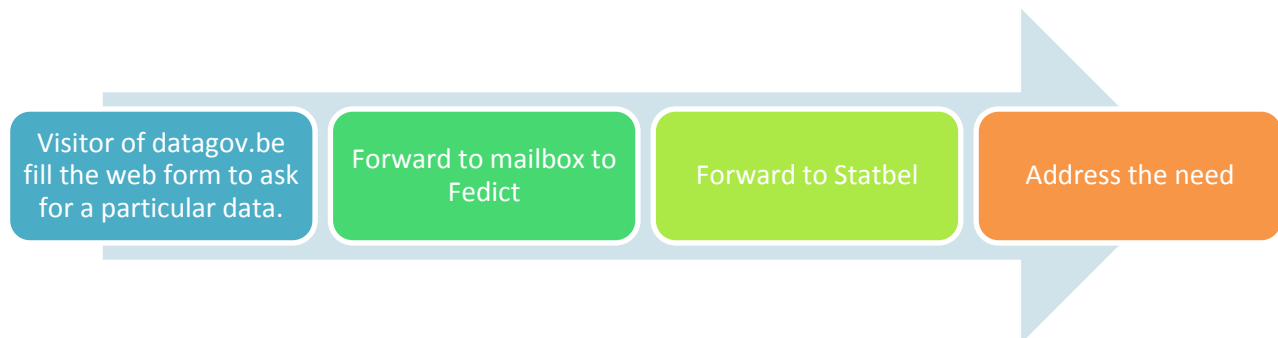
Interactions and contact are existing between local levels, federal level and regions. There are quite informal. Contacts as well with different actors like Raf Buyle, working at Informatie

Vlaanderen and involved at the moment on an OSLO project: Linked Open Data Standards for Local Administrations. They see what they could do at the federal level with the findings.

Interactions are informal. Some administrations have excellent feedback loops like Statbel who are active since 20 years in data analytics and are very professional, with a 'can do' attitude. They are really used to interact with people and deal with their demands. Also in the geo world (even if sometimes need to pay certain fees) the sector is doing quite well in term of metadata etc if you have questions. They have good relations with Open Knowledge, trying at the moment to participate to the 'Open Summer of Code'. Contacts also with regions but sometimes it is still new for certain administrations to have people who ask them to release data they are collecting for internal use in a machine readable format.

Still many excuses not to open data. Things that make an administration opening is often due to 'word-of-mouth'.

Example of feedback loop existing with Fedict and Statbel:



In this example, the main problem is not in the system requiring a transfer by e-mails and more human actions but the main problem is to address the need and open the dataset for which the party providing the feedback is looking.

#### *D. What factors make Statbel working properly?*

They have lots of statistics needed in many different types of researches. Ex: information about population, about average income, about the number of quores, lots of data that can be reused in other application. They don't have many 'stand-alone' data. They have a lot of experience

about how to anonymized data, how to publish data and get feedback from it. They have a good team of people skilled. They know what people would like to have as data. They know what people could expect from a specific dataset and they would be able to do with it.

*E. Internal structure of the open data section (in Fedict): organisational structure*

Federal level is redesigning or joining the smaller departments including Fedict. As a consequence, the way it will be organised in the short term is unclear at the moment. But Fedict and the open data remain important as part of the digital agenda.

*F. What are the main challenges faced by your organisation? (resources, harvesting, providing data quality, standardisation, interoperability, ...)*

For awareness, there is two sides: the people who want the data and the service who provides the data. Some administrations are still afraid of releasing the data like 20-30 years ago with the first websites. Some put forward that the quality is not high enough but are collecting the data for their internal uses which is certainly not better than if they open it. In any case, there is no datasets without any mistake (even very high quality datasets). But it still has to answer to a minimum level of quality. The goal is not to publish everything without paying attention to the quality.

The people who are into data are already aware of data or open data in general. For example if you look at larger companies for instance professional publishers who re-use laws or official publications that have been published for 20 years (Moniteur Belge). They were already aware and using open data before the concept was widely spread. At that level, awareness is not a problem.

The problem is more about where to find the data, even for data providers.

The quantity of data published is not always a good yards to measure the re-use. It is the same with all the apps existing on smartphone, most of them are not used by a lot of people. It is a reality in geographical data or public transports, there is a huge demand for re-uses. But at the same time, these are also the datasets which are hard to get. So pushing for more datasets is

maybe not a good way to increase re-use unless you make also available those really 'core' datasets.

LOGD can be really useful. There is a discussion whether the government should do it internally or not. He could just publish the data and another party would link it. The government has some services into place in any case (outside open data projects). But if one would like all the statistics as linked open data, it would probably not be the government in charge because there is no emergency in the demand for this at the moment. What is certain is that governments must publish their data at least in a downloadable format. "It is very strange. At one point, people from the open data community were screaming: 'Yeah you should have an API, you should create a web service'. And then, a few years later: 'Yeah but an API is not the way, you should publish your data in a machine-readable format so that we can make our own. Or you should instead of SPARQL, you should use data fragments'. My guess would be that, if it is feasible of course, you should create a dump of your data. And if you want to create a web service on top of it, that is fine, but I'm not sure if it is part of the tasks of the government."

Synthesis:

- The main challenge is to make the different governments releasing the same kind of data or using the same set of specification class standards.
- The second challenge would be getting some re-uses, some concrete collaboration with the whole community for making the re-use actually work. It is still really vague and unstructured until now. Difficult for data providers to see what they are building with data or what they would like to see the providers to improve.
- The third is releasing the core (high-value) datasets like public transport data or geographical.

G. *What are the main benefits you assessed thanks to open data initiatives for different levels (politicians, management, you, help potential re-users)? (costs reduction, social value, economic value)*

Fedict would like to know what re-users are doing with open data, mainly for the management, in order to be able to show them some benefits. Not in a purpose of control ('tell me first what you are going to do and then I might release it') but more in an interest way. Concretely, some good examples of re-uses could be really helpful for data providers to justify the investments and the continuity. At the moment, even if they ask for such feedbacks during events or through social media, saying that they could provide some exposure for these re-users, it is still not happening.

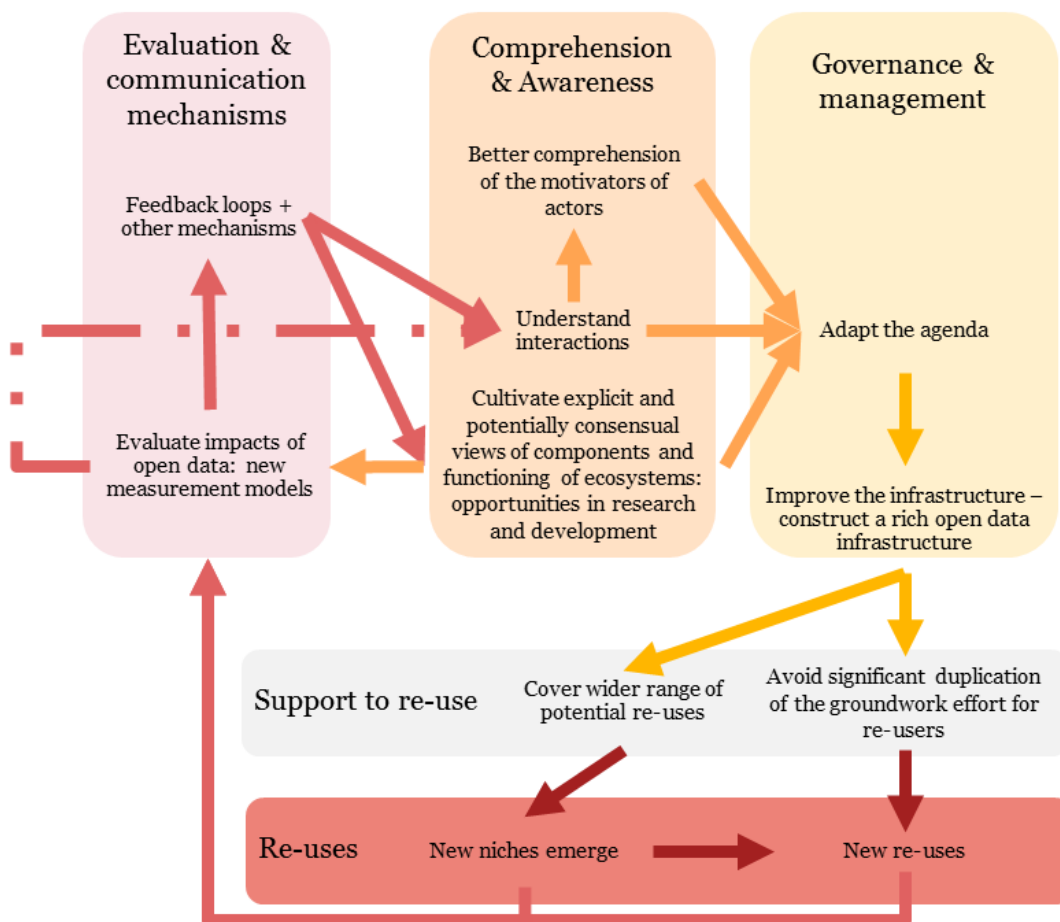
It has to come from the re-users. It is time for them to give an informal return regarding the last years when it was a one-way relation: data providers were asked to publish more data without any confirmation. "If they could provide two or three really nice examples I would be happy. But in my opinion, it does not have to be the data community or the re-users who have to make some efforts or some statements that say 'ok we are re-using your data and look what application we have built'."

No idea how they should improve their support to increase this informal return (links, good examples).

At the moment there is no financial benefits, it is more about creating awareness for different actors like the data collector himself.

### **1.3.2.2 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



*H. Have you implemented or plan to implement those mechanisms?*

It has not yet been implemented, at least formally. The feedback mechanism and other mechanisms are a bit ad hoc, it is not really built into their system or into the organisations. This is definitely something that has to be improved.

Current action is to go to the different departments and make them aware of what has been done in the open data.

In term of governance and management, a future action could be to put the open data and data management in general as part of the respective management agenda at more prominent places than currently. How to manage data, how to keep track of the metadata and new

evolutions of data (who owns the data, the license applicable and so on). Something that has to be more emphasized at the management level.

Concerning organisation of hackathons or support events for re-users, Fedict would like to participate if they find the money to the 'Open Summer of Code'. When questioned about the mitigated results from hackathons and the apparently good results from other formula like in the UK with ODI and Nesta, Bart Hanssens said that such kind of events have been organised in Belgium as well, like in Gent with <http://www.appsforghent.be/creativering> or to a lesser extent <http://www.geopunt.be/actualiteit/2016/maart/doe-mee-aan-de-wedstrijd-rond-open-data> .

Hackathons were nice to create awareness a few years ago. A difficult part for the new type of challenge organised is to find enough participants willing to invest their time for such project. Maybe because it is seen as too risky or not interesting enough or too narrow because focusing only on open data when start-ups would like to link it to IoT or other domains. So a good project on paper but difficult in the reality to put into practice.

Even if it is an opportunity for them to receive a fully integrated support. Could be a good feedback to receive from start-ups, to know why they have or have not participate to such events.

### **1.3.3 Yves Peeters – GeoLives – 24/05/2016**

#### **1.3.3.1 First part: interviewee's perspective without influence of our previous work.**

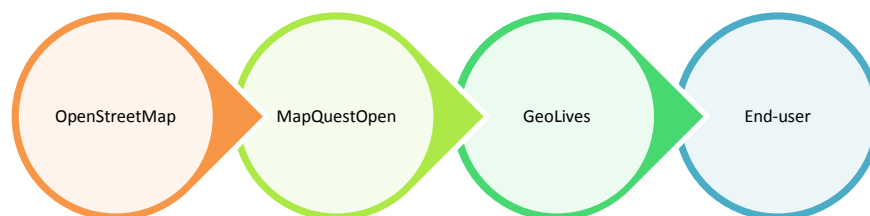
*A. What is the general state of the Belgian open data ecosystem from your perspective?*

He feels that there is potential but at the same time, Belgium is laying behind in the open data movement. For instance, he saw recently that all topographic maps in Austria had been made available when it is not the case in Belgium. Another point he has noticed in this example was the rapidity with which the Austrian organisation was asking them more information about their re-use and to provide some visibility on their website for this re-use with a sustainable link on it.

*B. Position/role of GeoLives in the Belgian open data ecosystem*

Company developing services on top of geolocalisation data, first and historically from a partnership with the IGN (NIG) and second from open data, mainly from websites like

OpenStreetMap or Wikipedia. Hereunder is one example of data-chain for open data in the case of GeoLives.



**Interactions with actors of the ecosystem:** what are the dependencies for services providers; what are the services proposed to actors; what are the relations with actors with the same profile, what are the relations with potential entrants inside the ecosystem; What metrics or measures do you think are the most adapted to assess the interactions between actors inside the ecosystem.

Contacts with other actors are mostly indirect.

*C. How would you like this ecosystem to evolve in the newt coming years?*

Development of open data is interesting but not halfway. He has the impression that some open data initiators are willing to open because open data is a hype.

In his opinion, in the process of supporting open data re-uses, open data initiators should consider the whole process from developing a business to the generation of feedback. In their case:

- 1) Business development of the re-user. Trying to reach the critical mass of consumers (clients)
- 2) Critical mass is reached. Start to develop statistics on the end-user utilisation. (use Google analytics but limit of number of applications for free) Develop

added-value services. Determine to what consumers/partners are paying attention

- 3) General statistics are available to share and transparent for feedback to other companies (like OSM) or data providers. For example, could provide information to OSM about the use made by consumers.

*D. What are the main motivators for you to start open data projects?*

Motivators: Historically associated to national partners collecting data which have an added-value for consumers (confidence).

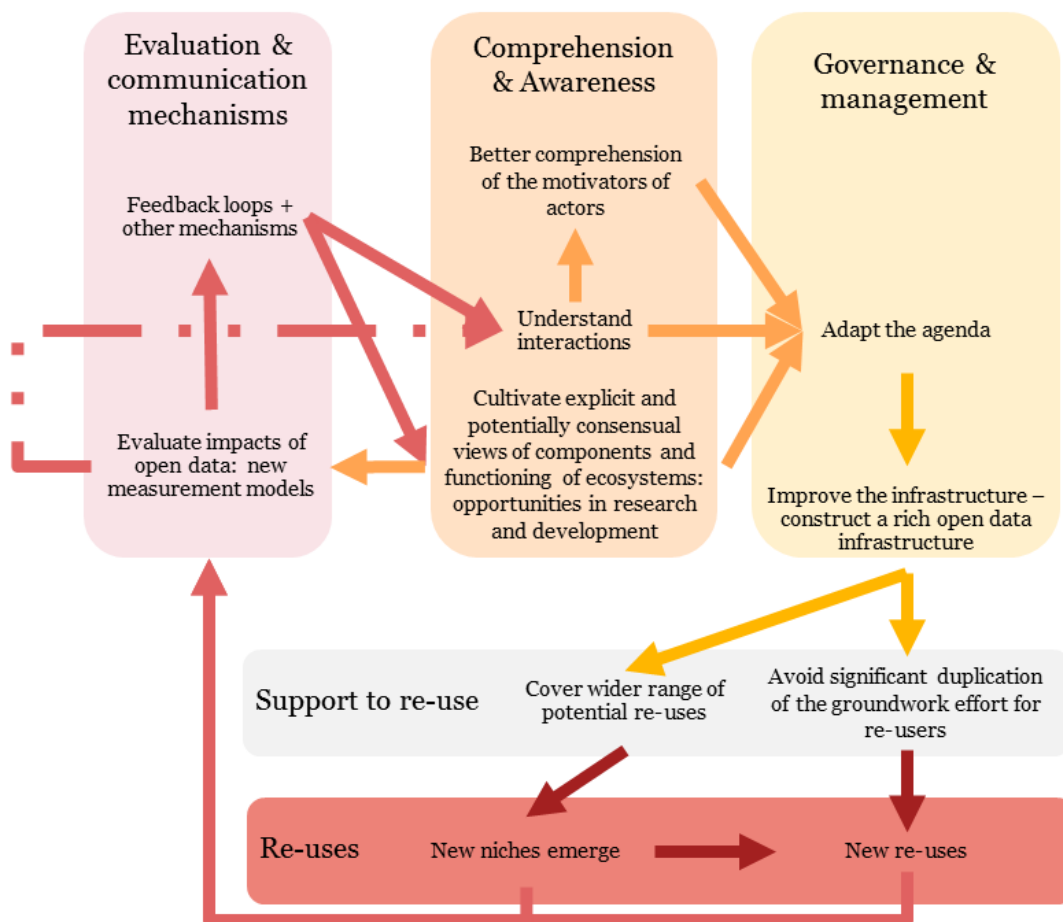
Open data have no boarder. It allows to move to one integrated product around the world.

*E. What are the main benefits you assessed thanks to open data? (costs reduction, social value, economic value)*

Not huge since they are coming from a partnership (and so open data is not their main focus at the moment) but it could change in the future to adapt to the market. If free enabler services (MapQuest for example) start to ask for royalties or if the IGN is becoming completely open. But it is not threatening their business as a whole. Their competitive advantage is based on their client-base and all the services constructed on top of raw maps. One possibility in the future could be to develop their own infrastructure around the maps internally. Importance to be sure of the sustainability of the providers (data or services) and importance of the measurability for the pricing strategy of providers of services (B2B). For example if MapQuestOpen asks tomorrow a price determined by the number of clicks made by the final consumer, it will become difficult for them to accept keeping working with MapQuest. At least until when the final consumer would be ready to accept to pay at the level of his consumption (similar to communication pricing strategies – subscription + add price if too high consumption).

**1.3.3.1.1 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



*F. Have you seen those mechanisms implemented in Belgium?*

At the local level, GeoLives has implemented a feedback mechanism between consumers, tourism offices, territory services and finally local services. Consumers report problems encountered during the walk to tourism offices who gather the information at the territory services level which spread the information to the responsible local service.

More related to our subject, GeoLives has feedbacks by e-mails and through a forum. E-mails are there to feed an ideas box. And the forum is available for any problem with the services provided. Later on, if the scale changes, the ideas box could be transformed into punctual (or even systematic) survey for instance ten propositions of the next implementation. The forum could be partially replaced by a new website with an online automatic helpdesk (similar to Microsoft etc). And the forum itself would be kept for concrete and unsolved problems.

*G. If not, would it be interesting if they were implemented? And would you be ready to participate to this dialog between actors or not and why?*

Interesting for him to have a national cover. To see round tables organised, forum or tools to establish a discussion between actors of the ecosystem.

*H. You told in the survey that you were accessing data by e-mail or API, do you sometimes use Open data portals? If not why? Would you like to access all the data with one tool?*

A centralised access would be helpful of course. But at the moment, data are not opened enough through the existing way.

*I. What are the main factors on which innovation is dependent in the open data domain? Or in other words, lots of data are available now from public entities, what, in your opinion, is missing to allow more innovation to appear?*

Skills to develop or to learn, critical mass, no interference from the public sector (transparency, no favouritism), favour competition (no volatile promises from politicians).

*J. What the government should know about challenges faced by re-users?*

It is certain that the number of interlocutors is a big challenge. Ex: Product 'Sityzen', too many interlocutors plus influence between local governments, tourism offices, provinces, region etc.

If the metadata could be standardised and being implemented in a large area (mass market).

Some data could be interesting: Natura 2000, private public areas, forest areas, land use map, urban plan, urban areas.

### **1.3.4 Thimo Thoeys from the City of Ghent – 25/05/2016**

#### **1.3.4.1 First part: interviewee's perspective without influence of our previous work.**

Definition of open data ecosystem

*"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)*

When applying the concept of ecosystem to the open data field, the literature describes it like:

*“A multi-level and multidimensional entity where raw material, as far as distribution and developing are concerned, is the target of cooperation.” (Anneke Zuiderwijk, Marijn Janssen, et al., 2014)*

*What is the general state of the Belgian open data ecosystem from your perspective?*

One thing that Thimo Thoeys is witnessing currently is the **place of consulting companies** making profit out of open data. Actually, he said that today those companies are making most of the value assessed to open data, not by directly re-using open data but by enabling an ecosystem to evolve, standardising data and services, constructing linked data etc.

*Position/role of City of Gent in the Belgian open data ecosystem*

The city of Gent can be seen as a **data provider** but also as a **community builder**. They invest a lot of money in organising events, workshops, mailing list...

In Gent, they have the opportunity to receive many feedbacks from re-users of open data. Any re-user of open data will automatically get in touch with the data owner and also get visibility through the portal or by other manners thanks to the city. From the point of view of Thimo Thoeys, about 25% of the investment in an open data initiative for the public sector is technical and 75% of the investment is on building a community. Next to the time and resources invested in building the local ecosystem, at each event, they remind the participants that they are really willing to receive feedbacks.

So why is it not working with Fedict when they do the same (reminding participants of the importance of their feedbacks)?

A reason advanced could be that they should ask companies active at the national level since they are providing data at this level. And for companies working at this level, it can be really difficult to get in touch with the appropriate agency.

A second reason is that local initiatives represent the first contact with re-users most of the time, it is then easier for local initiatives to be aware and to give visibility those re-users.

For Thimo Thoeve, the City of Gent was able to construct a proper local open data ecosystem.

*What are the main challenges faced by your organisation? (resources, harvesting, providing data quality, standardisation, interoperability, ...)*

- Lots of money will be invest in **anonymising data**, lots of academic as well as other actors will be involved in this process. It is needed because not only non-personally related open data need to be released.
- **Aggregating data wide**
- **Data ownership**
- **Linked Open Data**
- As well, it is very hard for the City of Gent as a data provider to **convince the management** of the importance of opening up data for a transparency purpose. They would be easy to convince if financial returns were easy and quick to demonstrate. Fortunately in Gent, a number of managers from different levels, including top management, strongly agree that transparency is an important goal as a public administration.
- The **European level initiative** provides also a **legal base for local governments** to build open initiatives.
- For him, the **5-star model** of open data was a big **mistake**, not in technical terms because it is right but in strategic view. With Linked Open Data, you limit the range of re-

users who can use it and you also limit the time during which data providers can open the data. He thinks that governments have been advised to linked data even if the benefits in the short term are really unsure.

- For re-users, the **type of format** for open data is important but **depends on the type of re-use made from it and the domain of data**: some will need real-time data and API when others will need aggregate data or data standardised at a wider level than region.
- When looking for an economic impact, it is of course interesting to provide additional services on real-time open data but it implies that if you are a business dependent on real-time data provided by government, you need to have a **service level agile** for data providers which is often inexistent.

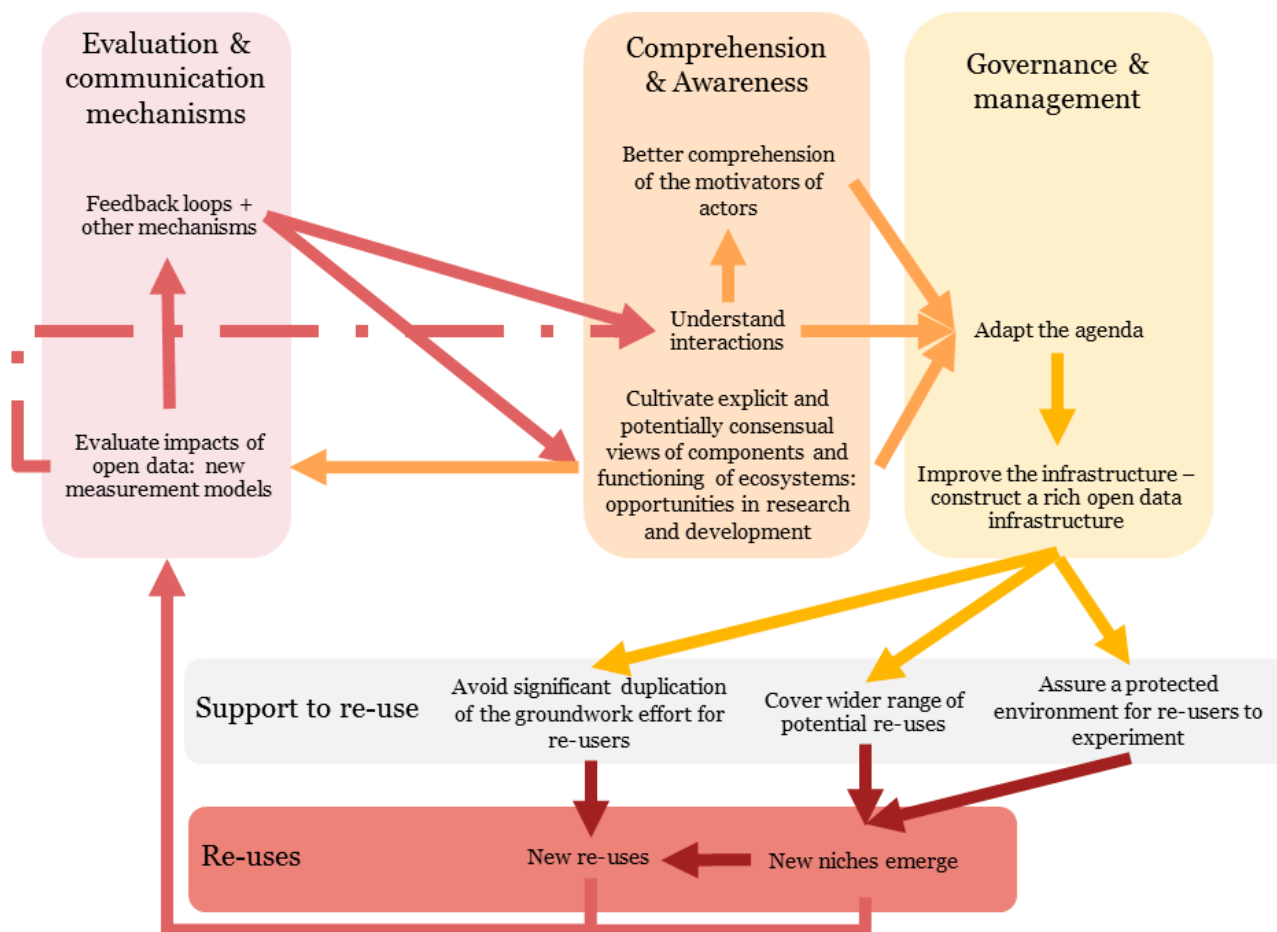
*What are the main benefits you assessed thanks to open data initiatives for different levels (politicians, management, you, help potential re-users)? (costs reduction, social value, economic value)*

From his perspective, he thinks that it is a mistake to try to quantify or measure the benefits from open data re-uses with economic value or costs savings.

There is a logical gain in transparency through the opening up movement from data providers which is an important aspect of modern democracy even if it costs more for the public sector.

#### **1.3.4.2 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



*Have you implemented or plan to implement those mechanisms?*

Timo Thoele: The big arrow from Re-uses to Evaluation & communication mechanisms is mainly informal. There are two main **types of feedbacks**:

1. One oriented towards the **evaluation of government open data policies in terms of technical feasibility and relevance**. Measured by: number of errors in datasets, format used, requirements from developers, services present online. Open Data Barometer and Open Knowledge Index are doing this for example.
2. The other more oriented to **economic outcomes** from re-users to data providers. This one is informal.

### **1.3.5 Jan Vansteenlandt – WeOpenData – 25/05/2016**

#### **1.3.5.1 First part: interviewee's perspective without influence of our previous work.**

*A. Could I consider you as an infomediary?*

Intermediary definitely in the public side and enabler so yes. WeOpenData consolidates entities that want to publish data through a number of software packages.

*B. How would you like this ecosystem to evolve in the next coming years?*

Something lacking today are common types of data provided on a larger scale than cities and regions. For example in Antwerp for public transport data: the city is publishing the transport locations and the distances between them. Data that are really interesting but limited to Antwerp area. So it is really hard to re-use it.

For Jan Vansteenlandt, it is more a government task right now. Governments need to talk much more to each other to agree on certain types of datasets, at least on core/high-value datasets in order to have more impacts on a larger scale. To agree on certain vocabularies is important as well or create groups like the 'Open Tourism' group. In this, linked data could a good support.

*C. What are the main motivators for you to start open data projects?*

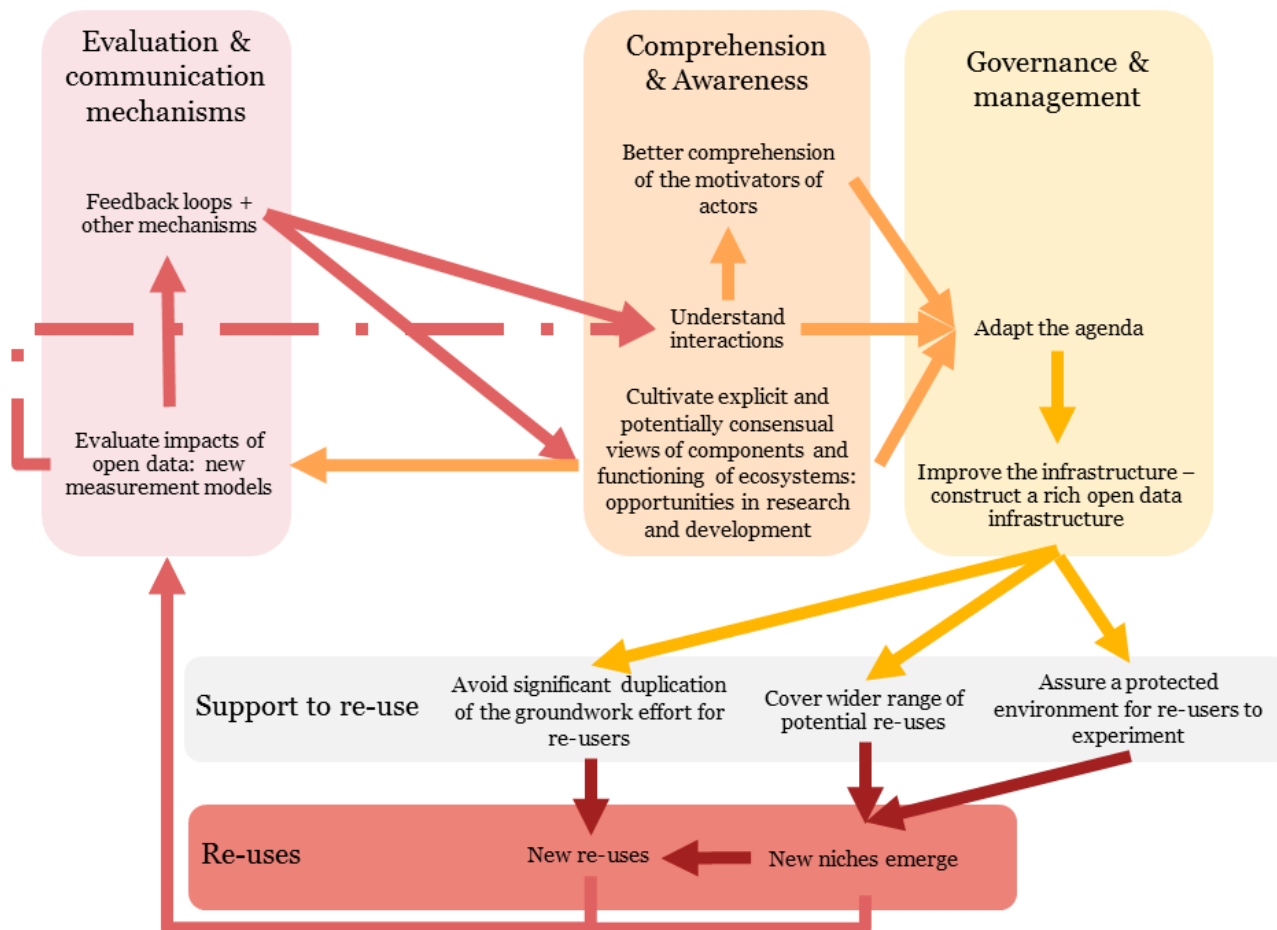
For him, main motivators to get involved in open data was the transparency concretise by the open spending initiative. Something that is really hard to achieve. In Belgium we are still not at a fully transparent point.

*D. What are the main benefits you assessed thanks to open data? (costs reduction, social value, economic value)*

Personal benefits (doing what he likes) and transparency purposes thanks to open data are the two main benefits for him.

#### **1.3.5.1.1 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



*E. Have you seen those mechanisms implemented in Belgium?*

For him, the ecosystem is what is happening at the moment: actors start discussing, people get more aware of what open data is.

About feedbacks of open data re-users to data providers: from public re-users, they don't provide any feedback because they are not paying attention to it, for re-users more generally it is sometimes because they don't know that they are using open data, and finally for private re-users, they don't want others to see or get information about what they are using as data. It is why events like Open Belgium and others are important to give the opportunity to various actors to meet and see each other.

*F. For your activity, do complementarities emerged easily or not? Are active efforts required? (Complementarities are human or machine artifacts composing the process from data to service innovation).*

For his case, it depends on the data sources, some are easy and some are not. The most difficult part is building the architecture design behind the service proposed.

*G. What are the main factors on which innovation is dependent in the open data domain? Or in other words, lots of data are available now from public entities, what, in your opinion, is missing to allow more innovation to appear?*

Today it is mostly use case driven. Once you have a use case and a business process or service and that open data are adding something to your business, you will integrate it because it is there and for free. So if you don't use it to add a value on your business, someone else will do it and obtain a better service than you.

See where the cost benefits are.

*H. What do you think about hackathons?*

*New forms of hackathons like in Gent: <http://www.appsforghent.be/creativering>*

For longer term hackathons, a focus or restriction on a specific framework is needed as well as really good quality datasets available (not perfect but useful). Why not thinking about releasing new high quality data only for one long term hackathon to see what the potential outcomes are.

*I. Even if it is a clear opportunity for them to receive a fully integrated support, why are they or not participating to such events?*

In his opinion, potential re-users are not participating because for the data that are today available, private entities making money out of it are already there. And it is difficult to get data that have a quick beneficial value for private entities.

In the place of government, he would choose one particular focus, contact all the different entities that he knows are working around this, publishes the data related to the topic, even for the event itself if needed and finally see what they can come up with.

*J. Who are for you the enablers of an open data ecosystem in Belgium at the moment?*

For foreign companies, quite hard to write software for other countries because the internal working and the governance of data is completely different most of the time. Modules are therefore not really transferable.

CKAN can be considered as an enabler.

### **1.3.6 Ides Bauwens – Nazka Mapps – 27/05/2016**

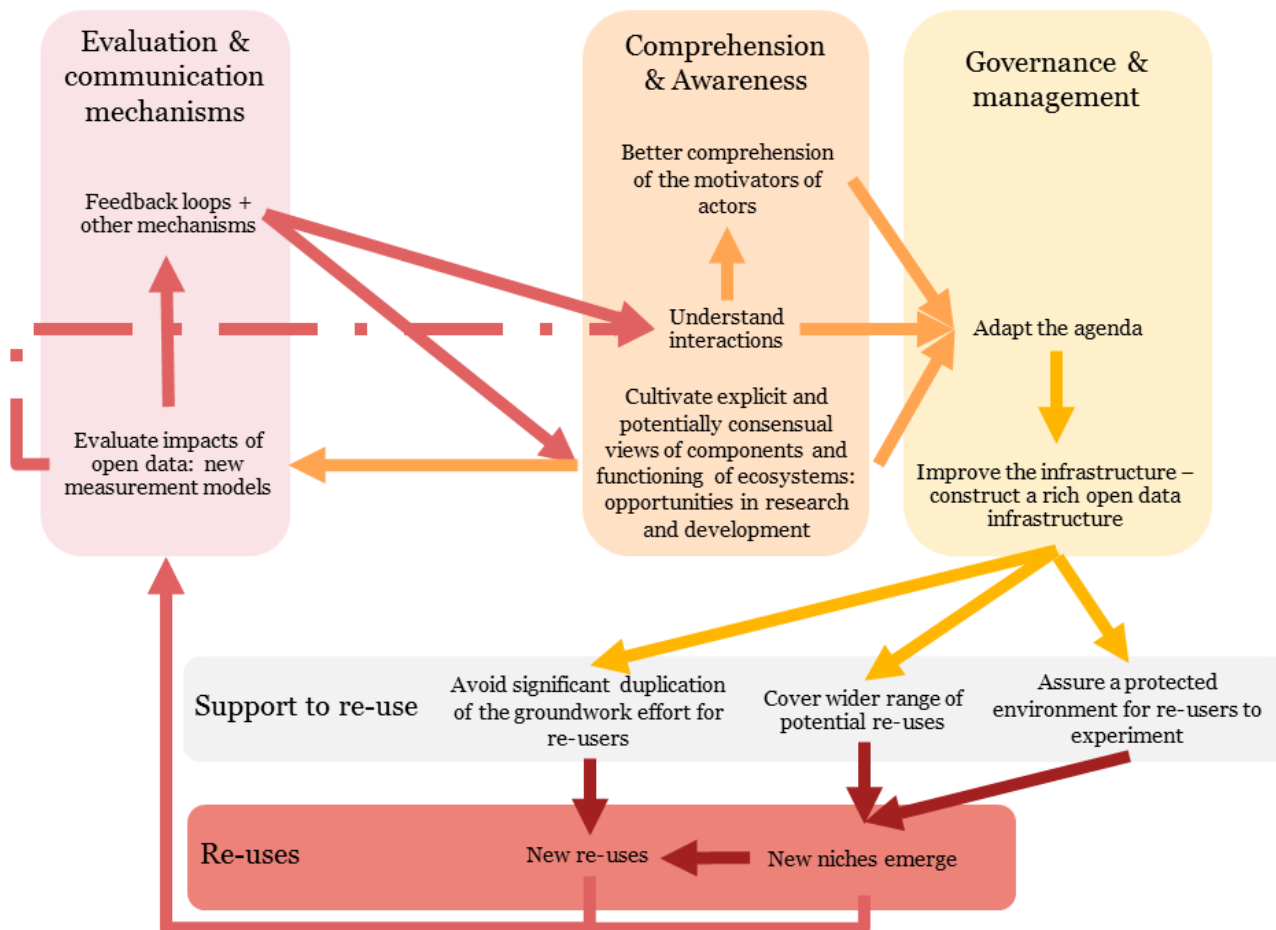
#### **1.3.6.1 First part: interviewee's perspective without influence of our previous work.**

*A. Position/role of Nazka maps in the Belgian open data ecosystem?*

**Interactions with actors of the ecosystem:** what are the dependencies for services providers; what are the services proposed to actors; what are the relations with actors with the same profile, what are the relations with potential entrants inside the ecosystem; What metrics or measures do you think are the most adapted to assess the interactions between actors inside the ecosystem.

#### **1.3.6.1.1 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



B. *Have you seen those mechanisms implemented in Belgium?*

*If not, would it be interesting if they were implemented? And would you be ready to participate to this dialog between actors or not and why?*

In Flanders, he has the impression that a dialog is already established. For example: AGIV is really active, in the geographical world as well (Geopunt).

<https://www.agiv.be/>

<https://www.agiv.be/opleidingen/trefdagen/trefdag-informatie-vlaanderen-2016>

<https://www.agiv.be/producten/geopunt>

<http://www.geopunt.be/>

<https://overheid.vlaanderen.be/magda-platform-toegang-tot-ondernemingsgegevens>

<http://www.vlaamseinfolijn.be/>

And if he agrees that data accessibility is important, he is not a supporter of the establishment of a dialog between different regions as a first objective since for him the Flemish government is ahead of the others in the opening up process and creating this discussion would mean for Flanders to duplicate efforts that have already been undertaken (attending round tables on topics that have already been discussed for example).

*C. What do you think about hackathons?*

*New forms of hackathons like in Gent: <http://www.appsforghent.be/creativering>*

Even if it is a clear opportunity for them to receive a fully integrated support, why are they or not participating to such events?

From Ides Bauwens' perspective, data providers need to come with concrete problematics (1 or 2) and quality data available. From there, companies could be interested to participate.

*D. What are the main factors on which innovation is dependent in the open data domain?*

Data accessibility and quality.

### **1.3.7 Toon Vanagt – Data.be and Lex.be – 30/05/2016**

#### **1.3.7.1 First part: interviewee's perspective without influence of our previous work.**

*A. What is the general state of the Belgian open data ecosystem from your perspective?*

Toon Vanagt is part of the digital agenda working group at federal level. Moving in the good direction, catching up compared to various neighbouring countries (the Netherlands, France, and UK) mostly when it comes to making business with open data by making the right datasets open for example. Legislative level is almost up to date, everything is at the Parliament now. Working on the license to make everything clear. It is one point that is still difficult for re-users.

Because data providers still need to be convinced that opening up data is important. Lots of excuses in the administrations, the main ones are privacy and security related. And those are tough ones to overcome.

Example of 3D data in Flanders: it has been opened but under a license specifying not for commercial re-use. "And when you put that sort of disclaimer hidden somewhere in the license agreement, then it is of course open data but it is not really open data." Reasons behind this choice are economical. Running a GIS system is expensive and open data are just a by-product that the states need to run anyway. 3D data have then a huge cost and potential re-uses on top of it are wider than just open data. What exists in other countries is government collecting the 3D data and private parties pay to use it and adapt the data if they change something in the reality (so they support as well the maintenance cost). So in a fully open system, private parties will no longer have to pay. A typical trade-off could be to put certain layers only for private uses (paid) and public layers are open.

"So there is still work to be done in the exact license terms. There are still departments that want their special cases because they are special and because they still don't embrace fully the concept although I see lots of people inside agencies who are supporting the cause but on the other hand there is the opposite movement for example with departments that used to make money on top of their own datasets and for them is it really tough to give this up because by opening up those datasets they are losing these revenues. So there is a conflict of interests for them. But when you look at lower government entities like cities and regions, you see many enthusiasts who are active in the open data field for years."

*B. Could we consider you as an enabler?*

Yes, at least he hopes so. When looking at Data.be for example, it is using three main different sources, some are informally open (ex: Moniteur Belge/Staatsblad is not providing data as structured data so it is not formally open), others formally.

What Data.be is doing: "We do some heavy lifting ourselves, by going fetching a copy every morning. We enriched it, we actually have to all see yards, do entity extraction to enrich what

would like the entities and the metadata we need to extract and to use data afterwards. And then we mashed that together with for example the Kruispuntbank van Ondernemingen (KBO/BCE) which is an open dataset, the National Bank of Belgium from where we have got all the balance sheets of companies. And so, when we are talking about the 'Staatsblad', then we are really talking about the law. Everybody in Belgium is supposed to know it, it is public information. It was open ages before the PSI directive etc. but it is not yet in an open data format (structured). So you will not be able to find it in a typical data portal. What Data.be is offering is then an easy search engine for anybody to look at information mashed from these three datasets on every company in Belgium.

Revenue model: freemium with lots of success (11000 daily visits on the website), and then advanced tools. For example, a subscription fee (20-100€ per month) for full text search inside pdf document, or analytics and consulting fees for tools, link detection, that they offer mainly for the financial industry.

*C. How would you like this ecosystem to evolve in the next coming years?*

For data privacy: many tools to anonymised data, to make sure you cannot extract who was the person, entity or where were the data created. One technology available to ensure identity protection. One example in the same sector of Lex.be, in Belgium very few cases law are released when in other countries it is different. In the Netherlands, every law decision made by the government is fully published in an anonymised format and it is the core itself that anonymised within a few days after the decision has been taken because they have a software that supports them. The software takes all names, addresses and other identifiers elements and then a human verify the cleaning up process of the software. There is not much work because the whole document is already cleaned out. So it is a way to do this, even about touchy data as here in the example. In Belgium, we are laying behind because we don't give the right IT tools to people and agencies. So we have to be conscious that at the same time technology can help but we have to be also aware that it is possible to deanonymised if you have datasets large enough you don't need that many points to find somebody back. In a more general perspective, privacy as a concept is culturally defined and time dependent. Another example Toon gives sometimes

is if you ask people how much they have bought their house, and if they are ready to share this information, they answer 'oh no, that's really private information'. In this case, Realo says that they are seeing a shift since a few months. But so, the cultural influence is high because in the Netherlands or Sweden, these types of datasets are released and it is all right because people are confident with what governments are doing from it. In Belgium, there is a clear tendency for the less governments know, the better.

So, open data and privacy:

1. Important to shield the information
2. Keep in mind that some datasets can be deanonymised
3. Have a discussion: what do we want as society today and in the near future. If we look at people disclosing information on social networks etc, they clearly don't really care about some private data to be published.

*D. How could the discussions about the cultural shift take place?*

Showing to people who have those concerns, those arguments not to open up, first what has been done in neighbourhood countries, and to provoke public debate about the value behind this, not only for companies but also for society as a whole.

*E. Is it important to assess the re-uses made on top of open data?*

It is important from a government's perspective to show good examples of open data that have been re-used and where successful in the re-uses or had a big impact. So at some point you have to evaluate all the data portals and all the people who are opening datasets to know how it has been used.

But Toon Vanagt is against the fact that you have to disclose information and ask for permission to be able to get the dataset. For example in the case of Lex.be, they provide the data sources even if they are not legally obliged to do so. And Toon Vanagt is quite sure that the data sources are quite aware of this and even more are able to follow the traffic on Lex.be's website. In

consequence of what they certainly have an idea of who are interested in the data they are producing (ex: le Moniteur Belge) and with which data people are interested to see it combined.

It is very hard to predict which data will be useful to third parties. There, it is a matter of the re-user to show his creativity and provide a feedback saying 'hey I have combined these three datasets and I have been able to construct this'.

Often one of the required step before completely opening up internal data is to allow the data owner to have a license agreement more restrictive than for completely open data. So that the data owner keeps partially the control over what is made from their data etc. This happens even more when people are not sure about what could be built from the data released.

But let's not forget that this is the first step in today's open data world. A world where data are mostly published by government organisations. To what we are going is a world in which data are released also by other organisations like private parties or non-profit organisations. For example with utilities companies who could make aggregate statistics or data on many things like energy use, locations of solar panels or wind turbines etc.

But at the moment, actors are still busy with opening governments.

*F. What are the main motivators for you to start open data projects?*

Opportunity. He was naïve when thinking that it was going to take two years for everybody to embrace the open data movement. It took longer than expected but there is still a huge opportunity.

Shared knowledge and social value with economic benefits as well.

Early days, impacts and influence to create a market, interesting for an entrepreneur.

*G. What are the main benefits you assessed thanks to open data? (costs reduction, social value, economic value)*

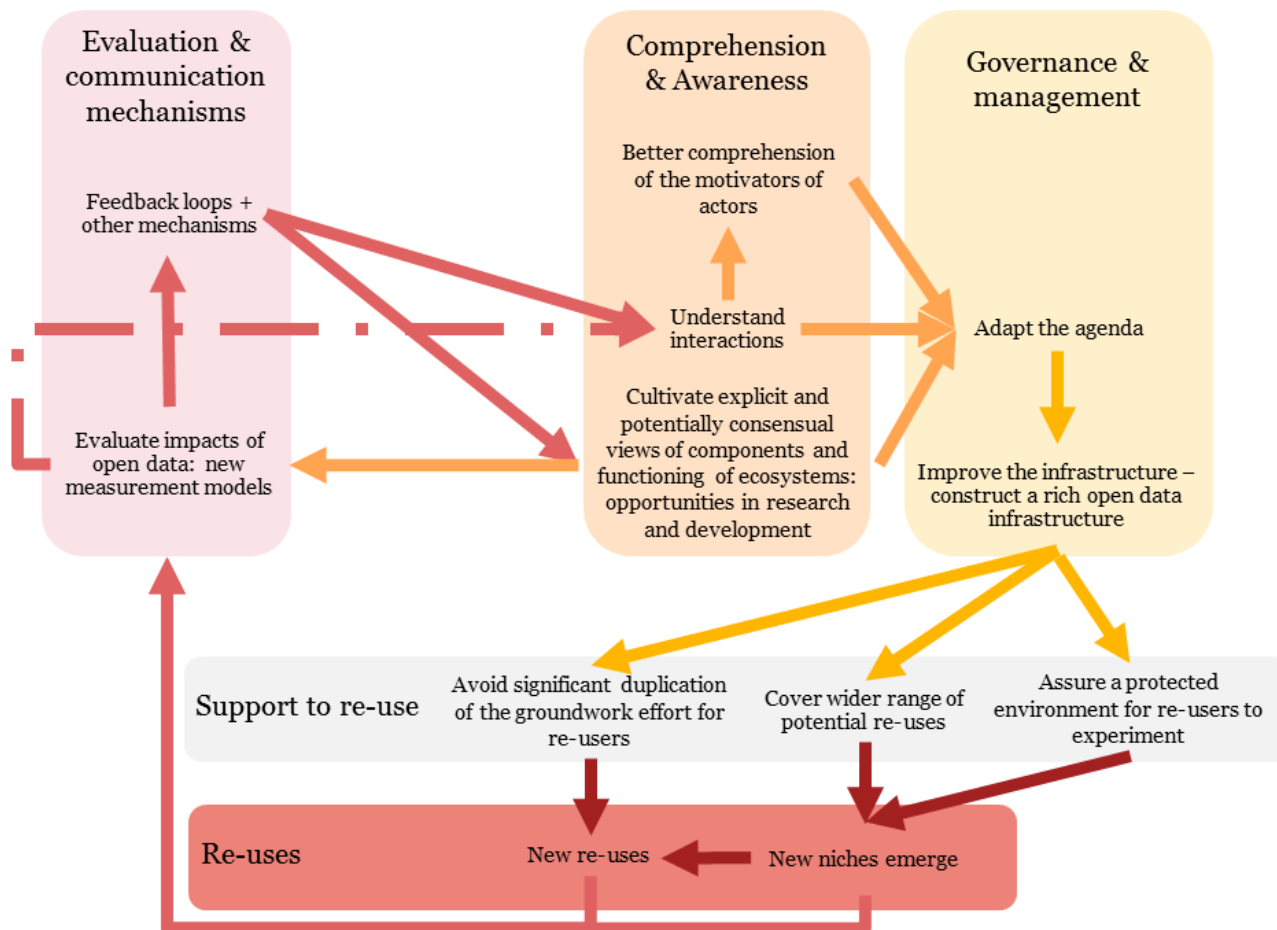
Toon Vanagt is seeing a lot of re-uses. But it is not always clear for a side based upon those datasets. For example, big companies like Google or other are proposing new and more accurate services and products like better routing path for taking public transports, implying huge changes at the moment at the societal level. But those are changes that we don't really see because those datasets have been made available by the SNCB/NMBS.

For Toon Vanagt, the real value behind open data is laying in the mash up of different data, opened and closed data. Where consulting companies have a role to play. We have lots of insights to gain there.

There is a huge opportunity for consultants, developers and IT experts to propose services to companies and end-users who don't have insights about open data: how to integrate open data as an add-on to their current internal and external applications and services.

#### **1.3.7.2 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



#### H. Have you seen those mechanisms implemented in Belgium?

It is the next step, to provide structured feedback loops. For example, for Lex.be and Data.be when they detect errors, they provide 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. So it is really informal. But this feedback should be provided directly to the data providers and not the re-user of open data as it is the case right now. And that's what Toon is meaning by a structured feedback: an easy way for end-users and re-users to give feedback to the data collector (for example a button on which they can push) it must be easy as well for the data collector to assess quickly if it is relevant or not.

But, from his perspective which is an entrepreneurial one, round tables and other events organised to bring actors together are not really needed because when he needs something, he

uses existing communication channels on his own to get answers. On this, he says himself that he has probably a skewed view since he is involved in many open data initiatives (re-user, enabler and provider sides). He has a privilege access and for him, annual events organised by each data provider are sufficient.

On data standardisation, he says that as long as there is no leading standards, he prefers that governments concentrate on releasing data at the best quality they can but with their own standards, even if it is different than other levels of governments. Why? Simply because a re-user who has built something on top of different providers of open data and has been able to integrate the different standards in one application does not want to see all standards changing all the time. What is needed is documentation available on the standard used and data that are available with the same standard.

Concerning the strategic view and discussions between different data providers to release same domains of data at the same time to support more efficiently re-uses in terms of scale: for him it is not a good idea simply because it will release the pressure from the ones that have already released this domain etc. He prefers a data provider who is releasing datasets and says that they will update every three months than one who is waiting six months for another data provider to release. So the first point of importance is to release data.

*I. So, is awareness at a good level in Belgium?*

Yes, definitely.

*J. What are the main factors on which innovation is dependent in the open data domain? Or in other words, lots of data are available now from public entities, what, in your opinion, is missing to allow more innovation to appear?*

The mash-up of closed and open data. For example to improve the back office or your customers experience. And there are new business models behind these services.

*K. What do you think about hackathons?*

*New forms of hackathons like in Gent: <http://www.appsforghent.be/creativering>*

*Even if it is a clear opportunity for them to receive a fully integrated support, why are they or not participating to such events?*

Those initiatives are still needed to give an extra push to potential re-users. It works because it gives a certain deadline and a feeling of urge to deliver something. They just need to be clear about which datasets can be re-used during and after the event.

*L. I would like to have your opinion on the findings of the survey, would it be possible for you to give me your impressions on it?*

Yes.

### **1.3.8 Pieter Colpaert – Open Knowledge – 30/05/2016**

#### **1.3.8.1 First part: interviewee's perspective without influence of our previous work.**

*A. What is the general state of the Belgian open data ecosystem from your perspective?*

<http://www.slideshare.net/OpenKnowledgeBE/the-state-of-open-data-in-belgium>

*B. Position/role of Open Knowledge in the Belgian open data ecosystem*

Open Knowledge is an enabler of the ecosystem. Encouraging the development of an open data ecosystem at different levels and with multiple initiatives. For example with different working groups (iRail, Open Tourism, OpenStreetMap, etc.).

Example of activity of iRail recently, partnership with TreinTramBus (Flemish federation of public transport users). This summer with OpenSummerofCodes, both would like to develop an application to predict overcrowded trains. Link:

Example: <http://datanews.levif.be/ict/actualite/traintrambus-et-irail-veulent-prevoir-les-trains-bondes/article-normal-490935.html>

*C. How would you like this ecosystem to evolve in the next coming years?*

Cost to start re-using is way too high. Too many interoperability issues. Pieter Colpaert is a supporter of Linked Open Data and in consequence of the work done by PwC with ISA Square etc. He thinks that interoperability promoted by ISA Square is the future of open data.

At the question whether a data provider should start immediately opening up data as a service, Pieter Colpaert answers that if the 5-star of open data are still applying, is already happy (to a certain extent) with pdf data as long as data providers publish data in an open license because it is already useful and good for transparency. Linked data needs more education to people. It is not really a problem of money to make linked data. He has already made some proof of concept for linked data.

With the 5-star of open data, organisations need to be resources-oriented and not services-oriented which is not the case for most of the organisations at the moment. For example, people say 'you can ask your question on my silo' which is not the right manner to think with open data stars. They should think first about the resources for which they are responsible. And it is a totally different way of working. On this point, organisations need to be educated. It is a difficult management shift.

*D. What are the main motivators for you to start open data projects?*

Pieter Colpaert has got suit by the Belgian railway company. It started from his frustration that no application with real-time data from railway was available on mobiles. Motivation from the frustration of the ignorance of the Belgian railway company.

*E. What are the main benefits you assessed thanks to open data? (costs reduction, social value, economic value)*

Really difficult to answer quickly. Many different parts. Analogy with science-fiction philosophy for which humanity is at level zero in term of civilisation advancement. And if you want to go from level zero to level one, you need a deep and global intelligence which could answer all questions in the best way possible. Open data can really help to achieve this. A global system, a global knowledge accessible for everybody. And it is as well one of his motivation.

*F. Should data providers know what is made on top of open data as re-uses?*

No. Because when willing to maximise the re-uses of open data, you need to make certain trade-offs and one of these trade-offs is that you won't be able to track every re-use that is made from your data. But you will see evidence. And certain consultants will come with some ways how to calculate the value behind and again a data provider will have some 'raison de vivre'.

But for Pieter Colpaert, it was the same with websites. It was really difficult to assess what were the impacts of such website on your sells. And now, with Google Analytics for example, you can measure certain things. It is not perfect but it is improving. Right now, we are at the start of a new type of web which is really hard to measure. And we have the same thing with open data. We are still in an early stage but we will see a network effect. It is the difficult thing with open data: the more people will get involved in this movement, the more people not involved will feel the disadvantages. Example of bus companies in a certain area, if all except one are present on Google Maps, the last one will feel the disadvantage of not being part of it.

*G. Then what should data providers do to support re-uses?*

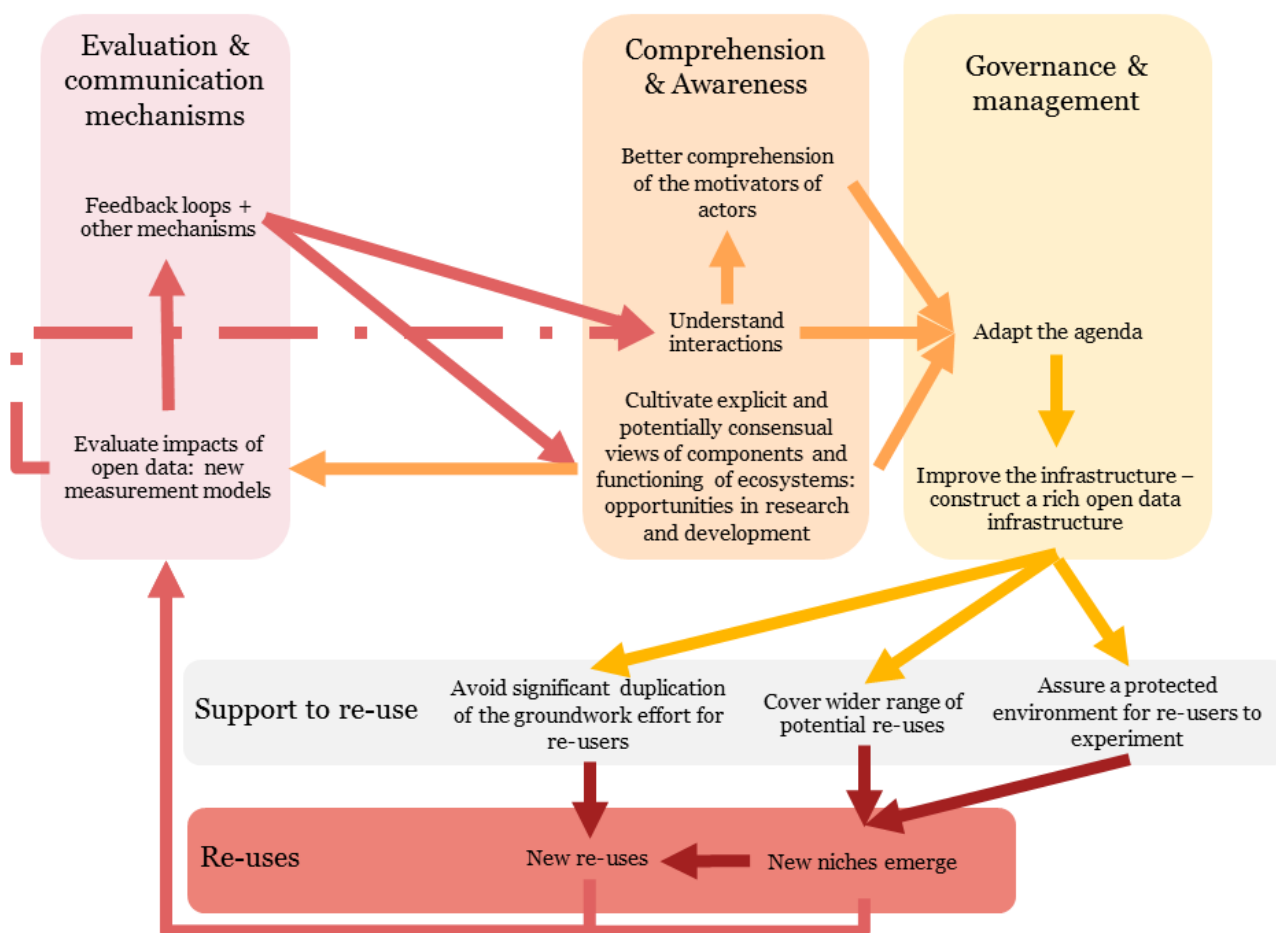
In his perspective, they should provide ways to work with their data. First priority is publishing data. Second is querying their data. For example, a querying for all Belgian railways stops, they will need then a way to resolve certain stations, for instance with all the different possibilities of querying one station name (Bruxelles-Midi, Brussels-South, Brussel-Zuid, Bxl midi etc.). It could be one of the small web services that they provide towards end-users. But it is not a priority since third parties can do that.

And about the investment priority of governments, Pieter Colpaert is not exactly following the same scheme than the City of Gent. As to know, a smaller investment for constructing and publishing data than for building a community. He defends that it is specific to Gent where every initiative needs the adhesion of citizens to work. The bottom-up and top-down processes that we can see at the same time in Gent cannot be generalized. In Gent there is this social aspect of linked data as well and also an ecosystem. Open Knowledge is still working on building

a national ecosystem. But the community building should not be the responsibility of a government but of Open Knowledge Belgium. Why OKFN? Because they are involved in this since the start and he will be really disappointed if tomorrow a new actor in the ecosystem organises for example a new hackathon to gather different actors of the ecosystem. Because it is what they have been organising for the last five years and there is no need for one more hackathon. What he would expect from new actors engaged in the ecosystem is to collaborate with existing actors and not doing their own small-scale add-ons.

### 1.3.8.1.1 Second part: interviewee's perspective and confirmation of observation

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



H. Do you agree with this graph?

Yes, if it is centred on governance.

*I. Have you seen those mechanisms implemented in Belgium?*

We can see the first evidences of this feedback loops. For example journalists that justify their articles with open data sources as data.gov.be and ask questions to governments from these articles.

The informal feedbacks that represent the reality at the moment are not only true for Belgium but worldwide. Right now, he does not see any good ways to make machine-readable feedbacks. It could of course be improved, read article on 5-stars of data portals:

<http://pieter.pm/5stardataportals/>

Improving feedbacks is not the top priority but it depends of course on the kinds of dataset.

But direct feedback is not the way it is going to happen for Pieter Colpaert. In his opinion, it will be more shaped like a git, in a decentralised way. Everyone is just keeping track of his own feedbacks. And it is going to be a global initiative. It is not Belgium who will be a front runner in processing feedbacks on top of data. So for him, once the tools will evolve sufficiently, we will see the development of machine-readable feedbacks. Example of initiative:

<http://dat-data.com/>

*J. What are for you the main challenges that the Belgian ecosystem should solution?*

From a technical perspective. Publishing datasets that are interoperable. Interoperability of the datasets is the main challenge that should be tackled by data providers: “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 kind of datasets together.”

You can summarize interoperability by 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 and that's all.
3. Syntactic: Are we using 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 towards that.
5. Querying: How can we ask questions over datasets? Because a question does not simply stop at a machine. And that is where we need to do work on: creating several ..., building blocks on the Web which provide certain functionalities on top of a certain dataset. Read [linkeddatafragments.org](http://linkeddatafragments.org)

*K. What do you think about hackathons?*

*New forms of hackathons like in Gent: <http://www.appsforghent.be/creativering>*

[For Pieter Colpaert, this new type of events are too early. Data providers need first to invest in better interoperability of their datasets.](#)

Even if it is a clear opportunity for them to receive a fully integrated support, why are they or not participating to such events?

*L. What are the main factors on which innovation is dependent in the open data domain?*

A less costly interoperability.

### **1.3.9 Sven De Meyere – Realo – 30/05/2016**

#### **1.3.9.1 First part: interviewee's perspective without influence of our previous work.**

*A. What is the general state of the Belgian open data ecosystem from your perspective?*

Evolving a lot in the last few years: 4 years ago it was almost non-existing except for some open data enthusiastic like Aerio. Governments have made lots of efforts like in data accessibility. But still much to be done compared to other countries: data quality is not always good, no convention about data structure, how to update, where to keep it. There is now a big opportunity for a third party to make the link between feedback providers and data providers.

Example: Realo is finding mistakes in the open data sources thanks to end-users and internal review. They have therefore a more updated and precise version on some parts of the dataset than the government.

*B. Position/role of Realo in the Belgian open data ecosystem?*

Realo is using open data and other data to build their algorithm (valuation engine) to determine what the value of different houses in Belgium are.

Culture was a big challenge: at first, people disagreed with their evaluation tool, saying that their houses were under evaluated. For example, if an owner was buying a new kitchen at a specific price, this owner was waiting the value of the house to increase from the same amount. But in the reality a potential buyer could evaluate the kitchen at a different price. In Belgium particularly, the valuation of a house is a really sensitive topic. It is what Realo wants to change. Ignorance comes at a very high cost in the real estate sector. People are interested by mitigating the uncertainty surrounding the purchase of a house. Concretely, information about neighbourhood or the right price for the house are information that interest them.

So culture was a barrier for the first months, but it is now changing completely due to different reasons: support from Alexander De Croo, Realo is also building the real estate portal of tomorrow not yesterday which means that the new generation is more data-driven. For instance, they will make online researches for a holiday booking etc. 'They are in a 'research mode'.' When the 'older' generation does not think they need it, people will ask a third service to book for them. This analogy can be applied to the real estate sector. The older generation thinks that it does not need extra data to find a house. But once it has started using the data, it

finds it useful. Realo's job is to convince this generation of the interest of their algorithm and services.

Concretely, in a first time, Realo was not explaining well enough the services they were proposing. People thought value put on their house was a judgment more than a valuation. To face this, they organise a customer team with the task of answering to clients' questions. And they also made a branding campaign about the valuation they were offering on social media etc. with the focus on "Do you know how much is your house worth?". By the same middle, they proposed a link to their website to allow people to calculate an approximate value of their house. They organise as well the branding campaign true TV and radio commercial as well. The main idea behind this was to change behaviour of citizens towards their understanding of Realo's business. Realo wanted them to think: information publicly available about the valuation of houses, including mine, it can be a good thing.

*C. How could the governments and data providers improve the culture of openness and transparency?*

What Realo is experiencing is the fact that people (citizens) don't know that data are already collected by governments and made available. And so when Realo explains to people that they are only using government's data available. Often people are surprised. Even if before open data, lots of information was already available online or in specific places. About the update to data, it must be data on which everybody can have an objective view like bus stops etc. And we could think about all the civil servants (police, workers etc.) for whom the content of the job is changing with the society. As open data is part of the society (data in general), it could be integrated as well into their job's content.

Sven De Meyere would also push for letting end-users update the data available because governments don't have always accurate or up to date data. So it could be possible to fix both problems at the same time. For updates, government has already some tools like authentication tool that could help checking if the update is right or not. For authentication, it could help notice if the person has the 'right' or the 'knowledge' to change the data or not. They could also work

with local governments which represent the first layer of contact with citizens. A bottom-up approach on this could be probably more efficient like in Gent. For him, the situation in Gent could be copy paste in other cities in Belgium. There is a community, but with nothing really official in the relations between members of this community. He says that for him it could be a problem.

*D. Interactions with actors of the ecosystem: what are the dependencies for services providers; what are the services proposed to actors; what are the relations with actors with the same profile, what are the relations with potential entrants inside the ecosystem; What metrics or measures do you think are the most adapted to assess the interactions between actors inside the ecosystem.*

No real interactions with the data providers because there is no real responsible person communicating with re-users and for whom re-users know they could contact in case of need. In Sven De Meyere's view, there is no single open data department managing everything about open data and keeping in touch with people using their data. There are many stakeholders in the case of governments, no single data ownership etc.

An example of communication between actors of open data, Cambio contacted Realo with a tweet, asking for simple collaboration (putting Cambio' sharing points on Realo's map). 4 days later, the map was updated with the changes on Realo's website. Of course Cambio is really agile and still small.

When asked if events already existing are not a good opportunity to share information and interact with actors of the ecosystem, Sven De Meyere answers that they are good opportunities but it is only once in a year like 'AGIV trefdag event'. It is still an occasion for Realo to show to data providers that should not be afraid of re-users but instead be proud of re-users. For him, one of the main goal of governments should be to allow end-users to take better and smarter decisions. And one way to do this is by releasing as much data as possible. But he thinks also that people should have the opportunity to refuse that certain data are collected on them.

*E. What are the main motivators for you to start open data projects?*

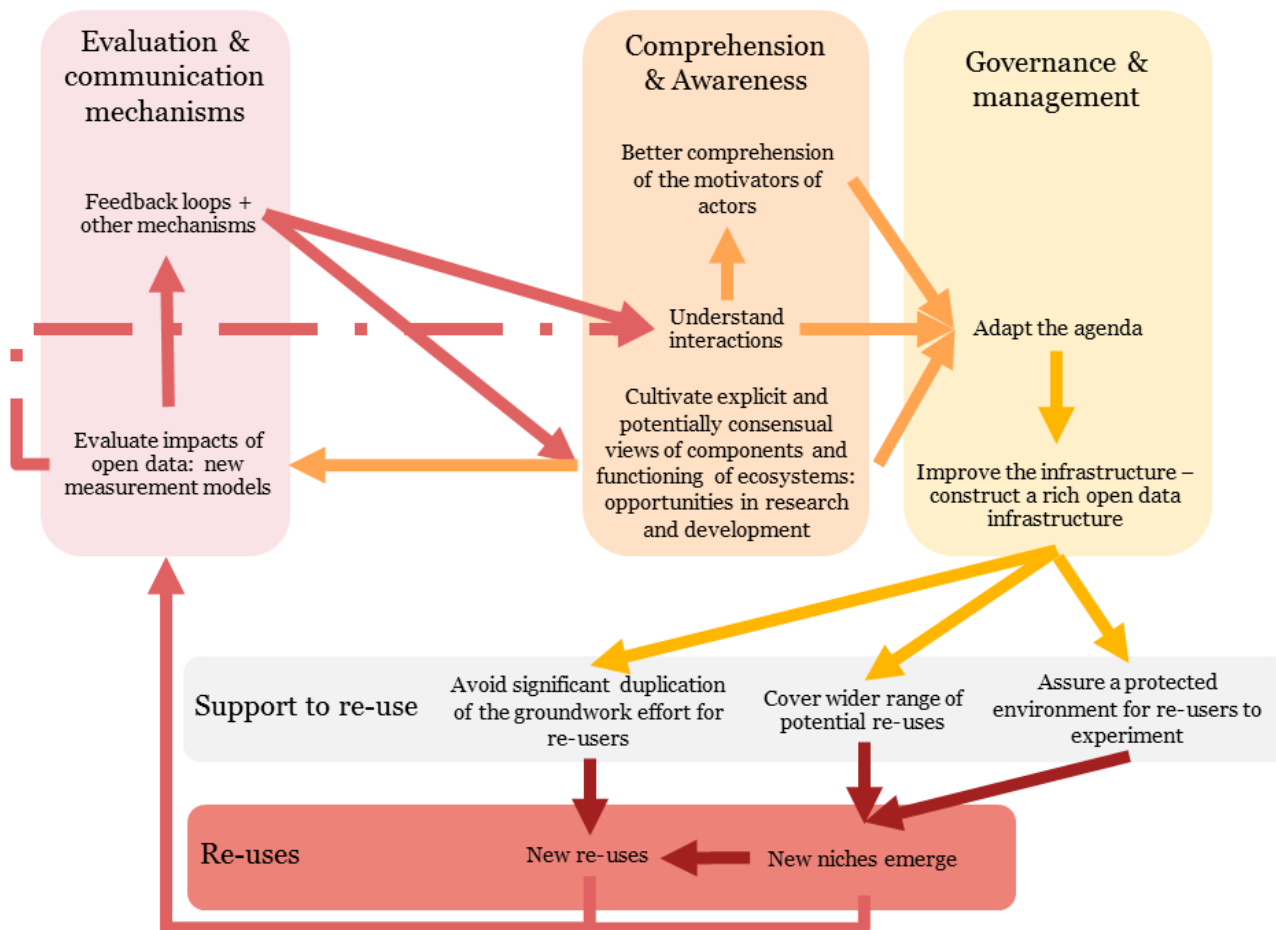
The quality and quantity of information proposed to people in the real estate was not enough. So for founders of Realo, their motivation was based on their frustration. And working with trustful government data (in the end-user's point of view)

*F. What are the main benefits you assessed thanks to open data? (costs reduction, social value, economic value)*

Large scale of the information thanks to open data = Social value. Internally, they are able to measure traffic of visits or leads. But also a kind of neighbourhood rating which is not finalised yet but takes into account different rates like criminality, income, aso. The difficult part is: they don't know if the information provided would not be there, people would still contact a real estate agent (meaning that information has an influence). And there is still a difference between contacting an agent and buying a house. At the moment, they don't have the feedback from agencies on this information. This is something on which they are working at the moment, collaborating more and more closely with agencies (Realo Pro profile). "They should not build a service for themselves. They need to ask end-users, citizens or agencies to tell them what they want to see in their service proposition."

#### **1.3.9.1.1 Second part: interviewee's perspective and confirmation of observation**

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



G. Do you agree with this graph?

“Yes, I think feedback loops are the thing that we are missing in Belgium at the moment.”

H. Have you seen those mechanisms implemented in Belgium?

For example, Realo has a presentation in one event organised by a data provider next week. For them, it is a one shot possibility to share with the data provider the feedbacks on how they can improve the data and how they can make it easier for them to re-use it. Interactions could have help adapting the agenda and consequently have big impacts on re-users who would be really thankful for data providers. For him it could be very interesting if they had frequent contact with the Belgian or Flemish government.

*I. If not, would it be interesting if they were implemented? And would you be ready to participate to this dialog between actors or not and why?*

One possibility for Sven De Meyere to implement it could be to organise monthly or three-monthly meetings with working fields and data providers to have an open discussion about: what is available or not, how they all can improve the landscape, who should take responsibility for what (it should not be always the government just because they are at the moment the biggest players).

It should not necessarily take place by industry sector. The first step could be to welcome everyone interested to see if it attracts interest or not. It should not be a one way communication. At the moment, it is really difficult to know how many actors would be interested in spending their time doing this. After a few events for everybody, organisers should see if smaller groups are needed or not.

Organisers should be a third party in which every actors interested could take small financial participation so that we can really see who actors ready to invest time to construct interactions are.

*J. What do you think about hackathons?*

*New forms of hackathons like in Gent: <http://www.appsforghent.be/creativering>*

[It is a good idea. Realo is a good example of this. Because for him, without open data, Realo could be more compared to Immoweb for example.](#)

Even if it is a clear opportunity for them to receive a fully integrated support, why are they or not participating to such events?

Realo is participating in different events like the last Appsforghent. One problem for him is in the follow-up. The low number of participants is not a real problem. It can happen as well in other sector than open data. But the problem is in the follow-up. To make ideas concretely implemented. Events need to be clear since the start about the purpose of the event: is it

constructing a real re-use through a new company or just to build awareness and creativity around open data.

For example, after Appsforghent, the participants from Realo were really surprised when they get contacted by one of the sponsor of the event to develop further their idea proposed during the hackathon. Because Realo has no commercial value in participating in these events, they have already the necessary funds and team to grow. However, the team who had won the first price during the event received no proposition from any company.

*K. What are the main factors on which innovation is dependent in the open data domain?*

More bottom-up approach. Easier for re-users to create their own datasets or to provide open data or to contribute to existing open datasets or moderate open datasets.

One simple example is the geospatial website with a few thousands of people really interested in contributing by updating the map of the universe. Sven De Meyere thinks that it could be interesting to discuss how similar contribution could be made for other topics like hospitals etc.

Why not thinking of a neutral party moderating the whole ecosystem because now it is really unstructured, everyone is doing a bit of everything: governments releasing and supporting some re-uses and events, companies re-using data but not able to provide feedbacks, the Open Knowledge community... So everyone is doing something on their own but not talking enough to each other basically.

Final comments: Be careful for the different opinions about the ecosystem depending on the concrete place of the interviewee in this ecosystem.

Sven De Meyere thinks that this discussion should not be only at the Belgian level. For example, Realo would not stop growing and proposing services at the national level. After Belgium they will grow internationally. In consequences, it would be inefficient to construct one ecosystem in a certain way in Belgium to see that it has been constructed in a different way in neighbourhood countries and then to have to adapt everything again.

### **1.3.10 Stéphane Vince – Digital Wallonia, Walloon government – 02/06/2016**

#### **1.3.10.1 First part: interviewee's perspective without influence of our previous work.**

*Definition of open data ecosystem:*

*“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)*

*When applying the concept of ecosystem to the open data field, the literature describes it like:*

*“A multi-level and multidimensional entity where raw material, as far as distribution and developing are concerned, is the target of cooperation.” (Anneke Zuiderwijk, Marijn Janssen, et al., 2014)*

#### *A. What is the general state of the Belgian open data ecosystem from your perspective?*

At local levels, Ghent and Antwerp are good examples of open data projects, demonstrating the achievements of an open government process. In Wallonia we are still beginning with open data. The public agency responsible for open data, ‘l’Agence du Numérique’, is attentive to what is done in Belgium and internationally but the culture for openness is less developed than in the neighbourhood countries or in Flanders. The Walloon decree for ‘Open data by default’ is about to be voted at the parliament.

#### *B. Position/role of Walloon government in the Belgian open data ecosystem?*

‘L’Agence du Numérique’ is at the moment trying to convince public administrations to open up their data. It is really difficult since many public administrations have the conviction that data collected by public administrations belong to those and that they are the best placed to propose services on top of these data. The ideological debate is therefore around the role of public authorities inside the whole society. There is no awareness concerning the need to change the way public organisations are working in order to adapt to the imperatives of the society. The

problem in this debate is the support granted by many top managers and politicians to the old thinking side of the administrations. One example that Stéphane Vince is describing more in details later on in the interview is the auto satisfaction of one top manager for having spent large amount of money on a digitalisation project in which citizens can fill pdf documents. The problem being the semi-automation of this procedure. In consequence, lots of time is spent in human input to transform pdf documents into machine readable documents. In synthesis, he has no influence on public administrations to make them change their minds. He hopes that the new decree will help forcing mentalities to change.

*C. Internal structure of the open data section (in Walloon Gov): organisational structure?*

Stéphane Vince is the only employee working on open data, about 1/10 of his time. But the agency is planning to recruit one fulltime employee for open data projects.

*D. What is your view over the maturity of the Belgian open data ecosystem?*

In Wallonia, maturity is really young.

*E. Vision on short and long-term (5-10 years) over the evolution of the Belgian ecosystem and the position/role of the Fedict.*

They are thinking of opening data bit by bit in a similar way than in France with the 'start-up d'état' project or the creative open data event in Ghent recently. The 'start-up d'état' project is a collaboration between the French state and potential re-users who can come to the public administration with a specific project to simplify the administration procedures and a first draft of business model, asking for particular datasets to re-use. The administration selects the start-ups to which it could provide the data and helps them achieve their objectives with more support like coaching.

*F. What are the main challenges faced by your organisation? (resources, harvesting, providing data quality, standardisation, interoperability, ...)*

Awareness of the population, the politicians, and companies as well. At the moment, he has no real good example of re-users of open data who have a viable business to convince public administrations.

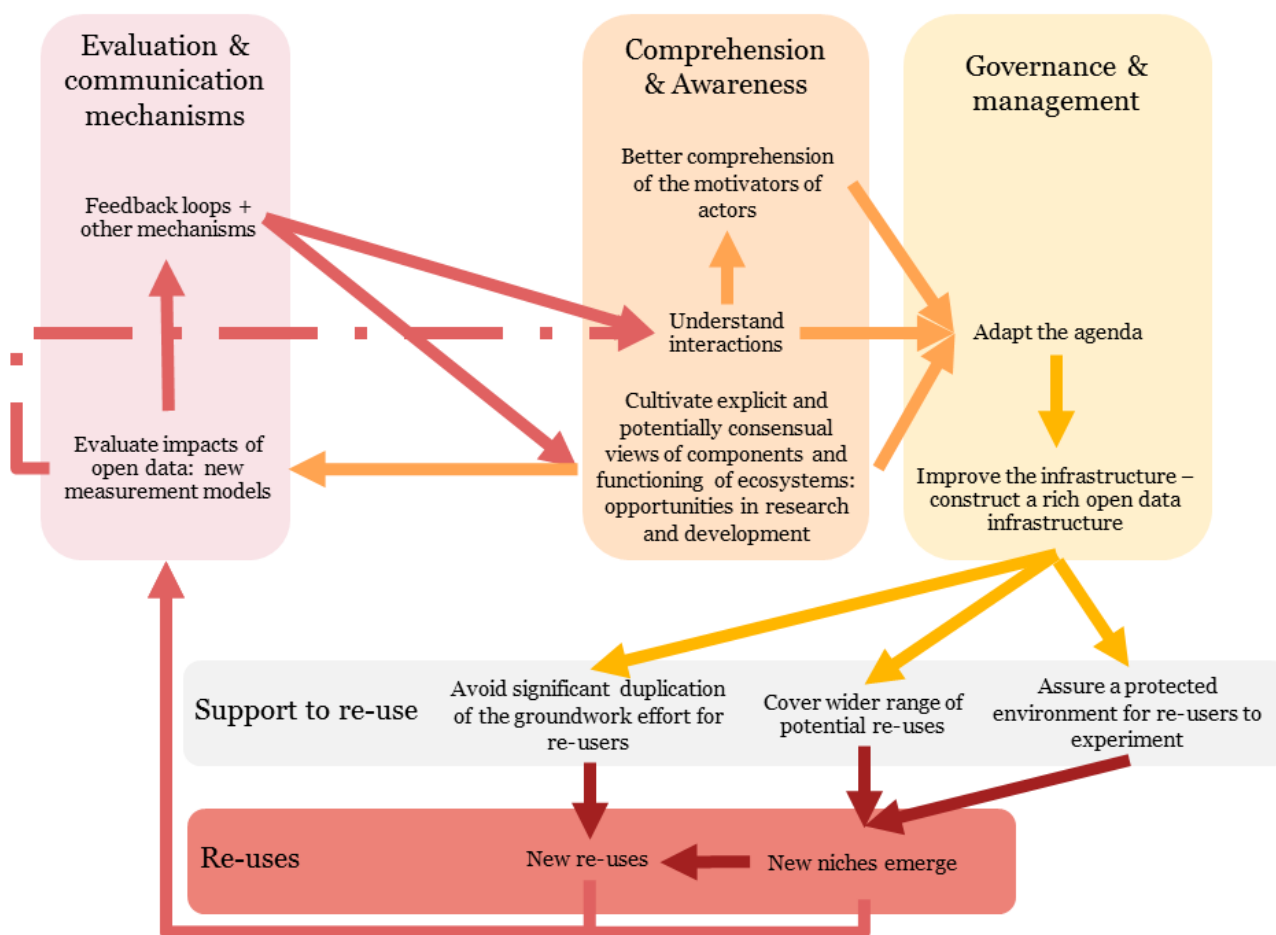
Mentality of the administrations. They are asking for results of open data before starting opening up their data which is the opposite way than common government supported open data initiatives.

Proposing immediately data-as-a-(technical)-service is not a reality in Wallonia. He hopes to be able to improve authentic sources digitalisation inside administrations. Because at the moment, three processes are co-existing: paper process, semi-automated process (pdf which can be filled online), and fully automated. Next to this, the issue to collect open data concerns the unstructured formats that most administrations possess. Stéphane Vince says that it is a broader issue than just open data. It is about digitalisation of administrations, the complete IT structure which is chaotic right now.

For what concerns open data, he was once in a meeting with a person who said being 100% for opening up data. But once he asked where the data were, this person pointed a cupboard full of sheets. Educating people inside and outside the public administration about data and data management is required in order to achieve results.

### 1.3.10.2 Second part: interviewee's perspective and confirmation of observation

*Potential benefits from the ecosystem to support re-uses of open data and innovation*



#### G. How to support adequately innovation?

For example, the 'digital Wallonia hub' is in the first steps to be created. It represents the Walloon equivalent of iMinds in Flanders: a junction between the business world and the academic world.