BEYOND BUSINESS MODELS

Stimulating balanced Open Data Ecosystems

Jonas Breuer
iMinds-SMIT, Vrije Universiteit Brussel, Brussels, Belgium
jonas.breuer@iminds.be

Keywords: Linked Open Data, Ecosystem, Inter-Stakeholder Relationships, Public Value, Business Models

Abstract: This article provides insights into inter-stakeholder relationships, which shape (linked) open data initiatives and their outcomes. The aim is not to quantify potential impact of open data on economic growth, or to focus on business models per se, but to provide suggestions for sustainable open data ecosystems that combine commercial incentives of private sector entities with public organisations striving for the creation of public value. This is regarded as essential at this stage of the open data movement; because multiple barriers still exist unfold its potential, both on a microeconomic and a macroeconomic level. Interrelations between actors are analysed in the context of an ongoing project. The analysis reveals that the public sector probably needs to take a strong leadership role, also in terms of financing, but that this is not only negative, as it is in a position to shape decisive factors, such as underlying data governance.

I. Introduction

Open data has evolved to a much discussed notion, in politics, business, and beyond. Among the leading arguments of its advocates is the potential, in terms of public and economic value: "The Commission has launched an Open Data Strategy, expected to deliver a €40 billion boost to the EU's economy each year" (European Commission, 2011); "enormous potential to create more accountable, efficient, responsive, and effective governments and businesses" (G8, 2013, p. 1). However, not only remains its hailed economic impact insufficiently established on a macroeconomic level. Also on the microeconomic level uncertainty prevails, in terms of adapting existing and especially creating novel, innovative business models (see e.g. Buchholtz, Bukowski, & Śniegocki, 2014; Networked & Electronic Media Initiative, 2013; Vickery, 2011).

This is the reason why the aim of this paper is neither to quantify potential impact of open data on economic growth, nor to focus on business models per se. Rather, the paper aims to provide suggestions for establishing sustainable open data ecosystems, with the intention to reconcile the definition of open data (and public value arising thereof) with business models and commercial activity. Despite multiple barriers for to unfold the potential, sustainable value generation with open data is feasible in manifold ways, although usually not straightforward. The main objective should be, therefore, to instigate suitable business models through well-configured open data ecosystems.

Who, then, takes up which tasks? And what does the operation of certain tasks mean for the allocation of control in the ecosystem? So far, it is only certain that it is necessary to combine private and public resources. The leading question guiding this paper is: *What are the power-relations within open-data environments, and do these interactions shape the results and outcomes of respective initiatives?*

First, an in-depth assessment of (linked) open data ecosystems is presented. This is followed by a use case of an ongoing linked open data project, which provides the context for analysing concrete inter-stakeholder relationships, assessing interactions and

interdependencies in wider ecosystems. The subsequent discussion addresses some of the bottlenecks revealed in the case study, illustrating also the decisive issue that 'publishing alone' does not generate impact directly: facilitating accessibility and stimulating reuse seem to be most relevant. Finally, some concluding remarks complete the paper.

II. (Linked) Open Data

The following pages present an in-depth assessment of (linked) open data, establishing a starting point for subsequent analysis. This is necessary, as various issues exist within the complex field of Open Data (OD). We begin with the question about what OD actually is. Because, in principal, any data can be(come) open, the prevailing definition being that "a piece of data or content is open if anyone is free to use, reuse, and redistribute it" without restriction (Okfn, n.d.). It is claimed to be good for society, in terms of democracy, economy, efficiency and other aspects along that line.

At the heart of the current debate lies Open Government Data (OGD), "produced or commissioned by government or government controlled entities" (Goedertier, 2013, p. 10): statistics, information about public-service-delivery, science, finances, etc. (Okfn, n.d.). In fact, the commonly used notion of open data relates to OGD more often than not: Janssen et al define OD "as non-privacy-restricted and non-confidential data which is produced with public money and is made available without any restrictions on its usage or distribution." (ibid, 2012, p. 259) Thus, OGD is public sector information (PSI) made available openly for reuse, as defined and regulated by the Revised PSI Directive in the EU (Directive 2013/37/EU, n.d.). OGD can therefore be defined as a public good, marked by non-excludability and non-rivalry: access to the data "cannot represent a source of competitive advantage *per se"* (Ferro & Osella, 2013, p. 2).

Linked open data (LOD) goes beyond openness in its nature, its benefits, but also in its complexity. Linking OD increases the potential by semantically enriching it, extending fragmented parts of information to bigger knowledge domains (Berners-Lee, 2006). It also renders processes more difficult, in particular on the supply side: the setting of rules and standards, their implementation and application, the creation of links and the integration of data. All this can probably not be achieved by one organisation alone. In fact, it might only be properly effectuated by a coordinated ecosystem of diverse actors, if it is not even too complicated to be sustainable at all (see e.g. Milicic, 2011).

In an ecosystem, the public sector plays a structural role for (L)OD, as governments and adjacent organisations already possess substantial amounts of data, collected as part of their function. They also constitute the main funding source for respective projects. Today, "the prevalent business model […] is the one where investment and maintenance costs are covered through on-going public funding" (Goedertier, 2013, p. 6). Public resources are limited, though; financially, and more importantly also in terms of expertise. For establishing and maintaining a sustainable LOD environment in the long term, therefore, combining public and private efforts probably yields most preferable outcomes, also in terms of the hailed impact in society.

But how can a public good such as OD be subject to commercial activity, without curtailing openness and forestalling the generation of public value beyond revenue for some companies? For (L)OD, a gap prevails between publishing and reuse, i.e. supply and demand, which needs to be bridged through appropriate infrastructures and tools, to create value also for a

mainstream audience (Dadzie & Rowe, 2011). In other words, to have any impact, the data must not only be open, but accessible and usable (see e.g. Goldstein & Dyson, 2013, p. xi). Facilitating the making accessible and usable could be outsourced to the private sector. Ferro & Osella (2013) offer a meaningful framework for identifying roles of commercial actors in this process. They argue that, on the one hand, *enablers* facilitate use of data (e.g. through retrieval, storage, categorisation, exposure) "behind the scene". *Re-users*, "on the front line", on the other hand, utilise the data as part of their own value proposition.

A prominent example of *enablers* of Open (Government) Data is Socrata[1], a Seattle based company that offers an advanced open-data platform. Their solution is highly successful in the US and implemented on all levels of government, and the company is extending its customer base to Europe's public sector. Also the collaboration between Portland's transit agency (TriMet) and Google is a best practice, from which *GTFS* originated: the General Transit Feed Specifications, nowadays the most widely used standard for open transit data (McHugh, 2013). The UK's *Statute Law database* is another example of how commercial actors can collaborate with governmental and non-profit actors in *enabling* the use of OD (in this case UK's statute law). They then build and sell products and services on, or with that usable data (Sheridan, 2012).

Plenty examples of "OD front line" actors are around; *Stat.io*, for instance, a young company that offers open socio-economic data in a highly accessible fashion (based on a map) - for free - and extends services for those who are willing to pay subscription fees (Stat.io, 2013). Mapbox[2], Google Maps' rival, offers highly customisable maps used by Foursquare and other popular services. It is mainly based on OD from OpenStreetMaps and NASA, but combines it with proprietary sources. Such intersection of OD with private, restricted information for new goods or services promises viable business models. Through the right licenses[3], companies could be charged for combining data in ways that generate revenue for them.

The current situation suggests that commercial incentives add substantially to the viability and sustainability of O(G)D-related initiatives. The story of GTFS' origins in Portland is a good example. The project would not have been as successful without the contributions of Google as a strong commercial partner. It should be regarded as "a success story of a new model of cooperation in order to solve a problem that cannot be addressed directly with either market forces or a classic government solution." (McHugh, 2013, p. 133) The interaction between stakeholders for "creating and capturing value from technological innovation." (Chesbrough & Rosenbloom, 2002, p. 532) is the decisive factor.

In the context of OGD, responsible public sector organisations occupy a central position, but they can only oversee a limited number of roles, and they are dependent on expertise from the outside. *Enablers,* i.e. companies holding core competencies in the ecosystem, will execute decisive roles. Interaction between private and public efforts is necessary, but allocating too much control to a corporate partner might also risk that the partner and its commercial objectives become the dictating driving force (McHugh, 2013, p. 132), as the ecosystem tends to co-evolve and align itself with the directions set by the central actor: "the role of the leader is valued by the rest of the community. Such leadership enables all ecosystem members to invest toward a shared future." (Moore, 1993, p. 76) Strong and visionary leadership on the

[1] http://www.socrata.com/
[2] https://www.mapbox.com/
[3] OKFN's Open Database License is a good example: http://opendatacommons.org/licenses/odbl/summary/

part of the government seems to be indispensable, as existing examples of successful OD initiatives demonstrate (see e.g. Goldstein & Dyson, 2013, Chapter 2; McHugh, 2013).

The Open Data Ecosystem

An ecosystem is a community of interacting organisms and their physical environment (see e.g. Moore, 1993), the term used to indicate a complex network or interconnected system; entailing also that linear value chains make way to multi-directional dependencies. This is applicable to OD, where the nature of actors' roles shifts, rendering the distinction between supply and demand, between providers and consumers unsteady. Data consumers become providers of data solutions, which the original data providers then use. In fact, an important incentive for public authorities to open up data is, in many cases, to gain organisational efficiency internally (see e.g. Rosseau, 2013; Goedertier, 2013, p. 6). Such dynamics renders OD-ecosystems complex ab initio, and an assessment of OD's benefits needs to take them into account.

On the one hand, there need to be effective supply-side mechanisms, involving factors beyond simply opening data, such as technology, standards, licenses, operational guarantees, awareness. Goedertier et al (2013, p. 6) suggest in their study that increasing amounts of released data, in appropriate form and with adequate technology, will entail more (re-)users, "contributing to a thriving ecosystem." In this light they state, though, that 3rd party reuse of provided LOD is currently far from widespread (ibid, 2013, p. 6). On the other hand, "thriving ecosystems" for use and reuse of the data (Chui, Farrell, & VanKuiken, 2013, p. 169) must be fostered collaterally.

Taking this into account, we adapt a definition of a business ecosystem (Moore, 1993), as a network of *actors*, which are inter-related and mutually dependent, the ecosystem being defined and delimited by its relation to certain objectives. The overall strength and weakness of such an ecosystem depends on how each *actor* – both public and private – contributes (adds value) to the ecosystem (Ballon, 2007). Interactions between *actors* are key. They co-evolve over time with *actors* and their *roles*, are influenced by the wider environment, by each actor's objectives and interests, and also by inherent power relations. According to Ballon et al (2010), *stakeholders*, *actors* and *roles* can be defined as follows.

Stakeholders are concrete and current real-life organisations that control key variables of project related processes. This entails that strategic objectives of a stakeholder also shape the project itself. ***Actors*** are entities that encapsulate a certain set of *roles*, although concrete *roles* can be shifting. The concept of an *actor* resides at a level of abstraction that not necessarily corresponds fully with current stakeholders. ***Roles*** are defined as a discrete set of responsibilities, actions, and authorisations that together constitute a logic for an activity. As mentioned before, roles are not coherent here, they shift and the actor that collects data and opens it up might or might not also publish it, and probably becomes a re-user or user. A re-user might develop an application or offer a service to the entity that has collected data in the first place. Thus, actors in the generic OD ecosystem occupy different roles simultaneously.

First, there is the role of *collecting/gathering,* which is conducted mostly by public sector organisations. The second role concerns the o*pening/providing* itself, which should take place 'by default', laid down by the PSI-Directive (Directive 2013/37/EU) as primary principal. The *publishing of data* is the third role, unfolding through some kind of online portal, which provides access to the data for everybody; this goes beyond providing data, as it concerns the enabling of access. Fourth, the *linking* and semantically representing of OD is a role that

substantially increases the OD potential. The latter roles are intrinsically connected and need to be assessed in conjunction, although they are separated here. The reason for this is mainly that they constitute different stages in the process of value creation with OD, but also because they might be separated organisationally in an ecosystem, i.e. outsourced to other actors.

Afterwards, the *operating/maintaining* of technology & data takes place once the data is 'out there'. Sixth, the *re-using* concerns everything that utilises data for other ultimate purposes. It includes value-adding (e.g. making data more accessible and understandable for citizens, as Stat.io, Mapbox and many others do) and value-creating (e.g. business models that are based on the content of OD). The *consuming/using* data is the 'last' role in the OD ecosystem, held by all those entities whose primary objective is the content of the data, i.e. the contained information.

III. Methodology

To analyse the inter-stakeholder relationships in a concrete setting, the MACTOR (*Méthode ACTeurs, Objectifs, Rapports de force*) method was borrowed from the field of strategic planning (Godet, 1991, 2008), and modified to match the context. It is well-suited to provide an overview of interrelations between actors and objectives. MACTOR is adapted here to concrete parameters of a **Business Model Configuration Matrix** developed by Ballon (2007) and expanded by Walravens (2012), which emphasises interrelations between control- and value-related aspects (see Table 1). The matrix was created for assessment of inter-actor processes in wider ecosystems. It also enables us to integrate positions of both private stakeholders (focussing on revenue generation), and public sector entities (aiming to create other kinds of value). This dynamism and ecosystem-centrism are particularly relevant in the context of OD, not in the least to consolidate disparate strategic perspectives in one comprehensive and encompassing project (Ballon, 2007; Walravens, 2012). The matrix thus provides the framework to phrase strategic objectives for the analysis. MACTOR then functions to simulate actors' positions on these objectives, towards each other, and how that impacts the outcomes of the project.

	Control Parameters		Value Parameters	
Business Design Parameters	Value Network	Technical Design	Financial Design	Value Proposition
	Combination of Assets Vertical Integration Customer Control	Modularity Interoperability Intelligence Distribution	Investment Structure Revenue Model Revenue Sharing	Intended Value User Involvement Positioning
Public Design Parameters	Public Governance Parameters		Public Value Parameters	
	Stakeholder Management Good Governance	Technology Governance Data Ownership	ROPI Public Partnership Model	Public Value Creation Public Value Evaluation

Table 1: Business Model Configuration Matrix

An on-going project provides the context and material for the analysis at hand. The **case study** approach was chosen for its ability to describe "a contemporary phenomenon in its real-life context." (Yin, 1981, p. 59) Data was collected directly related to the project (notably through stakeholder interviews) and desk research (literature review and document analysis).

Case Study

The open data initiative, from which the content of this work originates, is characterised by both potential benefits and drawbacks for all involved stakeholders. Initiated by a government

department, the objective was set to publish data openly, and to establish the structures necessary to actually unfold OD's positive potential. A primary focus on information and data around research (both research itself and corresponding metadata) was taken for practical purposes, driven specifically by demand of the department itself: this data is an important knowledge asset, needed for its own monitoring and policy preparation activities.

Several steps were undertaken. First, an OD environment for publicly financed research information was created. Here, the data remained in silos (corresponding to organisations such as universities), and information was thus not as complete and accessible as could be. Second, infrastructure was established to link data beyond silos, to semantically represent it, and to enrich structured data with unstructured sources (publication text, project descriptions etc.). Still, accessibility remains curtailed, in particular for non-expert users. Third, tools were developed to add accessibility to the complex LOD, and understanding of meaningful correlations emerging from links (notably in the form of visualisations).

Stakeholders and their Objectives

In line with the descriptions of OD ecosystems, respective actors and roles (see section II), several stakeholders (i.e. involved real-life organisations) compose the project. Variables established for following analysis are based on their expectations and requirements. Stakeholders are:

The initiating **government department** of economy, science, and innovation, which is primary stakeholder, as financier and data-demanding entity. Accordingly, it presents a clear vision and objectives, such as the returns on public investments: in the particular case of research information data, public institutions fund research and demand the output to be available and usable for everybody. An additional objective is to increase visibility and promote cities, regions, or even the country in terms of research activity. The main vision is, however, that the provision of data and access thereto creates economic value, especially due to its supposed potential for enabling innovation. Economic benefits then, are seen as societal benefits in a broader sense. **Universities** benefit foremost themselves from having accessible and qualitative data. Their objective is in the first place to have better access to their own data and that of other institutions, for organisational and budgetary reasons. Also stimulating visibility and promoting their work is important. **Researchers**, as the producers and subjects of the data, have similar interests as universities, even though they might only be indirectly affected by their institutions' strategic objectives.

A major **technology firm** was involved in the project at an early stage. It contributes knowledge, and services, mainly because it has the appropriate resources at its disposal anyway, but also because it has already been a common service partner to the public sector. **Citizens and SMEs** are actors, primarily because the project is publicly funded. They should neither have to pay too much as direct customer nor indirectly as tax payer. However, as the arguments go, LOD should also somehow affect them positively, i.e. create some kind of public value. They are combined here under the roles performed of "end-users", which also includes other researchers. Due to the content of this LOD, the academic **publisher** is a decisive stakeholder (as the most influential re-user of the data), not least due to the substantial income it gains from reusing the work of the researcher, university, and indirectly the government (as research sponsor). The publisher's objective is revenue generation, which in this case conflicts other objectives, such as openness and accessibility.

IV. Multi-Actor Ecosystem Analysis

The key steps involved in performing the analysis are explained in the form of a flowchart (see Annex) and implemented in detail for the project's ecosystem.

Step 1: Actor Profiles

The first step is to establish "those actors which control key variables, as previously identified" (Godet, 2008, p. 64). For the analysis at hand, they were abstracted from the project's stakeholders, as listed in Fig. 1. It is to be noted that this is a simplification, ignoring some (potential) actors and combining others into one.

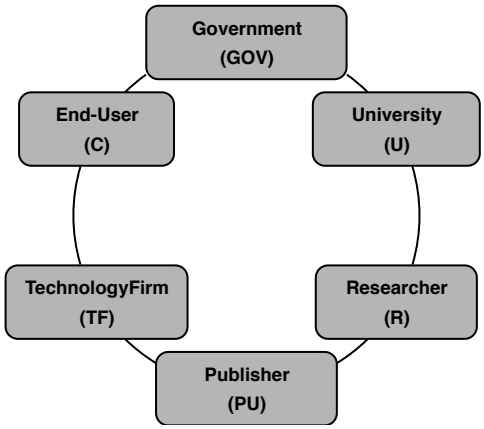

Fig. 1: Actors in the MACTOR ecosystem

Step 2: Influences & Dependencies

The second step evaluates the relationships among actors, based on influences and dependencies. In order to do so, each actor's degree of influence on the other actors is classified, according to five levels from 0 (no or little influence) to 4 (jeopardise the very existence of another actor). The various actors' positions are calculated as direct and indirect influences: the influence of actor A on actor B is the sum of its direct influence on B and of all indirect influences it gains through all other actors (Godet, 1991). Fig. 2 then reveals four position types with respect to the system under study; dominant (very influential, little dependent), dominated (little influence, highly dependent), intermediate (influential and dependent), and autonomous (not influential or dependent).

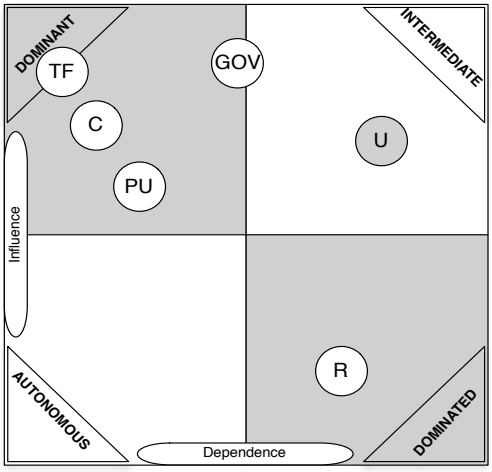

Fig. 2: Map of Influences & Dependencies

The high levels of influence in relation to relatively low dependencies demonstrate not only the strong position each actor holds in the ecosystem. It also means that the dominant actors can probably impede the success of the project. The researcher's weak position arises due to several direct and indirect dependencies; from university, government, but especially the publisher.

Using a scalar, the software can represent the competitiveness of each actor (i.e. the likeliness of an actor to enforce its own

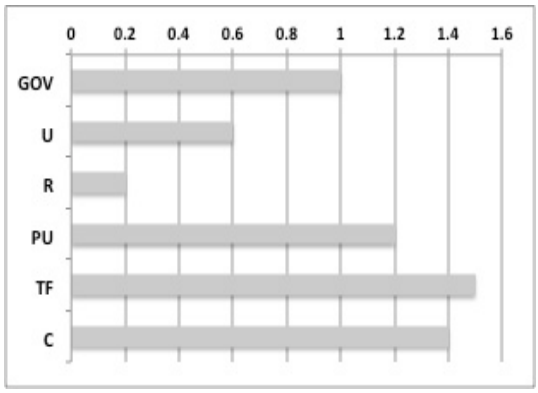

Fig. 3: Actors' Competitiveness

objectives and strategy), taking into account its direct influence and dependence, as illustrated in Fig. 3.[4] Notably, the private actors rank highest here. Alongside them, the citizen ranks high. This is probably the case due to its theoretical (but not always real) influence on the government in a democracy, but also because this actor includes the economic impact of SMEs.

Step 3: Strategic Issues & Objectives

Based on the Business Model Configuration Matrix (Ballon, 2007; Walravens, 2012), key project objectives are derived and defined. In general it can be stated that feasibility arises from a strategic fit between key parameters in the ecosystem. Using the business and public design parameters highlighted in Table 1 and Table 2, six principal issues could be derived, which are expected to have direct or indirect impact on the actors operating in the ecosystem and consequently the project as a whole. Table 2 clarifies the relationships between the parameters and derived strategic objectives.

Table 2: Deriving Strategic Objectives

Objective 1: Creation of Value for the Public
***Intended Value**: A business design parameter dealing with primary attributes that a service is intended to possess. In our case, this is combined with and supplemented by following public design parameter.* ***Public Value Creation:** The justification a government has for taking initiative. Public value is regarded from an end-user/citizen perspective.*
This objective captures the main intention of opening-up/linking data, even though it is not always clear what kind of public value is or can be generated. For the creation of public value, a significant question is whether there is a specific need in society for which the government needs to get involved, where other actors cannot (e.g. market failure). The European Commission prioritizes in this regard direct and indirect economic gains, whereas the Obama administration in the US emphasizes transparency, participation and collaboration, thus consolidation of democracy (Zuiderwijk & Janssen, 2013, p. 1).
Objective 2: Outsourcing Roles to Private Actors
***Combination of assets**: The usage and combination of tangible and intangible resources. This aspect constitutes the relative weight between actors, i.e. hierarchies.* ***Stakeholder Management:** Public design choices made about which stakeholders to involve (public, non-governmental etc.), distribution of competences among them, how they are organized and interact.*
Public sector bodies cannot carry out all roles; they outsource tasks to private stakeholders. Combinative capabilities are indeed decisive, as combinations of assets and competencies leads to synergetic effects. However, stakeholders, their competencies and interactions need to be managed carefully, as the allocation of core competencies to a single structural role leads to concentration of control. The viability of outcomes concerning objective 1 might be undermined if commercial interests constitute the driving force.
Objective 3: Data Governance
***Public Data Ownership**: The terms under which data is opened up, as well as to which actors. It is technological in the sense that selecting or limiting data formats affects which parties can work with it.* ***Technology Governance:** The issues that emerge from technological choices by public entities. Choices for a particular technology or platform, for instance, may exclude certain parts of the population.*
This objective captures the importance of choices regarding data and technology itself, as well as related stakeholders. Data Governance is about how the data is handled. It is highly significant that the data stays public and open to use for everybody. Exclusive partnerships with one actor could curtail openness, for example through too much control over technical functionality or proprietary formats. Effective data licenses are significant, as they ensure public data ownership in the long term.

[4] Competitiveness is calculated using Ii, Di and the Matrix of Direct and Indirect Influences (MDII)ij: Ri =[(Ii–(MDII)ii)/S]x[Ii/(Ii+Di)], where S =ΣiIi =ΣiDi. Ri takes into account actor i's manoeuvre range (Ii–(MDII)ii), i.e. its net direct and indirect influence (Ii) minus its retroaction (MDII)ii. The relative manoeuvre range [(Ii–(MDII)ii)/S] of actor i is then deflated by the coefficient Ii/(Ii+Di), which is between 0 and 1 and allows to integrate the actor's dependence in the equation. The Ri value is compared to 1: if an actor has a competitiveness value of more than 1, it is more competitive than the average.

Objective 4: Accessibility
Technical Architecture: *Important aspects are interoperability (i.e. the ability of the system to interwork with other systems, also in terms of data), and modularity.* ***Public Value Evaluation:*** *The reach of the initiative (who will be reached/affected by it), its quality, impact (will it create benefits for citizens/society as a whole), and whether it delivers value for money.* This objective represents a key issue: there cannot be any value in open data if it is not used, and that it cannot be used if the data (and respective infrastructure) is not accessible for the public. This concerns the reach and impact of the initiative. Creating 5-star open data portals and linking data represents a big step towards accessibility. However, complexity of the system risks at the same time to foreclose adoption and use by the public. In our case, Open Access (free availability of publicly funded research) is linked to the objective.
Objective 5: Risk Distribution
Partnership Model: *Financial relationships between the private and public participants (PPP, PFI, PC), including politics & political tensions, financial implications & risk distribution effects.* ***ROPI*** *(Return on Public Investment): Expected value generated by a public investment, purely financial, public, direct, indirect or combinations of these, and how investment is justified.* This objective captures whether the public sector undertakes the main investments. Financing and advancing initiatives can be difficult. Therefore, the public is obliged to carry a main share of the risk, also to convince commercial actors of the initiative's value. ROPI needs to be adequate, either direct or indirect in the form of multiplier effects (increases in GDP, economic activity, job creation...), but real "returns" of open data are not yet clear. To ensure quality and sustainability, contractual partnerships between different stakeholders are prominent and often successful options.
Objective 6: Business Models for private company
Investment Structure*: The cooperation of actors, how costs (investment) are shared. It also relates to sunk costs and up-front investment.* ***Revenue model:*** *Concerns the income streams (direct/indirect) for an actor's activity.* This objective captures the options for independent actors to build business models around open linked data. This is significant, since fostering of economic activity is central to the supposed public value to be created.

Step 4: Actors towards Objectives

Based on the instantiations described above, the position of each actor in the OD ecosystem is mapped into Matrices of Actors and Objectives in this fourth step, on the basis of possible alliances/ conflicts and the hierarchy of objectives of actors.

To do this, we examined each strategic objective and ranked each actor's convergence, divergence, or neutral stance as (+1, -1, 0) in the MAO matrix. Additionally, in the 2MAO matrix, we ranked if the level of opposition or agreement with the objective is high, medium or low with (-3 to +3). In other words, the higher the importance of the objective would be for an actor, the higher the absolute value recorded. Note that, at this stage of research, this ranking of positions was made based on initial inputs and needs to be validated by stakeholders as a next step.

Based on this, Fig. 4 illustrates the position of actors in relation to each objective. It shows that there is a relatively high degree of convergence, although some objectives are more controversial than others. Notably O3, concerning the issue of who owns the data, seems to entail divergences between actors, as does O5, i.e. who should make the investments and carry associated risk. These objectives are significant in terms of sustainability and outcomes of

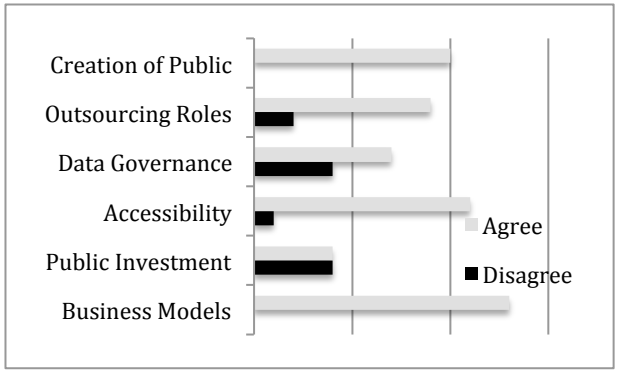

Fig. 4: Actors' Agreement over Objectives

the project and need to be carefully examined.

Fig. 5 then identifies for each couple of actors the degree of objectives on which these actors disagree. The underlying matrix values do not measure simply the number of potential conflicts, but rather the conflict intensity with the objectives' preferences of the couple of actors. Somehow unsurprisingly, strongest disagreements exist between researcher and publisher. Government's preferences are, however, shared with most other stakeholders, hinting at a high potential for realisation of its objectives.

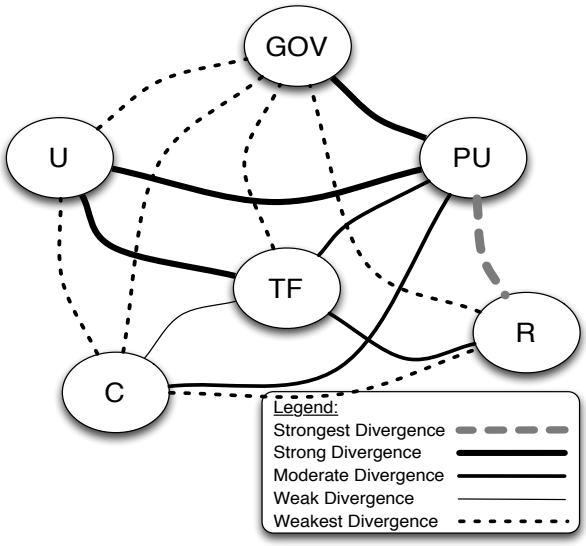

Fig. 5: Actors' Disagreement over Objectives

Step 5: Actors towards Actors

The next step is to formulate the Actor Interaction Matrix (MAA), where we derive how actors' relationships align or not over the objectives. The two MACTOR matrices describing convergent (2CAA) and divergent (2DAA) positions of each pair of actors over objectives are combined through the multiplication property of matrix calculation: the multiplying of a matrix with its transpose yields a number of factors in common for each pair of lines in the original matrix.

The MAA maps actors against actors, based on actors' preferences related to a certain issue as described before. The resulting matrix is made up of both positive and negative scalar products. Each element (nCij) in lower half and (nDij) in the upper half of the Matrix represents the number of objectives towards which actors i and j have a common convergence or divergence respectively – generated by retaining positive and negative scalars only. The result is depicted in Fig. 6, incorporating the average intensity (in absolute values) of convergences and divergences between actors.

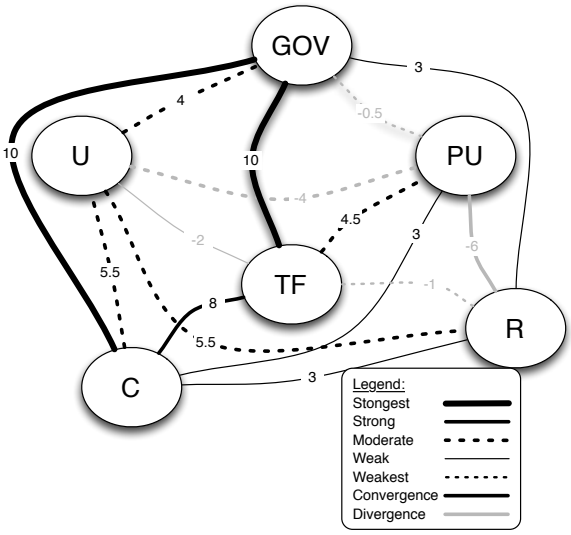

Fig. 6: Actor Interaction Map

Discussion

The aim here was to determine to what extend inter-stakeholder relationships shape not only the OD project, but also its outcomes. The findings have shown that the actors that control key variables demonstrate a relatively high degree of agreement between themselves and towards the objectives. The MAA, visualising the average intensity of conflicts between actors, shows that conflicts exist (especially between the Research/University and the

Publisher). But it also illustrates that they are clearly outbalanced by convergences. The government, as primary decision-maker, is in a position of relative assertiveness of its own objectives. Interestingly, its positions seem to be more in line with the Technology Firm than with the University. This might be the case because the government's and the firm's rhetorical input could have been pre-decided.

We are inclined to attribute this solid degree of convergence between actors to potential benefits that all of them expect in the context of the case study (except the Publisher). They exhibit willingness, because they see the need for structured, linked, and accessible data concerning publicly funded research. In this regard, particularly the increased quality of the data – likely to originate from the coordination between stakeholders on all scales - seems a good incentive. Universities and researchers, for instance, need to invest some effort into providing their data in appropriate format and quality, but benefit in return in terms of gained organisational efficiency.

Step 2 of the MACTOR analysis identified foremost that most stakeholders hold a strong position in the ecosystem, according to influences/dependencies. This entails, on the one hand, in correlation with the high degree of convergences that the project has good chances for success. On the other hand, it means that most of the stakeholders can impede outcomes of the project. The notably strong positions of private actors require particular attention. In addition, they score highest in terms of their competitiveness, i.e. the chances to enforce their objectives. In other words, the ecosystem is highly dependent on the collaboration of companies. The strong position of the Publisher is self-explanatory, as all research output published via its channels is legally owned by it.

Still, interdependencies seem to matter, possibly because project objectives are similar and all realise how important cooperation is to reach them. An interesting related aspect is that, even though the relationships between University, Researcher and Publisher are marked by divergences, they are the most interdependent actors in the ecosystem. Even more so as some universities begin to implement SciVal[5], a research management software offered by the academic publisher Elsevier. In how far this decision will impact the OD development, as promoted by the project, remains to be seen. However, indirectly entrusting this information to a company that already owns the output of the work of researchers seems to be in stark conflict with expectations related to OD. This illustrates that stakeholders are indeed rather dependent on private actors in the ecosystem.

Most of the concrete objectives, which have been established for the project, seem to be characterised by agreement between stakeholders, such as the creation of public value. However, objectives 3 and 5 are more contested. Not by coincidence, these are decisive for the sustainability of the overall initiative and potentially entail main bottlenecks for business models.

Objective 3 (Data Governance) demands safeguards that the data stays public (i.e. open). For the Technology Firm and the Publisher this seems to be more problematic than other objectives. The objective presents an issue as difficult as it is significant, also beyond the case study; companies work with the data as 'raw material', which does not belong to them. They invest, refine it and thus add value, but cannot own it. This is the reason why open data licenses are highly decisive for the functioning of a balanced ecosystem. OKFN's open

[5] http://www.elsevier.com/online-tools/research-intelligence/products-and-services/scival

database license and especially the Creative Commons 4.0 license suite[6] are said to be highly functional in this regard.

The question of which stakeholder owns the data is of even higher significance, due to the powerful position of the private actors in the ecosystem. Outsourcing research management technology to an actor like Elsevier as described above could entail that data, which was supposed to be open and accessible, becomes restricted in the absence of clear legal arrangements.

Objective 5 (Risk Distribution) mainly deals with the investments that the public sector has to carry for the initiative to function. It was mentioned before that today the public carries a major share of investments for LOD. The citizen as taxpayer prefers to distribute the risk over the whole ecosystem, especially because most people will not see any need for OD. This is understandable, especially since the true impact of the movement is not really measurable yet, due to its rather short existence. This presents a significant dilemma, as the lack of uncertainty about potential returns also prevents many companies to create business models based on OD.

Private actors demand the Government to carry the risk, at least to demonstrate that the hailed economic potential is real. This then creates incentives for companies to get involved. Thus, the Government must take a leading role and invest to get things started, even though the return on public investment is far from clear. Because without such investments, any chance for sustainable OD ecosystems generating public value will be oppressed before it even started. In other words, governments need to motivate companies, citizens, SMEs and developers to make use of the data.

Objective 6 (Options for OD Business Model) is straightforward (although the outsourcing of roles in Objective 2 is not agreed upon by all stakeholders). In fact, the current situation suggests that commercial incentives add substantially to the viability and sustainability of O(G)D-related initiatives. The examples of enablers such as Socrata, GTFS and SciVal, or re-users such as Stat.io and Mapbox (see Chapter II) demonstrate that.

This leads to Objective 4 (Accessibility). All stakeholders naturally agree strongly on this aspect, except the Publisher, whose business models are based on limiting accessibility. This is of utmost significance, because publishing does not generate impact directly. Creating accessibility matters even more. And here, the structural or value-adding roles of actors outside the public sector are irreplaceable.

V. Conclusion

The question leading to the analysis above concerned the interactions and interdependencies between actors in open data ecosystems. Relations of power are based on roles that allocate control to an actor; the stronger a role, the more an actor is in a position to actually shape results and outcomes of respective initiatives. The case study and corresponding analysis confirm that, alongside the central position of the government, the private players occupy strong and influential roles, being able to impact the important objectives. It has to be noted that, at this stage of research, this ranking of positions was made based on initial inputs and needs to be validated by stakeholders as a next step. Therefore, the findings should not be

[6] http://creativecommons.org/version4

regarded as definitive, but rather as indications and suggestions to consider for all OD initiatives, which intend to go beyond merely publishing governmental data.

Possibly the most important aspect in this regard is formal data governance that can ensure that the data remains under public ownership, i.e. truly open. This needs to be regulated and secured from the very beginning, as well as communicated in a fashion that demonstrates that this is not necessarily a real barrier for the creation of business models. Even if the data is public, there are not only manifold possibilities, but also a necessity for commercial incentives to contribute in the ecosystem. Still, it is clear that the public sector has to take a lead concerning open data, notably in form of investments. The commercial potential of open data alone will in many cases not suffice as incentive for companies to take the risk. On the other hand, companies need to be involved, not least because the facilitation of access to as many citizens as possible is most significant.

This is not necessarily a bad thing, especially because the government can be in a position here to shape the foundations of open data culture in this fashion. An important aspect in this regard is also public procurement: due to companies' dependence on the government as a substantial source of money, it has a powerful tool for shaping respective processes.

In the future, it will be most important to study true impact of open data initiatives. So far, the movement is still to fresh to produce many tangible changes in society. After answers are established in this regard, the true public value and realistic returns on public investment can be determined, as well as true economic potential.

Acknowledgment

The research leading to these results has been performed within the project "Een Experimenteel Open Data Platform Voor Vlaanderen" and received funding from the Flemish government Department of Economy, Science, and Innovation (EWI).

The software that facilitated the analysis was created and provided as an open-source tool by Michel Godet. It was downloaded from http://en.laprospective.fr/.

Annex

The key steps involved in performing the analysis are explained in the form of the flowchart in Fig.7. They are implemented in details for the project's OD ecosystem, in the Multi-Actor Ecosystem Analysis in chapter IV.

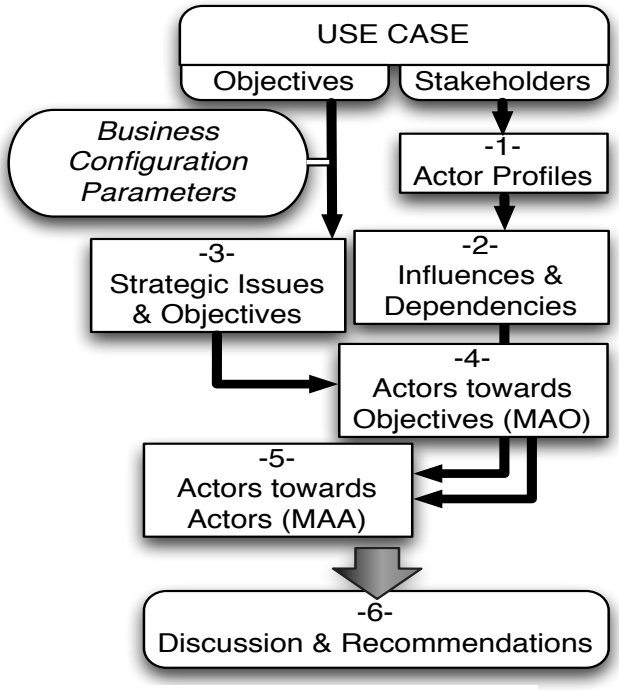

Fig. 7: Steps of MACTOR analysis

References

Ballon, P. (2007). Business modelling revisited: the configuration of control and value. *Info*, *9*(5), 6–19. doi:10.1108/14636690710816417

Ballon, P., Walravens, N., Spedalieri, A., & Venezia, C. (2010). The Reconfiguration of Mobile Service Provision: Towards Platform Business Models. In M. Falch & J. Markendahl (Eds.), *Promoting new telecom infrastructures: markets, policies and pricing* (pp. 197–216). Cheltenham, UK; Northampton, MA: Edward Elgar.

Berners-Lee, T. (2006). Linked Data Design Issues. Retrieved March 18, 2014, from http://www.w3.org/DesignIssues/LinkedData.html

Buchholtz, S., Bukowski, M., & Śniegocki, A. (2014). *Big and open data in Europe - A growth engine or a missed opportunity?* Warsaw Institute for Economic Studies (WISE Institute).

Chesbrough, H., & Rosenbloom, R. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. *Industrial and Corporate Change*, *11*(3), 529–555. doi:10.1093/icc/11.3.529

Chui, M., Farrell, D., & VanKuiken, S. (2013). Generating Economic Value through Open Data. In B. Goldstein (Ed.), *Beyond Transparency-Open Data and the Future of Civic Innovation* (pp. 53–78). San Francisco: Code for America Press.

Directive 2013/37/EU. Directive 2013/37/EU of the European Parlament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the Re-use of Public Sector Information. [2013] OJ L175/1. Retrieved from http://eur-lex.europa.eu/

European Commission. (2011). PRESS RELEASE - Digital Agenda: Turning government data into gold. *Press Releases*. Retrieved November 18, 2013, from http://europa.eu/rapid/press-release_IP-11-1524_en.htm

Ferro, E., & Osella, M. (2013). 8 BM archetypes for PSI re-use. Istituto Superiore Mario Boella. Retrieved from http://www.w3.org/2013/04/odw/odw13_submission_27.pdf

G8. (2013). G8 Open Data Charter and Technical Annex - GOV.UK. Retrieved from https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex

Godet, M. (1991). Actors' moves and strategies: The mactor method. *Futures*, *23*(6), 605–622.

Godet, M. (2008). *Strategic Foresight/La Prospective - Use and Misuse of Scenario Building*. Paris: Cahier du LIPSOR. Retrieved from http://innovbfa.viabloga.com/files/LIPSOR Strategic Foresight.pdf

Goedertier, S. (2013). *Study on business models for Linked Open Government Data* (BM4LOGD). Retrieved from https://joinup.ec.europa.eu/community/semic/document/study-business-models-linked-open-government-data-bm4logd

Goldstein, B., & Dyson, L. (Eds.). (2013). *Beyond Transparency-Open Data and the Future of Civic Innovation*. San Francisco: Code for America Press.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, Adoption Barriers and Myths of Open Data and Open Government. *Information Systems Management*, *29*(4), 258–268. doi:10.1080/10580530.2012.716740

McHugh, B. (2013). Pioneering Open Data Standards: The GTFS Story. In B. Goldstein (Ed.), *Beyond Transparency-Open Data and the Future of Civic Innovation*. San Francisco: Code for America Press.

Milicic, V. (2011). Problems of Linked Data (3/4): Publishing data. Retrieved from http://milicicvuk.com/blog/2011/08/02/problems-of-linked-data-34-publishing-data/

Moore, J. F. (1993). Predators and prey: a new ecology of competition. *Harvard Business Review*, *71*(3), 75–86.

Networked & Electronic Media Initiative. (2013). NEM Position Paper - Big and Open Data. Retrieved from http://www.nem-initiative.org/documents/position-papers.html

Okfn. (n.d.). Open Definition. Retrieved August 30, 2013, from http://opendefinition.org/

Rosseau, B. (2013). Interview with Bart Rosseau, e-strategist of city of Ghent.

Sheridan, J. (2012). Legislation.gov.uk: Putting APIs first. *Government Digital Service*. Retrieved November 21, 2013, from http://digital.cabinetoffice.gov.uk/2012/03/30/putting-apis-first-legislation-gov-uk/

Stat.io. (2013). Stat.io. Retrieved November 21, 2013, from http://www.stat.io/

Vickery, G. (2011). *Review of Recent Studies on PSI Re-use and Related Market Developments*. Paris.

Walravens, N. (2012). Mobile Business and the Smart City: Developing a Business Model Framework to Include Public Design Parameters for Mobile City Services. *Journal of Theoretical and Applied Electronic Commerce Research*, *7*(3), 21–22. doi:10.4067/S0718-18762012000300011

Yin, R. K. (1981). The case study crisis: some answers. *Administrative Science Quarterly*, *26*(1), 58–65.

Zuiderwijk, A., & Janssen, M. (2013). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, *31*(1), 17–29. doi:10.1016/j.giq.2013.04.003