Open Data and the Core Competences of Government: Lessons from Flanders, Belgium

Dr. Nils Walravens
imec-SMIT, Vrije Universiteit Brussel
nils.walravens@imec.be

Dr. Pieter Ballon
imec-SMIT, Vrije Universiteit Brussel
pieter.ballon@imec.be

Mathias Van Compernolle
imec-MICT, Ghent University
mathias.vancompernolle@imec.be

Dr. Pieter Colpaert
imec-IDLab, Ghent University
pieter.colpaert@imec.be

Dr. Lieven De Marez
imec-MICT, Ghent University
lieven.demarez@imec.be

Abstract: *As part of the rhetoric surrounding the Smart City concept, cities are increasingly facing challenges related to data (management, governance, processing, storage, publishing etc.). The Smart Flanders program was initiated by the Flemish Government (Belgium) at the start of 2017 to research and support cities in this field. The goal of the program is to support the 13 so-called center cities in the Region (by and large the biggest cities) and a representation of the Flemish Community in the Brussels Region with defining and implementing a common open data policy. As part of the program, a "maturity check" was performed, evaluating the cities on a number of quantitative and qualitative parameters. This exercise laid to bare a number of challenges in the field of open data and led to a checklist that cities can employ to begin tackling them.*

Keywords: *open data; smart cities; e-government; co-creation*

1. Introduction and context

The Smart City concept has been popularized (and hyped) for over at least a decade but turning its promises into practice remains a challenge for cities today. Most agree that technology has *some* role to play in supporting or implementing policy, but how that role should be filled remains unclear and often the result of trial and error. The Smart City concept has also been thoroughly criticized, *inter alia* for its self-congratulatory tendency, the vast commercial interests at play, as well as its push of ICT and the potential consequences towards reinforcing a digital divide (Graham, 2002; Hollands, 2008). Handing over too much control over the public domain to private companies raises concerns regarding democracy and the commodification of the public space (Greenfield, 2013; Townsend, 2013). Both are a far cry from what would be labelled as 'smart'.

At the same time, city governments are exploring what the concept can actually contribute to their daily practices and which role technology can play in providing better or 'smarter' services to citizens. Even the staunchest critics of the Smart City concept agree that *data* increasingly has a role to play in policy making (Hollands, 2008). While this of course has always been the case to more or lesser extent, the sheer amount of data that is becoming available today, as well as the combination of data from different sources and domains, can provide new types of tools and insights to policy makers. This can be data that comes from Internet of Things solutions (e.g. sensors in public parking garages), structured information in internal reporting systems, detailed data on the public domain (e.g. from satellite imaging) and so on.

In order to fully unlock the potential of this data however, it needs to be more easily available and accessible than today. This is where open

data comes in. The idea is that governments currently own (but do not use) a wealth of information related to divergent aspects of life in the city, but that this data is neither publicly available, nor easily interpretable. This has sparked a movement to encourage the opening of datasets in a structured and machine-readable way, under the 'open data' moniker, which has gained significant traction across local and national governments. The Open Knowledge Foundation is one of the strong proponents of open data and has come up with what has become the generally accepted definition of open data: *"Open means anyone can freely access, use, modify and share for any purpose (subject, at most, to requirements that preserve provenance and openness)"* (OKFN, 2015). This means that open data can be used for any goal at no cost, with the only (potential) exceptions being that reusers mention the source of the data or do not in any way prevent the data from being shared further on.

The idea here is clear: public organizations open up all kinds of data related to their operations, with the goal of having external developers create new services and applications ('apps') based on this data. In principle, this can mean a cost reduction for the public organizations that open data, as they do not need to build and maintain their own services and apps, an activity that is generally accepted as being highly cost-intensive (Walravens, 2015).

In practice however, a number of challenges remain and 'merely' opening up data has not always proven equally successful (see e.g. Peled, 2011; Lee et al., 2014). Opening up data already entails significant challenges to governments and public organizations before any data "leaves" the organization (e.g. setting up internal processes to safeguard internal data hygiene and quality control or implementing new or updating existing database systems). Relevant data can also be distributed over different government organizations or levels of governance, and some data applicable to the public may be under the control of private players that are less inclined to open it. After data are made available, the role of government is not necessarily played out. Ensuring that data is actually reused and relevant applications are

built should also be considered a concern for these public organizations and open data policy makers.

In order to tackle some of these challenges, the Smart Flanders program[1] was initiated by the Flemish Government (Belgium) in early 2017. Smart Flanders is coordinated by IMEC, the largest nonprofit technology research institute in Belgium, by an interdisciplinary team of researchers from communication sciences, organizational science and computer science. The goal of the 3-year program is to support the thirteen so-called center cities in Flanders (by and large the biggest cities) and a representation of the Flemish Community in the Brussels Region (referred to as the 13+1), with defining and implementing a common open data policy. The program is followed up by a steering group consisting of representatives of the cities, the cabinets of the Flemish ministers for Urban Policy and for Innovation, the Flemish agencies responsible for Interior Policy and Information, the Knowledge Centre Flemish Cities, the Organization of Flemish Cities and Towns, and IMEC.

To achieve the goal of defining and implementing a joint open data policy, these cities needed to find common ground and collaborate in ways and on themes that were quite new to them. This paper will present some of the most significant challenges at play when it comes to open data in a city context today. It will summarize these points of attention in an Open Data Checklist that cities may reuse to assess their "open data readiness".

2.   Methodology

In order to establish a state of the art around the topic of smart cities, a thorough, written, open questions survey was conducted with the cities. This survey asked the participating cities how they looked at the Smart City concept, whether and how they currently organize around it, how they spend resources on Smart City projects and how they think about technology and data. The survey also aimed to document whether any smart city policies were already in place and what these may entail.

---

[1] https://smart.flanders.be (Dutch only for now)

This initial written survey was then complemented by a round of in-depth expert interviews with representatives of the 13+1 cities. These semi-structured interviews allowed us more insight into the motivations, concerns and challenges raised by trying to establish a smart city strategy. Fourteen interview sessions were held between April and October of 2017, with multiple representatives of the cities present. The profiles that participated to the interviews range from politicians, civil servants responsible for data management, ICT, geographical information, local economy, mobility and so on. Representatives from the following cities were interviewed: Aalst; Antwerp; Bruges; Genk; Ghent; Hasselt; Kortrijk; Leuven; Mechelen; Ostend; Roeselare; Sint-Niklaas; Turnhout and the Flemish Community Commission in Brussels. The interviews lasted between two and four hours and were transcribed for analysis. The data gathered in 2017 (Van Compernolle et al., 2017) is currently being updated during a new round of interviews taking place in Summer 2019. Where possible, we will complement the analysis with this new material. Later publications will focus on these new results and the evolutions we can derive from them over a two-year period.

Based on the insights coming from both this quantitative and qualitative data, a number of critical aspects were identified that cities can actively work on, with the goal of making a smart city and open data strategy more concrete. It became clear a number of general challenges remain when it comes to implementing sound open data policies. These challenges came to the foreground during the Smart Flanders steering group meetings and were shared via the website[2] to generate wider debate (in Dutch). The following section will present and discuss these challenges.

3. Challenges and questions related to (open) data policies

This section will outline four key challenges that need discussion and answers if cities want to establish more sustainable open data policies.

Summarized, we will deal with the following topics:

- Data hygiene in the organization: In some cases, digitization still is a significant challenge, but how can we generate awareness to the level of key individual public servants that work with data? How do we change working with data into an operational process that leads to good open data?
- IoT and open data: In the hype surrounding the Smart City concept a lot is made of the data generated by sensors and other IoT devices, but how do we publish data from these sensors in a proper way, dealing with the real-time aspect, the sheer volume of the data, archiving of data and so on?
- Centralize vs decentralize: As a principle, open data lends itself quite well to decentralized publishing and the technical solutions are available, but how do we turn these into processes that work? This requires agreement on the roles of different levels of government.
- Government and the market: where does the role of government end? When do private actors come into play? This is particularly relevant in the field of open data as well.

3.1. Data hygiene in the organization

The first challenges for most of the cities that were interviewed still relate to the digitization of internal processes and services towards citizens. This also entails having processes and procedures in place when it comes to working with data in the organization. It may seem counterintuitive but open data can actually offer significant short-term efficiency gains in this regard. By reusing data from other organizations or departments within the city, public workers can avoid wasting time looking for the most recent or complete information. This does however require that everyone in the organization that needs to work with data is aware of the importance of doing this in a structured, traceable and repeatable way. That also means a data management plan at the level of the whole organization becomes an important tool to manage these processes. Very

---

[2] https://smart.flanders.be

often, this is not or only partially present in the interviewed cities. It is however recognized as being of key importance and is under development in almost all cases. Keeping data hygiene within the organization under control and at a high level is a first long-term challenge and requirement to implement a sustainable open data policy.

Cities also recognize that interoperability will increasingly be of great importance in this context. Making clear agreements on the ownership, use and publishing of data will only grow in importance, but it requires an investment on the part of the organization to ensure sufficient technical expertise and to make the right decisions in this complex area. Interoperability and the concept that data and applications *can* be seen as separate from each other should prevent data becoming "locked up" in applications provided by third-party vendors. Avoiding so-called vendor lock-in means that the relationship between a local government and its service suppliers can evolve from a typical client-supplier relationship into a partnership in which data are easier to move from one system to another when this is needed or desirable.

3.2. IoT and Open Data

The Internet of Things (IoT) is often mention in one breath with Smart City services and can mean an extra complicating factor when viewed from the perspective of the data these systems generate. The concept links to the idea that we can understand reality better by measuring as much as possible and by equipping the public space with all kinds of sensors that collect different types of data, policy can be informed by more evidence than ever before. Policy could be tailored to what is observed in the public space, even in real-time.

However, the idea of data-driven policy making comes with a number of complexities on different levels. Divergent actors need to collaborate in new ways and in new fields. One real life example from Flanders is using ANPR cameras to enforce a low emission zone in a city in which certain types of polluting vehicles are not allowed or need to pay a fine when they enter the zone. The sensors in this case are the smart cameras that can detect license plates and determine whether a car can enter the low emission zone or if a fine needs to be sent. To enable this, an elaborate collaboration between different actors needed to be realized, as data needs to be shared between different government organizations, police databases, companies deploying the infrastructure (the cameras in this case) and related software platforms, citizens who need to be informed about which types of cars can enter in the zone during which period and so on.

Next to the often-complex forms of collaboration or partnership between diverse actors, processing all the data generated by IoT solutions is another significant challenge. Clearly, when more sensors are deployed in the city, the amount of data these systems generate increases dramatically. All this data needs to be processed, a task often given to the third-party vendor supplying the solution, but what remains often unclear today is if and how the collected data should be archived. Historical analyses can yield very interesting insights to inform public policy or even allow for predictive analytics, but how long should these large datasets be stored? After which time period should data be erased, especially if personal information is included? Who is responsible for storing and providing access to the data? Who pays for these services? It is important to consider these questions when procuring IoT solutions from third parties and including these arrangements in contracts and agreements. Very often, this is not the case today.

Finally, and to the core of this paper, a significant challenge related to IoT data is how to publish this data for reuse in a sustainable and cost-effective way. In the spirit of open data, providing potential reusers with real-time information coming from IoT solutions has the potential to generate all kinds of innovative services and applications (e.g. in the domains of mobility, air quality crowdedness and so on). This means however, that infrastructure needs to be made available to allow for a swift processing, publishing and archiving of said data. Some solutions are available today, but they are often tied to a single vendor or solution. Furthermore, with the speed at which more IoT data is becoming available, this challenge will quickly become more prevalent and need to be addressed sooner rather than later.

3.3. Centralize vs decentralize

Another pertinent challenge or question in the field of open data relates to the way data are published and which actor takes up which role. The question should be framed in a broader debate on centralizing data versus decentralizing them. What remains crucial is that data are easy to find and use for potential reusers. The success of any open data policy will depend on this. Hence, it is important that a local government communicates about the data it makes available, but also that the data can be easily found be anyone looking for it (e.g. also from abroad). When data are published in a decentralized way, for example on the website of the municipality, it is important to describe the data according to standardized principles. By applying standards (like DCAT for example) to describe data, information about that data can automatically be picked up by regional, national and international open data portals, making them easily retrievable by anyone looking to reuse them (including commercial data portals such as Google Dataset Search for example).

Publishing data in a completely decentralized way is technically possible but entails a number of organizational challenges. Clear agreements need to be made about the standards used, the ways in which they are applied and the processes that need to be put in place to ensure data is published in the proper way, for example on a municipality's website. This requires a significant investment by local governments and since open data is rarely a priority, this remains a challenge. Additionally, the resources and skills required are not always present, particularly in smaller organizations. For them, a more centralized approach will prove far more sustainable.

The question then becomes who should take up the role of supporting smaller local governments with this challenge. In Belgium, because of its complex and federated structure, the regional Flemish government, provincial government or intercommunal organizations could take up this role. Larger cities could take up some of the investment to support the smaller municipalities in their region. And new forms of collaboration between local governments are also coming to the foreground in different regions (e.g. around Brussels).

Today, none of these actors are clearly positioned to take up such a role, but it is becoming increasingly clear (and urgent) that more collaboration in this area is needed to set up more sustainable data (sharing) policies.

The first question related to the core competences of government is then; who does what and who has a clear mandate to enforce certain policies if necessary? Today, this situation is fragmented and unclear in Flanders and by extent, Belgium. A broad governance of the Flemish public data landscape should be developed and formalized as soon as possible in order to avoid further fragmentation and an inefficient use of public resources.

3.4. Government and the market

Next to the question of which level of government should take up which role, a second important question related to the core competences of government can be identified: which tasks should be for government and which should be taken up by private players? This is a political decision and choice for the most part and hence will evolve depending on dominant views at the time. As such, it is something of a moving target. This however does not mean this question should not be in the back of minds of policy makers, as a choice for "more" or "less" government can have consequences for the quality of service provision to citizens.

A key challenge in this area of balancing public and private interest in the context of open data relates to stimulating reuse of open data: should it be a task of government to ensure that data are actually reused? Most cities agree the local government has a role to play here, by (1) serving as the authentic source for published opened up data (2) ensuring data can be easily found and the threshold for reuse is kept as low as possible and (3) that local government engages in a dialogue with potential reusers so that the data that are published are relevant and of value for reuse. Since data are also made available for commercial reuse, it is not possible to exclude companies from this dialogue. A challenge then becomes how to avoid giving any company a competitive advantage (e.g. by giving them insight into available data or a roadmap for publishing

certain datasets). Transparency on both the process and result of a dialogue are crucial here.

Another challenge is the relationship between government and third party vendors: what are the options as a public organization in enforcing certain behavior from its suppliers? A number of basic demands can be included in the contracts between the two, are e.g. penalty clauses also foreseen ? What is the recourse when the systems of two vendors turn out not to be compatible even though this was ensured during the contracting phase and both suppliers point to each other? Very often, local government does not have the resources to engage in complicated lawsuits. There is no simple answer to these challenges, but the dialogue and transparent approach referred to in the previous paragraph can be part of the answer. Additionally, traditional procurement could be abandoned in some cases where innovative procurement allows for more flexibility on the part of the procuring organization.

A public organization is expected to serve the public interest. When working with and on data, this role becomes even more important, but also far more complex. More than ever, local governments should inform themselves on good practices in this field and clearly position themselves towards third-party vendors that promise the single solution to all of their challenges. By starting from a stronger base of information as well as some shared principles, local governments can evolve away from a traditional client-supplier relationship towards a partnership with market players. When it comes to open data, the role of government here is to strive for a maximal and broad reuse of data, through a transparent process and dialogue.

4.   Open Data Checklist

The survey and interviews with the 13 cities have led to a number of insights related to publishing open data, some of which were outlined in the previous section. To make these insights accessible for reuse by other (local) governments, they are presented as a checklist[3]

in what follows. Government organizations that are exploring open data initiatives can use this checklist to ensure to cover some of the most significant challenges related to publishing open data in a sustainable way. The checklist consists of 6 main categories:

- Problem (re)definition
- Capacity and resources
- Organizational culture
- Governance
- Partnerships
- Risks

In what follows, we will very briefly list points of attention in each of these categories.

4.1. Problem (re)definition

Frame context and cause: do not just open data to open data, but start from a clear and concrete policy challenge.

Define problem and goals: make the policy goal more concrete by establishing measurable KPIs. Open data will never completely solve a problem but can be instrumental in speeding the process along.

Do "reuser research": understand the needs and pains of potential reusers by engaging in a transparent dialogue.

Redefine the problem: evaluate the initially identified problem and do not hesitate to rescope or redefine it if necessary.

Create an overview of the data: understand which data are available within the public organization and who is responsible for them.

4.2. Capacity and resources

Build data infrastructure: publishing data means the basic data infrastructure needs to function well first. For smaller municipalities this cost can potentially be shared through intergovernmental collaboration.

Develop expertise: working with (open) data requires skills that are today not always present

---

[3] Inspired by http://odimpact.org/periodic-table.html under a CC BY-AS 4.0 license.

within public administrations. Training and knowledge building in this area is important.

Provide sufficient resources: open data requires an initial investment and a translation into processes within the organization. This requires sufficient means and personnel.

4.3. Organizational data culture

Apply shared principles: whenever possible strive for using shared frameworks so that all partners understand terminology in the same way.

Stimulate "believers": identify public workers in the administration that see the potential of open data and actively involve them in implementing a policy.

Be open for feedback: reusers of your data will provide you with feedback on data quality, availability and so on. The organization needs to be prepared to tackle constructive feedback.

4.4. Governance

Guard standards and data quality: a good internal data hygiene requires the use of standards to allow for easier and automated sharing, linking and exchanging of data.

Set roles and responsibility: clearly defining who does what within and outside of the public organization is key in ensuring efficient use of resources. This is perhaps the most important challenge facing local governments today.

Strive towards an agile and flexible organization: working with data and technology requires flexible processes to allow for corrections when needed.

Develop structured evaluation: foresee quantitative and/or qualitative KPIs to evaluate both process and outcome. This means including a baseline measurement as well.

4.5. Partnerships

Approach data owners: explore new partnerships with owners or relevant data to support policy challenges.

Involve domain experts: include the domain expertise present in the public organization to ensure data is described and applied in correct ways.

Involve organizations with similar goals: use the knowledge and expertise of like-minded organization, whether they be other local governments, departments within other levels of government, civil society, companies, research centers and so on.

Procurement: when procuring new solutions or renegotiating contracts with third-party vendors, include clauses related to data ownership, processing, storage and open data.

4.6. Risks

Privacy: develop privacy-by-design solutions and applications and include privacy impact assessments when publishing data. Open data per definition does not include personal data, however scenarios could be envisaged where the combination of open data results in the identification of individuals. An a priori privacy impact assessment can identify this.

Security and data management: as local governments start processing more data, security becomes increasingly important as well. A data management plan can support this, but may require external capacity and support.

Digital exclusion: open data initiatives should never lead to an exclusion of those who do not have the skills or access to public services.

Data quality and policy decisions: evidence-based policy can only be as good as the data that support it. Data quality and verification are thus of high importance, also when opening up. A guiding principle here can be that if data are considered of sufficient quality to be used internally for policy development, they should be of sufficient quality to open up.

"Open washing": this risk refers to a situation in which public organizations claim to open up, but only do so to comply with regulations. This is not a sustainable situation and waist of resources. Starting from a concrete case or project can avoid this.

5.   Conclusion

Opening data remains something of a chicken-and-egg problem: sufficient investment is needed on the side of government in order to publish significant amounts or relevant data, but reusers will only generate innovative applications and services once enough data are available. The research presented in this paper shows that cities certainly see the potential value of open data, but a number of challenges remain. In order to develop sustainable open data policies, a number of conditions have to be met. These have been summarized as points of attention presented in an Open Data Checklist in the previous section. Taking factors related to problematization, organizational culture, governance, partnerships and a number of risks into account can help local governments make more informed decisions when designing or developing an open data policy for their constituency.

6.   References

Graham, S., 2002. Bridging Urban Digital Divides: Urban Polarisation and Information and Communication Technologies. Urban Studies, 39 (1), pp. 33–56.

Greenfield, A., 2013. The City is Here for You to Use. Wired, 5 February.

Lee, M.J., Almirall, E. and Wareham, J.D., 2014. Open Data & Civic Apps: 1st Generation Failures–2nd Generation Improvements. *ESADE Business School Research Paper*, (256).

Hollands, R., 2008. Will the Real Smart City Please Stand Up?. City, 12 (3), pp. 303-320.

OKFN, 2015. Open Definition. Retrieved from http://opendefinition.org

Peled, A., 2011. When transparency and collaboration collide: The USA open data program. *Journal of the American society for information science and technology*, *62*(11), pp.2085-2094.

Townsend, A., 2013. Smart Cities. New York: Norton & Company.

Van Compernolle, M., Waeben, J. & N. Walravens, 2018. *Eindrapport Smart Portrait*. Public Report for Kenniscentrum Vlaamse Steden and Agentschap Binnenlands Bestuur. Available online: http://www.kenniscentrumvlaamsesteden.be/overhetkenniscentrum/Documents/Eindrapport%20Smart%20Portrait_PUBLIEK%20-%20definitief.pdf

Walravens, N., 2015. Mobile city applications for Brussels citizens: Smart City trends, challenges and a reality check. *Telematics and Informatics*,*32*(2), pp.282-299.