Public bodies’ access to private sector data: The perspectives of twelve European local administrations
by Marina Micheli

Abstract
Public bodies’ access to private sector data of public interest (also referred to as business-to-government (B2G) data sharing) is still an emerging and sporadic practice. The article discusses the findings of a qualitative research that examined B2G data sharing in European local administrations. Drawing from semi-structured interviews with managers and project leaders of twelve municipalities, the study contextualizes access to private sector data in the perspectives of those working in the field. The findings examine the four operational models to access data that featured more prominently in the interviews: data donorship, public procurement of data, data partnerships and pools, and data sharing obligations. The analysis highlights the power unbalances embedded in B2G data sharing as perceived by representatives of local administrations. In particular, the findings address the gap between municipalities in the opportunities to access private sector data of public interest, the lack of negotiating power of local administrations vis-à-vis private sector data holders and the strategies envisioned to foster more inclusive forms of data governance.

Contents
Introduction
B2G data sharing: Variables, obstacles and operational models
Methodology
Results
Discussion
Conclusions

Introduction
Data from the private sector could offer valuable opportunities to public bodies in supporting their efforts to address both short term and long-term societal challenges. Access to data collected by businesses — such as telecommunication companies, online platforms, transport services, accommodation Web sites, energy providers and financial operators — can improve situational and causal analysis of societal challenges, thus supporting more focused and evidence-based policy-making. B2G data sharing usually concerns information collected by private companies as a by-product of providing their main services (such as behavioral data and digital footprints) shared with public sector agencies and institutions operating at
Public bodies' access to private sector data: The perspectives of twelve European local administrations

Examples of B2G data sharing range from statistical authorities accessing sources of private-sector data to improve quality of official statistics, to commercial banks sharing financial-transaction data with intergovernmental organizations to understand how people behave during natural disasters (European Commission, 2020b).

This article focuses only on one kind of public sector bodies: local administrations. Typical examples of B2G data sharing involving local administrations are city authorities accessing data collected by ride-sharing companies and mobile phone operators to develop statistical models to better plan and manage urban mobility and related infrastructures (Susha, et al., 2017; European Commission, 2020b; Martens and Duch-Brown, 2020). Although desirable according to policy-makers, data sharing between commercial entities and local governments occurs rather infrequently (Susha, et al., 2019; Verhulst, et al., 2019; Martens and Duch-Brown, 2020; European Commission, 2020b). Local administrations attempt to access private sector data of public interest adopting various approaches, but for the most part they do so in pilots or experimental projects, eventually involving other actors (e.g., research institutions, civil society organizations, start-ups) (Sarasa Funes, 2017; Hardinges, 2019; European Commission, 2020b). The actual practices adopted by local administrations to access private sector data are far from the overstated vision of the ‘data city’ that dominates the public imagination but are highly relevant from a socio-technical and policy perspective (Shelton, et al., 2015; Madsen, 2018; Löfgren and Webster, 2020; Calzada, 2018; Cardullo and Kitchin, 2018; van Zoonen, 2020).

Current debates on digital innovation in cities tend to adopt either an “engineering perspective” — how can the use of new data sources and innovative technologies be fostered? — Or a “critical perspective” — what are the risks associated to increased data use by municipalities? [1]. An additional line of inquiry consists instead in scrutinizing the “actually existing” (Shelton, et al., 2015) practices situated in specific contexts, enquiring how “ideals” are brought into “practice” (Madsen, 2018) and emphasizing the “complexities” underpinning innovative forms of data sharing (van Zoonen, 2020). Informed by the latter line of inquiry, the article increases understanding of B2G data sharing in Europe drawing from an empirical study based on semi-structured interviews with city’s managers and project leaders working in the field of data and/or urban innovation (e.g., chief data officers) in twelve countries. The study collected experiences and opinions about the operational approaches adopted (or considered) by local administrations for getting access to private sector data of public interest. Through the analysis, it identified four prevalent models of access and related perspectives.

The study focuses specifically on the European context in which several data governance initiatives are currently being prepared (European Commission, 2020a). A set of principles for B2G data sharing was initially published by the European Commission (2018). More recommendations were included in the final report of the High Level Expert Group on Business-to-Government data sharing (European Commission, 2020b) and a specific regulatory intervention (Data Act) is forthcoming as a part of the European Strategy for Data (2020a). B2G data sharing enhances the re-use of data across actors and sectors, which is one of the key tenets of European data policy (European Commission, 2020a, 2020b). It falls into a wider tradition of policies that foster a “data sharing culture” in Europe, such as through public sector innovation and open data (Huyer and Cecconi, 2019). The theme has also a geopolitical connotation: access to private sector data might improve public sector’s strength and competency in the current digital transformation, while fostering the European digital single market (European Commission, 2020a, 2018). Therefore, although the study is not representative for Europe, the European setting, with its values and contextual features, might explain why the data practices (and related perspectives) unfold in the way they do in these local administrations.

After clarifying the motivations for the study in the current introduction, the next section provides an overview of B2G data sharing, including a review of the different operational models through which it can be implemented. The third section presents the research methodology and its theoretical underpinnings. The fourth section illustrates how local administrations in twelve European countries experienced and perceived four operational models for B2G data sharing. The discussion focuses on the implications of the different operational models for what concerns inequalities in opportunities to access data, and control of modalities through which private sector data of public interest is shared. The conclusions set out the limitations and
B2G data sharing: Variables, obstacles and operational models

Increasing volumes of data about citizens, including “passively generated” data created as a by-product of human actions online and offline, do not belong to national statistical offices or governments, but to various entities in the private sector (e.g., digital platforms, telecommunication operators, utility companies, retailers and other commercial actors). Public authorities have gained interest in the information such data assets provide and have established different kinds of partnerships or agreements in order to get access to it. Such data can only be shared under certain conditions, in a controlled form that accounts for a range of legal, technical and operational factors. The specific way through which companies share data with public bodies differs according to many variables and it is implemented through a variety of operational models (Susha, et al., 2017; Huyer and Cecconi, 2020; European Commission, 2020b; Verhulst, et al., 2019).

A key variable is the level of detail and the granularity of data, which defines the possibilities of public bodies to use it and derive meaningful information from it. Data can be shared at very different level of details: raw (as it is collected from the source), pre-processed (e.g., cleaned, re-sampled, normalised), processed (aggregated and combined) or packaged in data-driven insights resulting from internal analysis (European Commission, 2020b; Verhulst, et al., 2019). Usually, companies are more comfortable sharing data at its lowest level of granularity, such as data-driven insights instead of actual datasets as this allows keeping control of the shared information and reduces legal and economic risks (Shkabatur, 2019; European Commission, 2020b; Helderop, et al., 2019).

A further key issue is the “level of engagement” of the interested parties, intended as the degree to which data supply and data demand actors “co-design” how data assets are shared and used (Verhulst, et al., 2019). Companies might keep full control, for instance distributing only data-driven insights or allowing data to be used for pre-defined purposes in the context of particular events (Verhulst and Sangokoya, 2015). On the opposite end, a private-sector data holder might have little direct involvement in the way data is used and exchanged, for instance if it is subjected to data sharing obligations (Huyer and Cecconi, 2020; European Commission, 2020b). A third way is cooperation: private companies and public bodies collaborate throughout the whole process of data sharing jointly deciding which data is needed and how it is analyzed (Verhulst, et al., 2019).

The “level accessibility” of private sector data is an additional issue. Data holders (especially ‘big tech’ corporations) might share certain information in an open access format, allowing the general public to view or download datasets, through application programming interfaces (APIs), online public interfaces, or prizes and challenges (Verhulst and Sangokoya, 2015). On the other hand, business might allow only restricted access to pre-selected partners, for instance through data sharing agreements (Verhulst, et al., 2019).

B2G data sharing at the city level is still very infrequent because of a series of obstacles (Susha, et al., 2019; Verhulst, et al., 2019; Martens and Duch-Brown, 2020; European Commission, 2020a, 2020b). Companies are hesitant or refuse to allow local governments to access their data on the grounds of commercial confidentiality and security concerns (Helderop, et al., 2019; Susha, et al., 2017). The lack of appropriate regulatory frameworks partially explains private companies’ lack of willingness to share data (World Economic Forum, 2015; Taddeo, 2017; European Commission, 2020b).

Businesses face various uncertainties in relations to liability regimes (e.g., who is responsible if inaccurate or biased data is shared that leads to discrimination), intellectual property, and competition law (World Economic Forum, 2015; Martens and Duch-Brown, 2020). Furthermore, they encounter operational and technical challenges for the preservation of sensitive commercial information and the protection of customers’ personal information (Taddeo, 2017; Susha, et al., 2017). There are also economic barriers,
including monopolistic data markets (data holders companies often have the possibility to charge high prices for data), high transaction costs and lack of incentives for firms (Martens and Duch-Brown, 2020). Public bodies, on the other hand, face specific obstacles: lack of “culture” on data sharing (e.g., how to create public value from it), limited resources and capacity, deficiency of skills among public servants, and limited trust from both the private sector and citizens (GovLab, 2017; Giest, 2017; European Commission, 2020b).

The relations established between businesses and governments for data sharing hold some similarities with public-private partnerships (PPP) through which private companies contract services to public bodies. Yet, B2G data sharing can also fundamentally diverge from PPP depending on the operational model put in place (European Commission, 2020b). Firms voluntarily engaging in B2G data sharing are usually driven by economic interests. As for PPP, through B2G data sharing companies reinforce their position in the market and vis-à-vis public entities. An operational model that exemplifies such economic interests is data procurement that is when public bodies purchase private sector data of public interest through ad hoc contracts (European Commission, 2020b). Besides generating new sources of direct revenues, B2G data sharing might also serve companies’ economic interests by helping them to increase know-how and reputation. Operational models supporting such goals include data pools, that is when private and public entities create a pool of shared data and jointly analyze it for mutual benefits, and data donorship (or data altruism), which consists of private companies sharing data pro bono on a voluntary basis usually for humanitarian or social causes (Huyer and Cecconi, 2020; Verhulst, et al., 2019; European Commission, 2020b). Even when data is offered for “free” and in an altruistic manner, data sharing can generate power imbalances and dependencies. For instance, national governments in lower-income countries are increasingly relying on private-sector data, often offered at no cost by Telcos and big tech corporations, to gain information about social and economic trends. This has implications for sovereignty since it generates new visibilities and “states are not — or are only partially — in charge of collecting and disseminating (...) representations of themselves” [3].

B2G data sharing, on the opposite hand, can pursue public benefit first and contribute democratizing the governance of data, redistributing value and tackling power asymmetries (Morozov and Bria, 2018; Micheli, et al., 2020; Mercille, 2021). Examples are data sharing obligations, clauses that local governments include in their tenders for subcontracted services. The clauses demand data collected by a company as a by-product of delivering a public service is made accessible to the municipality (Bass, et al., 2018; Verhulst, et al., 2019; European Commission, 2020b). It builds on PPP, but it entails data is treated as ‘a public asset or infrastructure’ (Morozov and Bria 2018; Eurocities, 2019). Through such method, public bodies generate societal benefits from data collected by private companies, whose thriving often benefitted from public assistance/funding and infrastructures (Mazzucato, 2018; Couldry and Powell, 2014; Bass, et al., 2018; Heldrop, et al., 2019). A related approach is data sharing by regulation (or reporting obligations) when government authorities mandate access to information necessary to carry out their functions, such as for standard business reporting, environmental or social reporting (Klievink, et al., 2016). Data is required for monitoring the implementation of, or compliance with, existing legislation, but generally it is not reused or combined with other datasets for other purposes (Huyer and Cecconi, 2020; European Commission, 2020b). Therefore, depending on the operational model, B2G data sharing could either be an additional source of economic revenues for companies or a (data) relation that enables data use for the social good (Bass, et al., 2018).

B2G data sharing models can involve other actors besides local administrations and private companies. For instance, in the case of “research data partnerships” with academic and research institutions, “open APIs and public interfaces” available to civic society and the public at large, and “trusted data intermediaries”, in which external organisations act as trustees, offering a data sharing platform or conducting the analysis. However, overall the current knowledge on operational models for B2G data sharing at the city level is scarce, mostly based on the review of few prominent cases (Susha, et al., 2019; Susha, et al., 2017; Verhulst and Sangokoya, 2015).
Methodology

The study examines access to private sector data through semi-structured interviews with chief technology officers, chief data officers, or project leaders working at a city’s department in charge of urban digital innovation. The interviews have been conducted during 2019 and collected both practical experiences and perspectives regarding access to private sector data of public interest. The wide-ranging objective was to understand what kind of power relations are established between actors, in the process of data sharing, through the point of view of local administrations. Instead of collecting information about noteworthy use cases or “best practices”, the study collected information about the views of one of the actors involved in B2G data sharing.

A combination of purposive and snowball sampling procedures has been adopted for the selection of the participants in order to have a diversified group by city size, geographic location in Europe and previous experience in data innovation. In total, the managers/officers of twelve cities from 12 countries took part in the study. The participants are not representative for all European cities, however the group is diverse enough to bring to the fore a variety of practices and discourses associated to access to private sector data in Europe (Table 1).

<table>
<thead>
<tr>
<th>City</th>
<th>Country</th>
<th>Size</th>
<th>EU zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amsterdam</td>
<td>NL</td>
<td>Large</td>
<td>North West</td>
</tr>
<tr>
<td>Barcelona</td>
<td>ES</td>
<td>Large</td>
<td>South</td>
</tr>
<tr>
<td>Ghent</td>
<td>BE</td>
<td>Medium</td>
<td>West</td>
</tr>
<tr>
<td>Ljubljana</td>
<td>SI</td>
<td>Small</td>
<td>Central East</td>
</tr>
<tr>
<td>London</td>
<td>UK</td>
<td>Large</td>
<td>North West</td>
</tr>
<tr>
<td>Milan</td>
<td>IT</td>
<td>Large</td>
<td>South</td>
</tr>
<tr>
<td>Rennes</td>
<td>FR</td>
<td>Small</td>
<td>West</td>
</tr>
<tr>
<td>Rijeka</td>
<td>CR</td>
<td>Small</td>
<td>East</td>
</tr>
<tr>
<td>Tallin</td>
<td>EE</td>
<td>Mid-size</td>
<td>North East</td>
</tr>
<tr>
<td>The Hague</td>
<td>NL</td>
<td>Mid-size</td>
<td>North West</td>
</tr>
<tr>
<td>Vilnius</td>
<td>LT</td>
<td>Mid-size</td>
<td>North East</td>
</tr>
<tr>
<td>Zaragoza</td>
<td>ES</td>
<td>Mid-size</td>
<td>South</td>
</tr>
</tbody>
</table>

The interviews, which lasted 70 minutes on average, have been conducted remotely via a synchronous Internet connection. Cities strive to be seen as innovative, therefore I attempted to move the conversation beyond the façade by explicitly asking participants to discuss about challenges, failed attempts and what is not working well in the area, and by being self-reflective during the conversations. The transcriptions did undergo a qualitative thematic analysis. To assure the validity of the interpretation, I shared an earlier version of the article with the people interviewed and presented the findings during a workshop with
Public bodies' access to private sector data: The perspectives of twelve European local administrations.

The study is guided by the following research question: What can we learn from the perspectives of local administration representatives about the power relations established for getting access to private sector data of public interest? The question is informed by the notion of the ‘politics of data’, intended as the power relations embodied in data control and use (Ruppert, et al., 2017; Metcalfe and Dencik, 2019; Micheli, et al., 2020). Power is always exerted in the process of controlling access to data and significant power shifts might originate within data exchanges (Taylor and Broeders, 2015). Not only access to data has the potential to influence which decisions are taken and how a public issue is perceived, but also to empower (or disempower) certain social actors vis-à-vis their capacity to use information (Gurstein, 2011). With such premises, I explored the power relations that characterize different models adopted to access data through the lens of a specific type of actor from the public sector.

Results

All respondents stated that accessing private sector data was challenging because companies often have no interest in selling data or in sharing it with a municipality. The examples mentioned in the interviews are mostly pilot projects, activities at the “early stages”, if not still in preparation (“we are still figuring it out”). For a few participants the topic was rather novel, as access to private data was not part of the existing or planned activities of the municipality. The most relevant types of data providers, according to the study participants, were utilities companies, online platforms and telecommunication companies. On the matter of data protection, participants asserted they got access to aggregated data assets with no identifiable information. The interviewees saw privacy issues mainly as the necessity to respect the EU General Data Protection Regulation and they expected companies to comply with their responsibility by sharing only aggregated datasets that do not allow re-identification.

In this section I illustrate the key perspectives on the models for B2G data sharing mentioned most often during the interviews, which are: data donorship (addressed in n. 6 interviews), public procurement of data (n. 7), data partnerships and pools (n. 7), and data sharing obligations (n. 5) (see Table 2 for a summary of the results).

<table>
<thead>
<tr>
<th>Table 2: Summary of findings from interviews.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Role of local administrations</strong></td>
</tr>
<tr>
<td>Data donorship</td>
</tr>
<tr>
<td>Public procurement of data</td>
</tr>
<tr>
<td>Data partnerships</td>
</tr>
</tbody>
</table>
Respondents mentioned the opportunity to access commercial sector data at no cost, as companies/data providers occasionally make it available for free on a voluntary basis (data donorship). Private companies might engage in data donorship during an emergency or to address a humanitarian cause (Taylor and Broeders, 2015; Huyer and Cecconi, 2020). More broadly, however, data donorship comprises corporate social responsibility (CSR) initiatives performed through high-level intermediaries within data for development initiatives. The operational model was associated to a specific perspective across most interviews. Instead of being perceived as a philanthropic action, it was understood as a marketing strategy private companies used to increase their reputation and to build marketable use cases. Companies share data with cities at no cost — and eventually collaborate with them to develop products or services valuable for the municipality — as a means to develop new data-driven services to be sold to other cities in the future. The following quote, from a representative of a major European capital city, highlights that no effort was needed on their part to gain access to the company data, because the city reputation was sufficient to entice their interest. At the same time, it shows the volatile nature of such relations, as they take place in pilots only and are defined “incidental partnerships” elsewhere in the interview.

“A company will approach us and say ‘hey, we’ve got this ...’, but it will be on a pilot form only, because they want to say that they work with the, you know, the Mayor of cityX, in order to market their products in other places (...) In this case they were gaining some free promotion from these experimental samples (...) And we, yes, we didn’t pay them any money.” (city09) [4]

Companies use the reputation as “smart cities” as a means to advertise the data-driven products developed thanks to the collaborations. Thus, well-known cities (such as capitals or municipalities renowned for being innovative) are in a favorable position for being recipients of data donorship initiatives. The reputation and know-how of cities are incentives for private companies, as they help to develop and market new data driven services.

Such mechanism creates a double source of disadvantage for smaller/less well-known cities: not only they lag behind in terms of data innovation, but they have also to pay if they want to benefit from a service that other municipalities got at no-cost. In the first quote below, a respondent working in a well-known “smart city” describes this phenomenon as an ethical dilemma and explicitly questions the city’s position of privilege in relation to that of many others in Europe and elsewhere. Eager to address the issue, the respondent considered such asymmetries of power between local administrations worrisome. In the second quote, an interviewee highlights how those that are in a less favourable position perceive the same dynamic.

“[Companies] can say to other cities, ‘hey cityX did this use case, our data is very valuable, so our product is also more valuable, so you can pay more’. This is for us a way to work with these companies. But again, there is the ethical question,
do we want to have a free lunch if others are paying for it? (…) Ethically we are always a bit strained to use this.” (city02)

[During a meeting of a national group of cities] “He said, ‘Okay, for us, in cityX, the conditions under which we deal with the great companies is that we deal for nothing. They come and they develop some solutions, and we work together in partnerships, and it’s free for us’. And the other one in the room, they said, ‘Okay, it’s free for you, but it’s not fair. You have money, more than we have’. And when they get to us, they say, ‘Okay, we developed a solution with cityX’.” (city06)

Public procurement of data: Lack of power

Almost all interviewees discussed the possibility of acquiring data directly from private companies through procurement procedures, making it the most mentioned mode of access. Yet, there was a consensus that it was very far from the ideal one. Those who had experiences in public procurement of data, for instance, questioned its effectiveness. The main problems, according to the participants, consisted in the lack of control on the format through which data is shared and the quality of data itself. Public procurement of data occurs through sharing of actual datasets, APIs or with data-driven insights (such as reports or dashboards). In practice, however, companies often sell data-driven insights with limited options for personalization (such as dashboards or PDFs) that leave little room to public bodies’ for setting the terms of the analysis, thus often curtailing the possibility of using data for fruitful purposes. Furthermore, companies seem to be not fully transparent regarding key issues for data quality, such as representativeness, reliability and resolution. Local administrations have a glimpse of the limitations of such data assets only once they get access or start working on them. The quote, from a big city with extensive experience and high reputation in data innovation, shows how challenging it is to obtain good quality data.

“We’ve also found companies over-promise. So, a little while ago we were looking at card transaction data (that) gives you an idea of what they’re buying, what activities. (…) And it turned out when we looked into it, that the accuracy of the point of sale, it could have been anywhere in possibly 100 shops. So, we couldn’t use it to answer the kind of questions we were trying to answer, and the company had sworn blind in a number of meetings that the data was perfect (…)” (city09)

A long process of negotiation between local administrations and companies is often needed before those at the receiving end of the relation are satisfied with the data assets they obtain. Municipalities struggle to find an agreement regarding the correct specifications for data sharing to include in the procurements. Respondents who obtained data through public procurements stressed the experimental setup of such data sharing activities described as “evaluation phases” to assess the quality of data and the opportunities it affords for the municipality. Great decision-making power lie in the hands of the company that sells data, making it demanding for public bodies to define which information they need, in what format and for what kind of analysis.

“The company has set the rules and we’re not at a stage yet where we’ve set the rules in any of the projects I’ve been involved in.(…) But until you agree to their terms, you can’t get your hands on the data. So, we’ve had quite a few discussions that have gone round and round and round, maybe for a year, maybe for a year and a half. Which is why we’ve decided (…) we’re getting a month’s worth of data as a sample (…) to assess it against other data sets that we know to be true or that we
have confidence in (...) I think maybe after the evaluation, well be in a stronger position to say ‘well, you said this, but actually we’ve tested it and it’s that’.” (city09)

Another recurrent perspective towards procurement of data is “rejection”. Respondents considered not advisable to pay to obtain private sector data of public interest and drew both from practical and ideological arguments to support their statement. On the one hand, many municipalities struggle financially, due to budget cuts and an overall lack of economic resources, therefore paying for such information was not deemed appropriate. On the other hand, many felt that paying to obtain private sector data could create a hazardous antecedent that would make public bodies dependent on certain private companies. The “data sovereignty” (Hummel, et al., 2021) of municipalities could be jeopardized if these become buyers of ‘big tech’ corporations and telecommunication companies. The quotes, from representatives of two medium-size cities quite advanced in the area of urban innovation, further clarify the positions of local administrations unwilling to pay for private sector data of public interest:

“I’m very reluctant to pay for data (...) first of all; we need to keep a certain amount of independence from third parties when it comes to information on your city. Because data is not neutral and if we become very dependent on a tracker, we know there’s not a lot of competition on the market, because the technology is expensive, scaling up is expensive, the knowhow is a long process (...) I never heard that a product is becoming cheaper over the years.” (city04)

“One of the most important things in the equation is that we are not putting money in it, so if we bought the data that would be easy (...) we don’t have that kind of money and it’s also some kind of a principle discussion that the data has been collected in a public space. Data collected in public space is from everyone, it’s not just from the company who happens to put a sensor.” (city05)

**Data partnerships and pools: Cooperative engagement**

The respondents showed a very different attitude for data partnerships and pools, which were enthusiastically described as “win-win collaborations” based on shared interests between local administrations and data holder private companies. Instead of consisting solely in a transfer of data assets, both parties were involved in the project at times also sharing the final objective for which the data is used. In the interviews, respondents often switched to “we”, instead of “them”, to refer to the actors of the private sector that hold the data with whom they collaborated.

“We try to talk to them, like, ‘what are your incentives and how can we help you?’ We didn’t go to them and say, okay, we need your data. We say ‘how can we work together?’” (city02)

In most cases, local administrations give administrative data in exchange of private sector data, creating a pool with data from one or more private companies, eventually including government-owned companies (Shkabatur, 2019; Micheli, et al., 2020). Alternatively, municipalities act as a partner for the development of a data-driven service. Respondents positively perceived this approach to B2G data sharing. They explicitly referred to data partnerships and pools as the most favorable way not only for getting access to private sector data, but especially for being able to take advantage of the information such data provides. For instance, respondents described data partnerships and pools as “productive relations” with private companies that allow addressing societal challenges more effectively. Both respondents who had engaged in it, and those who would have liked to do so, describe this operational model as a form of “co-creation”,
which stands in stark opposition to “buying data” and being “just a client” of data holder companies. The quote below is from a representative of a European capital city that was not able yet to establish a data partnership for B2G data sharing. Similarly to colleagues who instead had engaged in such partnerships, the respondent highlighted the advantages of collaborating as peers with companies, and considered this operational model as the most productive and sustainable approach.

“I think we should establish collaborative relations since the very beginning, we should find circumstances in which from the beginning it is clear what is the advantage for the municipality and which for the companies holding data. On the one hand, this leads to a co-creation of the goals and practices, and at the same time it lowers the costs for the local administration, because you (the municipality) are offering an advantage to the data holder company as you are helping them to improve the quality of their data.” (city01)

Yet, this model is not accessible to all equally. For instance, respondents claimed that such data partnerships are more likely to develop among people that already know each other, i.e., that are in one’s professional network and with whom exists a professional relationship.

“We have a history with the people. I mean the people working in company X, I know her for, I’ve knowing her for now five years maybe. Had discussion on different topics, and now I know where she wants to go. She knows where we want to go. We know where we could go together — it is easier. With company Y, it is the same. We are working with them on data since 2010.” (city06)

Another enabling factor is the societal relevance of the projects. According to the respondents, the new generations of developers are interested in working on socially relevant issues. Therefore, establishing personal relationships with them (for instance during hackatons) facilitates future collaborations. Furthermore, although partnerships might originate from a common goal, private companies mainly join to develop a business model or a commercial product to offer to other cities. Therefore, municipalities with advanced know-how in data innovation are more likely to find private companies willing to collaborate with them. The more qualified a municipality’s data office is, the more it has to offer to private companies, as it is highlighted in the statement below from an interview with a representative of well-known “smart city”.

“The collaboration so far is more that it’s a win-win, that they give us what they have, and they see what we do with it, how we enhance it, which makes their product better. So, it’s that iteration.” (city02)

Data sharing obligations: Sovereignty

Municipalities may include clauses in their tender contracts demanding that data collected by a service provider is made accessible to the city council. Usually, businesses subject to data sharing obligations have a contractual relationship with the municipality, such as for utilities, government-owned companies or mobility operators (e.g., public transport companies, waste management companies, telecommunication actors and ride-sharing companies). However, they can also include services built on public infrastructures and technology companies (Bass, et al., 2018). A few study participants already adopted data sharing obligations in some of the services subcontracted by their municipality, others instead were starting to consider it for the future. This approach was perceived as a systematic way to get access to private sector data of public interest, both in organisational and legal terms, since the same rules apply for different contractors and across departments. The quotes below, from representatives of a big capital and a medium-
size city with a history in data innovation, show how European municipalities are starting to acknowledge the advantages of data sharing obligations, but are still learning how to implement them in the most efficient way.

“We have done that [paid for data] in a couple of situations where it was not specified well in the tender (...) What we try to do now is prevent that by making our contracts better and have a warrant in our contract that says all the data being used in something we buy, belongs to the city.” (city02)

“We are thinking about something more systematic, like how to introduce data requests in our contracts, in our agreements on different public policies. It’s quite a different perspective. It’s not how to reach new partners on data, but how to introduce data with our historical partners.” (city06)

Data-sharing obligations grant public bodies the power to set the terms for accessing private sector data (“sometimes we don’t make it better for the private companies, but better for us”). Through such contracts, local authorities not only make sure that information collected as a by-product of delivering a service is made available to them, but they can also specify in advance the format for exchanging information keeping control of the modalities for data sharing (“we ask for a format that’s readable in a machine-to-machine data exchange”). Respondents justified data sharing obligations with the claim that data collected as a by-product of delivering public services should also be available to serve the public interest. The aspiration of keeping data sovereignty — intended as public bodies maintaining control of public interest data without becoming dependent from corporations or external bodies — is also part of the discourses adopted in favor of data sharing obligations and against the monetary purchase of data insights generated by corporations.

“Thinking about contract services, tenders, saying that you are providing services as you were the city council, so it’s not your business to collect data about the city. Okay, your business is to provide a service that you are contracted for, so the data you are collecting within the service needs to be available for everyone to provide, or for the city to provide, or to improve the service.” (city12)

**Strategies to enhance local administrations’ negotiating power**

The strategies discussed by the participants to enhance their negotiating power for accessing private sector data are transversal to the four models described earlier. Although the vast majority of experiences consisted in bi-lateral relations between a single municipality and a data holder company, the interviewees shared the belief that in order to enhance public bodies’ access to private sector data of public interest more actors had to be involved in the process.

A foreseen strategy consists in “joining forces” with other cities to set the terms for B2G data sharing as a means for municipalities to build collective bargaining power; for instance, acting within a collective of cities to better negotiate prices of data assets with private companies that have a dominant market position and to avoid lengthy negotiations. As the quotes below emphasize, such approach is instrumental to share costs, strike better deals, and/or avoid iniquities between cities.

“We think that cities have a real role in, basically, collective bargaining on this and telling companies what they pay for it, rather than the other way around.” (city09)

“With a collective, in a sort of a collective effort with other
Public bodies’ access to private sector data: The perspectives of twelve European local administrations

Furthermore, respondents took part in national or transnational networks to address some of the challenges of B2G data sharing. For instance, a respondent from a medium-size city was collaborating with an association of municipalities to develop a joint standard contractual framework for data sharing relations with private companies. Through a mutually agreed upon and standard contract, the municipalities hoped to increase their negotiating power vis-à-vis companies and suppliers and better clarify their demands:

“We are working together with the association of cities (...) and we want to come up with a model contract in which we can come up with the juridical text where we can use that to make a contract with these businesses upfront. So, there is no discussion about a data, but it would be every city in the country is using the same contract, so it’s no use to go shopping to another city because it’s very similar. I think it’s a very, very good initiative.” (city10)

Another strategy was involvement of third parties as data intermediaries to facilitate effective and efficient exchanges of data between multiple actors. The expertise brought in by intermediaries, mentioned by the respondents, was rather diverse, ranging from legal support, helping with negotiations with companies and overall decision-making processes over data, providing technical solutions related to security, data infrastructures or data analytics, and informing the development of ethical frameworks. Intermediaries could also be those accessing and collecting private sector data in the first place. Respondents mentioned a wide range of actors that operated as intermediaries: start-ups specialized in data-driven innovation, consultants, universities, and platforms for distributed peer-to-peer data sharing (e.g., Ocean Network [5]). Additionally, municipalities also included new professional figures in their own offices dedicated on some of the tasks above (such as data stewards). Overall, respondents sought to increase their negotiating power in B2G data sharing by seeking support in cities’ networks, data intermediaries or new recruited staff.

Discussion

Access to private sector data is an embryonic practice that takes place through different modalities, each distinguished by specific relations and levels of engagement between actors. Yet, little is still known about how local administrations implement and understand such practices. Through expert interviews I examined how B2G data sharing is understood and implemented in twelve European cities and the related perspectives of municipalities’ chief data officers (or similar representatives). Although B2G data sharing is an infrequent activity, four operational models were mentioned more often in the interviews: data donorship, public procurement of data, data partnerships and pools and data sharing obligations (Table 2). The article provides information about how the four models are perceived and what strategies are envisioned to enhance access to private sector data and what power relations emerge between actors for data access and control. The findings point to two broad critical issues: the inequalities between cities in the opportunities to access data, and the aspiration of public bodies to control the modalities through which private sector data of public interest is shared.

Although challenging for all, certain cities have more chances to access data collected by the commercial sector than others. Reputation as “smart cities” put certain municipalities in a favorable position to be contacted by companies for data donorship. Similarly, municipalities with more resources and capacity, such as a wide professional network among data scientists and data officers in business entities, greater data
Public bodies’ access to private sector data: The perspectives of twelve European local administrations

culture and expertise, are more able to find partners who are willing to collaborate and eventually share data at no cost in data sharing pools. According to the respondents, private companies benefits from pro bono collaborations by developing use cases to be marketed to other cities. The finding suggests that, if access to private sector data assets is not supported by policy measures, it is likely that municipalities with more resources will be those able to benefit from the new stream of information provided by companies — eventually enhancing their advantaged position in the field of data and urban innovation — according to a “richer get richer” dynamic, also referred to as Matthew effect (Merton, 1968). Such differential access to data could potentially exacerbate power imbalances in the digital landscape, providing the basis for a new instance of the “big data divide” (Andrejvic, 2014) between those who are able to collect and analyse large quantities of data and those who are excluded — typically data subjects, but also other actors, such as less innovative or well-known municipalities. Further inequalities derive from the different skills and resources available to public bodies, which lead to different levels of awareness and uses of the information such data provides. The multidimensional concept of digital inequality can help to better unpack the implications of power unbalances in both data access and use (McCarthy, 2016).

A second key issue, emerging from the discourses of the participants, concerns the importance of controlling how private sector data is shared. To simply gain access to private sector data, especially if through data donorship or public procurement, was not perceived as an achievement by respondents. If public bodies lack the means to set the terms of how data is shared, it is not assured the information they get will be useful to their public interest purposes. For instance, a company might only share reports, with figures and numbers that do not allow public bodies to integrate private sector data with administrative data for further analysis. Otherwise, data quality might be lacking, not corresponding to what companies have promised to deliver, hence requiring efforts on the side of public bodies for its assessment (which leads to further negotiations for data sharing). The modalities of access that support a greater level of engagement by public bodies are those more positively perceived by respondents. Respondents wish they could be actively involved in B2G data sharing processes, not as ‘recipients’ of donations or ‘clients’ of data products, but as key actors, such as ‘partners’ or ‘administrators’ (Table 2). To keep control of how data is shared, municipalities either establish collaborative relations with private companies through data sharing pools (as ‘partners’), or include data-sharing clauses in tender contracts with suppliers (as ‘administrators’). Such models of access, however, enable different levels of engagement of the actors. The first demands a well-established professional network upon which the collaborations are established (and that not all cities have), while the latter is mainly applicable to businesses that have a formal contractual relationship with the municipalities.

The enclosure of data sharing obligations in contracts with suppliers is understood as a key mechanism to systematize municipalities’ access to data of public interest. Respondents justified this mode of access with the underlying principle of data sovereignty: local administrations should have the right to access data about citizens, collected as a by-product of the delivery of public services (paid by the public pursue), and to use it for the benefit of the whole community. Data sharing obligations could empower the public sector allowing access to additional data and internal capacity. Data of public interest is understood as a public good; hence, the information it affords should not be at the mercy of private companies that use it for profit-oriented goals. This mode of access paves the way for the re-use of private sector data for the public interest. It has implications for the wider policy debate on data governance as it represents public bodies as powerful actors who set the agenda for data (re-)use, instead of acting as ‘clients’, ‘recipients’ or even ‘partners’ of data holder companies.

Respondents shared the belief that, in order to be more systematic, B2G data sharing should move from bi-lateral partnerships to more stable contractual frameworks that involve a broader set of actors. Great emphasis was placed in particular on the opportunities that collective forms of action would bring to municipalities. Local administrations’ representatives thought that cities could “join forces”, either to define common contracts for tender clauses and partnerships (to avoid companies “going shopping” from one city to the other) or to bargain and strike better deals for the procurement of data. Such associations or networks of municipalities could tackle cities’ unequal opportunities to access private sector data, while increasing control of data quality and the sharing of data. They could help level the playing field increasing cities’
Public bodies' access to private sector data: The perspectives of twelve European local administrations

negotiating power. Finally, involving third parties in B2G data sharing practices and hosting new human resources in the municipalities strengthens the negotiating power of public bodies for B2G data sharing. Such multi-actor approaches seem critical to rebalance the power asymmetries of current local data ecosystems.

Conclusions

Upcoming policy initiatives on data governance, which in Europe are under preparation, will shape the use of private sector data for the public interest (European Commission, 2020a). Besides key economic, organizational and legal challenges detailed in recent analyses of case studies (e.g., Huyer and Cecconi, 2020; Susha, et al., 2019; European Commission, 2020b; Martens and Duch-Brown, 2020), such policy interventions could also be informed by the experiences of public bodies and related socio-economic power dynamics. Making visible the perspectives of those working in the field, the challenges they face and the strategies they put forward, is a mean to promote the participation of a wider range of actors in shaping the digital transformation. This study provides empirical evidence to inform upcoming policy interventions by increasing knowledge on the practices and meaning-making processes of a key social actor: local administrations. In particular, the article explored the power relations embedded in B2G data sharing bringing to the fore three main themes: the contextual factors that shape access to private sector data and might generate inequalities between municipalities; the willingness of local administrations to actively engage and control how data is shared with them; and the potential of collective efforts to increase local administrations’ negotiating-power and address some of the asymmetries of current data ecosystems.

The study has limitations: it is small-scale empirical research conducted on a purposive sample of cities not representative of all European municipalities. The research collected the perspectives of one relevant social group only: chief data officers or similar representatives. Furthermore, it deliberately focuses on the models of access and does not include other key themes, such as the use of private sector data; the resources available to local administrations and how they impact experiences and perceptions within the municipalities; the challenges and risks associated to data use; and the role of citizens and public trust.

Although the article only addresses the relations between private and public sector, citizens are those who generated data in the first place. There is growing awareness that citizens have to be involved in local administrations’ (data) innovation initiatives. Such awareness partially showed also during the interviews with respondents voicing a desire of giving data “back” to citizens, allowing them to control it and benefit from it. However, in most instances, citizens are viewed not as participants or co-creators, but as data subjects whose data could be harvested for the public interest (Calzada, 2018; Cardullo and Kitchin, 2018; van Zoonen, 2020). On the other hand, pioneer initiatives across Europe suggest that public bodies could establish a fiduciary relation with citizens. Public authorities could be trustees of citizens personal data providing them the means to exercise data sovereignty and decide what to share and with whom (Bass, et al., 2018; Morozov and Bria, 2018; Open Data Institute, 2019).

Taking such limitations into account, the article concludes suggesting that further research on public bodies’ data sharing practices could examine to what extent “more access” to private sector data effects public services delivery, helps to address societal challenges and enables more inclusive forms of data governance in which both local authorities and citizens are powerful stakeholders.

About the author

Marina Micheli is Scientific Project Officer at the European Commission’s Joint Research Centre (JRC) at the Digital Economy Unit. She is interested in the social value of data, digital inequalities and data power.
and her background is at the intersection of media studies and sociology. Her current work explores data governance for the public interest, examining in particular the perspectives of social actors and the emerging models for the governance of personal data. Prior to joining the Commission, she was a Senior Researcher and Teaching Associate at the University of Zurich (Switzerland) and a Post-Doct oral Fellow at the Department of Sociology and Social Research at the University Milano-Bicocca (Italy).
E-mail: marina [dot] micheli [at] ec [dot] europa [dot] eu

Acknowledgements

The views expressed in this article are solely the responsibility of the author and do not reflect the official opinion of the European Commission.

Notes

1. Meijer, 2018, p. 203; see also Löfgren and Webster, 2020.

2. An earlier version of the article was presented at the Data for Policy Conference 2020 (Micheli, 2020) and a preliminary analysis has been included in the final report of the Digitranscope project, in the framework of which the study was launched (Craglia, et al., 2021).

3. Taylor and Broeders, 2015, p. 236.

4. To protect participants’ identities, the number of cities in quotes does not correspond to the order in Table 1.


References


Public bodies’ access to private sector data: The perspectives of twelve European local administrations


---

**Editorial history**

Received 31 May 2021; revised 30 November 2021; accepted 3 January 2022.
Public bodies’ access to private sector data: The perspectives of twelve European local administrations
by Marina Micheli.
First Monday, volume 27, number 2 (February 2022).
doi: https://dx.doi.org/10.5210/fm.v27i2.11720