Accessing privately held data:
Public/private sector relations in twelve European cities

Marina Micheli

European Commission – Joint Research Centre, Ispra (VA), Italy
marina.micheli@ec.europa.eu

Abstract

Access to data collected by the private sector could help local administrations to comply with their public-interest mission. Privately held data access by public bodies, however, is an emerging field with no established praxis yet. Its in-the-making status makes it a compelling subject for a research on the "politics of data", since it allows looking at the power (un)balances between the companies holding data and the local administrations interested in it as well as citizens' role. The contribution discusses the findings of a qualitative exploratory research that investigated the experiences and perspectives of local administrations regarding access to privately held personal data. Semi-structured interviews have been conducted with the innovation/data managers of twelve European local administrations. The interviews allowed contextualizing concrete experiences of data sharing in the discourses and viewpoints of the specialists working in this area, as well as the strategies they plan to put forward for facilitating access. The study illustrates how these cities' managers experience four operational modes to access private data: data donorship, public procurement of data, data sharing partnerships and tender obligations. The results highlight new divides caused by the unequal possibilities to access private data, as cities with more experience and better reputation are in a favorable position both to establish partnerships and receive data at no cost. The findings also suggest that public managers would like to retain sovereignty of data, co-create instead of acquire data-driven solutions, and that are imagining collective forms of negotiation to enhance their access to private data.

Keywords – cities; privately held data; qualitative research; data governance; data access.

1 Introduction

Privately held data might offer insights and valuable opportunities to public authorities supporting their efforts to address both societal challenges and emergencies. At the city level, attempts to access such data are made through various operational models, which often involve also other actors (e.g. research institutions, civil society organizations, start-ups). This study looks at how local governments in Europe are (or are not) gaining access to privately held data examining the perspectives of city's managers and project leaders (e.g. chief data officers) that are working in the field of data, technology and urban innovation. Through twelve interviews, we collected experiences and opinions about the various operational modes for accessing privately held data. A goal of the study is to shed light on the role of local governments in contemporary data politics, intended as the power relations embodied in data control and use (Ruppert, Isin & Bigo, 2017). Public bodies, especially including local governments, are in fact key actors, as they could have an important function in promoting a more balanced and just data economy based on the redistribution of value produced with personal privately held data (Couldry & Powell, 2014; Morozov & Bria, 2018; Bass et al., 2018; Adalovelace Institute, 2019).

2 Access to private sector data by public bodies: State of play

Privately held personal data – collected by mobile phone operators, social media platforms, transport services, accommodation websites, energy providers, and so on - is often described as part of the "data city" or of the "urban data landscape" (Powell, 2014; Kitchin, 2018). Yet, notwithstanding the societal expectations, the practice of data sharing between businesses and governments is currently sporadic and lack sustainability (HLEG, 2020; Martens & Dutch-Brown, 2020). The High Level Expert Group on Business-to-Government (B2G) data sharing - established by the European Commission to provide recommendations on how to enhance the sharing of privately held data for the common good - identified several challenges and obstacles. Among the critical issues stands out the lack of governance frameworks, which means that private companies have to face various uncertainties when sharing their data - in relations to liability regimes (who is responsible if inaccurate or biased

data is shared that leads to discrimination), intellectual property, and competition law. Furthermore, they face operational and technical challenges for the preservation of sensitive commercial information and the protection of customers' personal information (HLEG, 2020). While in a study on the economics of B2G data sharing, Martens and Dutch-Brown (2020) identified the following economic barriers: monopolistic data markets (companies can charge high prices for data), high transaction costs and perceived ex-post risks for data providers, and lack of incentives for private firms to contribute to the public good if it might affect them negatively (e.g. competition, market regulation). Finally, challenges specifically faced by public bodies include a lack of "culture" on data sharing (e.g. how to create value with it), few resources and deficiency of skills among public servants, as well as limited trust from both the private sector and citizens on public bodies accountable use of data (HLEG, 2020). The regulations currently available for privately held data sharing vary by EU country and are sector specific, while at the EU level regulations are currently being discussed (EC, 2020).

Notwithstanding the circumstances, some cities are implementing data sharing practices, experimental projects and other initiatives developing their own strategies in an uncertain terrain. The operational models adopted by local administrations to access privately held data are diverse and denote very different relations between these actors. Private companies, for instance, might share data with public bodies at no cost on a voluntary basis as corporate social responsibility, such as during an emergency or to support initiatives for the public interest. This mode for accessing private data is referred to as *data donorship* (Huyer & Cecconi, 2019; HLEG, 2020). Otherwise, public administrations might purchase data through *public procurement* (HLEG, 2020). In this case, triggered by specific needs, public bodies request to acquire a specified set of data, or data-driven insights, from a data supplier through a contract (Huyer & Cecconi, 2019: 16). A different approach is that of *data sharing pools* (Shkabatur, 2019; Micheli et al, 2020) in which a public authority establishes a *partnership* with other actors to pursue mutual interests, and commercial companies, government entities, data platforms, and/or research institutions exchange data in a collaborative way. They "analyse each other's data, and help fill knowledge gaps while minimizing duplicative efforts" (Shkabatur, 2019: 30). A related mode is that of *data research partnerships*, that is when public bodies collaborate with research/scientific institutions for a project of mutual interest to analyze privately held data that the latter have at disposal (HLEG, 2020). Another relevant mode for accessing data of private sector consists in the introduction of *data-sharing obligations* as part of subcontracted services (HLEG, 2020): cities might include clauses in their tenders "specifying that a service provider must make any data that may be of public value available to the city council in machine-readable format" (Bass et al., 2018: 28). All modes could lead to access different types of data: raw, pre-processed (e.g. cleaned, re-sampled, normalised), processed (aggregated and combined) or insights derived from the data (HLEG, 2020). Private companies might be more willing to sell (or donate) insights deriving from internal data analysis ("intelligence sharing", such as dashboards, apps, reports), instead of actual datasets, as a way to keep control of information and reduce risks (Shkabatur, 2019; HLEG, 2020; Micheli et al., 2020). Furthermore, data could be shared in various technical solutions, such as Application Programming Interfaces (APIs), limited release of data under conditions stipulated in a contractual agreement, remote access by a trusted intermediary, etc. (HLEG, 2020). The study presented in this contribution contextualizes the different operational modes for accessing privately held data described in this section in the experiences and opinions of European city's managers, directors and project leaders.

3 Methodology

This study examines how the socio-technical practice of accessing privately held data is "constructed" throughout the relationships established between actors. Data sharing is not examined as a technical issue, neither as an economic activity, but as a socio-technical practice that is shaped by the specific context in which it takes places and with certain meanings according to the social actors who engage in it. The methodological approach is informed by the tradition of research in media domestication and the social shaping of technology (Silverstone & Haddon, 1996; Lievrouw, 2006). This method is here adopted to understand how public actors envision their 'power to set the terms' on how privately held data is shared (allowing them to use it to pursue their public interest mission) and what strategies they put forward to facilitate access.

The study focuses in particular on the perspectives of specific actors from the public sector: cities' Chief Technology Officers, Chief Data Officers, or project leaders working on a city's innovation/smart city agenda. Twelve semi-structured interviews have been conducted with representative of as much cities during the course of 2019. A combination of purposive and snowball sampling procedures has been adopted for the selection of cities to be included. The participants were chosen in a way to have a diversified group by city size, area in Europe, and tradition of innovation. Together with the "usual smart cities" (London, Barcelona, Amsterdam), the list includes cities of different European areas and size (Table 1).

Table 1 – List of cities involved in the study[1]

City	Country	Size	EU zone
Amsterdam	NL	Large	North West
Barcelona	ES	Large	South
Ghent	BE	Small	West
Ljubljana	SI	Small	Central East
London	UK	Large	North West
Milan	IT	Large	South
Rennes	FR	Small	West
Rjeka	CR	Small	East
Tallin	EE	Mid-size	North East
The Hague	NL	Mid-size	North West
Vilnius	LT	Mid-size	North East
Zaragoza	ES	Mid-size	South

Large: > 1 million inhabitants; Mid-size: Between 300.000 and 1 million; Small: < 300.000.

The interviews, which lasted 70 minutes on average, have been conducted through VoIP (Voice over Internet Protocol) technologies that allow interviewing research participants using voice and video across the Internet via a synchronous connection (Iacono et al., 2016). Conducted by the author, the interviews investigated the "concrete realities" of working in this area, digging into the actual experiences of these professionals, and simultaneously analyzing discourses and imaginaries, investigating their opinions on the topic. The interview outline covered the following topics: data strategy of the municipality, overview of projects and initiatives run by the municipality based on access to privately held data (including 'failed attempts'), purposes for data use, private actors involved, relationships and negotiations with private companies, enablers and obstacles (technical, organizational, etc.), operational modes for data access, data analysis, regulations, human resources and skillset, citizens role and attitudes. The transcriptions did undergo a qualitative thematic analysis through manual coding; the documents have been coded following the main themes of the interviews. The findings discussed in this contribution focus in particular on the analysis of the following themes: "Operational modes of access", "Discourses and perspectives on the topic", "Relationships between actors", "Power to set the terms", "Strategies to negotiate power". The analysis aims to both delineate common trends in the experiences and discourses, as well as key differences and how these relate to the specific context. The interviews strived to move beyond the *façade* of promoting the innovative activities of the city, although it is not possible to assure that this was achieved equally in all cases. During the interviews it was taken into account that digital innovation is a highly marketed issue for cities, for

instance by explicitly asking about obstacles and unrealized projects, and generally being self-reflexive during the conversations.

4   Results

Local administrations' access to privately held data is a rather sporadic activity. All respondents, including those from municipalities with a reputation of smart cities, stated that accessing privately held data is challenging. Data companies often have no interest in selling data, and neither in sharing it with a municipality (cfr. Section 2). The practices mentioned in the interviews are often pilot projects, activities at the "early stages", if not still in preparation ("figuring out"). In a few cities the topic was rather novel, as access to private data was not part of current/planned activities ("nobody's talking about that and nobody's thinking about it").

In general, companies with mobility data were cited more often, both as potential, past or actual data providers, highlighting how access to privately held data might be promising especially for this sector.

The types of data providers most interesting according to those who participated in the study are: (1) utilities companies, and (2) online platforms and telecom operators. Utility companies are in general very interesting for municipalities. These are depicted as the ideal candidate for access to privately held data, but also as a difficult one to deal with, due to lack of human resources and interest in data sharing. Online platforms and telecom operators feature less prominently. According to a participant, big platforms are difficult to reach, also because they do not have representatives working at the local or even national level. Furthermore, they seem to have less to gain from engaging with a municipality. Several other actors are involved for access (and use too) of privately held data. Start-ups, public universities, research institutions and civic organizations also have an important role in this context, helping with data stewardship and analytics. Collaborations with research institutions or PhD students, is mentioned as an enabler for working in this field, especially by cities with less economic resources.

After this brief introduction, in the next, we contextualize the operational modes for accessing private data in the discourses and experiences of the professionals of twelve cities. Hopefully these findings could help to shed light on the hindrances, as well as promising avenues for public bodies' access to privately held data.

3.1   Data donorship: "Free lunch?"

Some respondents recognized the possibility to access privately held data or information at no cost, as companies/data providers occasionally make it available

---

[1] To protect participants' identities, the numbered cities in the quotes does not correspond to the order of cities in this table.

for free on a voluntary basis. This *data donorship* mode, however, was often associated to a specific discourse. Instead of being described as a philanthropic move, it was explicitly acknowledged as a marketing strategy private companies used, which inevitably favored already privileged "smart cities". Companies that freely share data with cities – and collaborate with them to develop products or services valuable for the municipality without asking anything in return - do so because this allows them to market new products and services to other cities in the future.

> "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)

Companies use the reputation of (smart) cities as promotional material. Thus, being a well-known city seems to be a key enabler for such form of access to private data. This creates a double source of disadvantage for smaller cities because they lag behind and are proposed the same service to a price. This phenomenon emerged as an 'ethical dilemma' in a couple of interviews in which managers questioned their city's position (its privilege or lack of thereof) in relation to that of others:

> "They (companies, ndr.) 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?" (city02)

> [During the 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)

This operational mode for access tend to be associated to smaller or one-time-only projects; one interviewer described that as an "incidental partnership" to stress the volatile nature of the initiative.

3.2 Public procurement of data: Resistance and caution

While almost all interviewees discussed the possibility to acquire data directly through *public procurement*, most were against this solution and the remainders experienced it with great circumspect. This operational access mode is contrasted with ideological arguments: (1) data produced in public spaces should be accessed by public bodies and not be treated as a commodity; (2) local administrations have to serve the public interest and should not invest economically is acquiring data; and (3) it is important that cities keep sovereignty over data, becoming a buyer to a private platform (especially if a big corporation) might undermine their autonomy.

> "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 (…) sometimes it is their business model, so they don't want to give the data for free, they want to have money. And well, we don't have that kind of money and it's also some kind of a principle discussion that the data has been collected public space. Data collected in public space is from everyone, it's not just from the company who happens to put a sensor" (city05)

Those engaging in public procurement of data also questioned its effectiveness. Companies, in fact, often sell data packaged with limited options for personalization, they send finished products (such as dashboards or PDFs) that curtail the possibility of intervention on data. Furthermore, they are not transparent regarding data quality and representativeness ("we've also found companies over-promise").

> "In the best scenario, we receive data in a PDF, but not in an XML format, so it's very difficult for us to process the data in our systems and well, basically it's not of any use at all if we receive a PDF." (city10).

Respondents who obtained data through public procurements most likely have done so through pilot projects. They stress the experimental setup of such activities, described as "evaluation phases" to assess the quality of the data and the opportunities it affords. This mode for access is criticized mostly because it does *not* allow municipalities to be involved in defining the information needed and the kind of analysis performed,

leaving great decision-making power in the hands of the company. Only after a long process of negotiation, local administrations seem to be eventually satisfied with data they obtain.

> "They are ready to sell them, but they're not ready to think with us how can they contribute." (city06)

> "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." (city09).

A couple of respondents suggested that cities should do collective bargaining to deal with the issues of the long negotiations and the high prices enforced by private companies. This strategy, which consists in cities creating a coalition and relating with companies altogether, would allow municipalities to strike better deals.

> "We think that cities have a real role in basically collective bargaining on this and telling companies what they'll pay for it, rather than the other way round." (city08)

> "With a collective, in a sort of a collective effort with other stakeholders, to share also the cost (…) we could also imagine that the data, which is for sale, we can go and buy it under a collective." (city06)

3.3 Data partnerships: Win-win collaborations

A completely different attitude emerged when respondents claimed that they established (or wish to establish) partnerships with private companies to cooperate, experiment and solve challenges together. In this operational mode local administrations identify shared interests with the private companies holding data, seeking a win-win collaboration. Both parties are involved in the project and in the analysis, at time also sharing the objective for which the data is used. Local administrations could give administrative data in exchange, creating a data sharing pool with the private company (Shkabatur, 2019; Micheli et al., 2020), or simply be a partner in the development of a product (e.g. a public service).

> "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)

This approach allows establishing productive relations with private companies and to address societal challenges more effectively. Significantly, this relation is described as a form of "co-creation" in stark opposition to "buying data" and being "just a client" of data holder companies. Several conversations and negotiations are needed to establish such relationships. A key element seems to be to work with people already in ones network with whom a personal relationship has already been established.

> "We have a history with the people. I mean the people working in companyX, 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 companyY, it is the same. We are working with them on data since 2010."

Another important element, in favor of public bodies, is the societal relevance of the projects on which these collaborations are based. According to some respondents, the new generations of developers are interested in working on socially relevant issues. Therefore, establishing personal relationships with them (for instance during hackatons and events) is an enabler for future collaborations. If/when these young developers will work for private companies, they will know that cities have interesting data and will be inclined to collaborate. Occasionally respondents highlighted that the partnership originated from the common goal (between the municipality and a private company) to impede the dominance of big tech corporations in a certain sector, such as the use of Google Maps as mobility app. Yet, it is also true that private companies join partnerships to develop a business model or a commercial product to offer to other cities/clients. Therefore, these partnerships might also lead to inequalities between cities (such as for data donorship), as municipalities with advanced knowledge and expertise in data innovation are more likely to find private companies willing to collaborate with them. The more experienced a city is, the more it has to offer to private companies in terms of data quality and support in developing and improving their products and services.

> "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)

3.4 Tender obligations: Data sovereignty

Another approach to access private data consists in putting clauses within tender contracts with suppliers so that data collected as by-product has to be accessible to the city council. Only a minority of respondents has put it in place, with a few others seriously considering it. This operational mode can be used for accessing data of suppliers and, theoretically, of any company that has a contractualised relationship with the municipality (public transport, waste collection, avenues rented for events, telecommunication infrastructures, etc.).

"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 of X." (city02)

"We are thinking about something more systematic, like how to introduce data questions 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)

The reasoning behind this is analogous to the explanations given for not using public procurement for buying data. Data collected for public services, should be available to public bodies as data sovereignty allow directing the digital transformation at the service of public interest.

"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)

A strategy put forward by a few respondents to enhance such mode of access consists in working collectively with other cities and jointly define the same contractual framework to be used with private companies:

"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."(city10)

4   Discussion

This discussion paper summarizes findings from a qualitative study with innovation/data managers of twelve European cities. Access to privately held data is examined as a socio-technical practice that is still "in-the-making", as no established praxis is fixed yet. The research contextualized four operational modes for accessing private data in the experiences and discourses of cities innovation/data managers. From the results we learned that, although it is challenging for all, some cities seem to have more chances to access privately held data. Their reputation, their professional network, as well as their resources and expertise, put them in a favorable position to be contacted by private companies (for *data donorship*) or be saluted as partners (for *data sharing pools*). Private

companies, in fact, use such "use cases" to market their services and products to other cities. Further research could see to what extent a divide between cities regarding their chances to access data and use exists and with what implications. Another underlying issue, emerging from the discourses of participants, is that of data sovereignty: respondents are wary in buying data through *public procurement*, both because that would place them in a dependent position (economically), and because there is a lack of transparency regarding privately held data quality as well as limited possibilities for controlling how it is formatted and used. To preserve control, some respondents envision a form of collective bargaining to strike better deals when acquiring data, but most often they look towards other modes of access. The best ways to keep control of data include establishing actual collaborative relations (co-creation) through *data partnerships* with private companies and, when possible, including *obligations clauses in tender contracts* with suppliers. Some of the strategies proposed to maintain control are collective efforts in which cities join forces for this cause: from collective bargaining, to develop a common contractual framework to use with businesses for tenders or partnerships. These tactics could indeed help to level the playing field, lessening the inequalities described above, and increasing cities' strength in demanding access to privately held data with a public interest.

This study aims to provide qualitative insights in the lived experiences of cities' innovation and data managers in relation to access to private data and to reflect on how to strengthen the role of public bodies in the current European data ecosystem. Clearly, a lot more could have been addressed in this discussion paper. The key omitted topics are: an in-depth recognition of the strategies adopted by cities to improve data access (which also include new roles, such as data stewards and intermediaries); a systematic analysis of the resources available to local administrations; a review of how cities are using or planning to use privately held data and what kind of discourses are associated to it; and how citizens, and public trust, are taken into account, for instance through the creation of 'city data trusts' that combine data from different sources and let citizens decide which to share and with whom.

Acknowledgements

References

Adalovelace Institute (2019). Rethinking data. Changing the data governance ecosystem. Report (pp. 1-26): Adalovelace Institute.

Bass, T., Sutherland, E. & Symons, T. (2018). Reclaiming the Smart City. Personal data, trust and the new commons. Report (pp. 1-76): NESTA.

Couldry, N., & Powell, A. (2014). Big Data from the bottom up. Big Data & Society, 1(2), 1-5.

European Commission (EC) (2020) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions "a European Strategy for Data" COM(2020) 66 Final. Luxembourg, Publications Office.

Huyer, E., Cecconi, G., & Capgemini Invent (2019). Analytical Report 12. Business-to- Government Data Sharing. Report (pp. 1-24) European Data Portal.

HLEG (2020). Towards a European strategy on business-to-government data sharing for the public interest. Final report prepared by the High-Level Expert Group on Business-to-Government Data Sharing (pp.1-112) European Union.

Kitchin, R. (2018). Data-driven urbanism. In R. Kitchin & T. P. Lauriault, Data and the city (pag. 44-56). New York: Routledge.

Iacono, V., Lo, P. S., & Brown DHK. (2016) Skype as a tool for qualitative research interviews. Sociological Research Online 21(2), 12.

Lievrouw, L. A. (2006). New Media Design and Development: Diffusion of Innovations v Social Shaping of Technology. In L. A. Lievrouw & S. Livingstone (Eds.), Handbook of New Media (pp. 246– 265). Thousand Oaks: Sage.

Martens, B., & Duch-Brown, N. (2020). The Economics of Business-to-Government Data Sharing. JRC Technical Report, JRC Digital Economy Working Paper 2020-04 (pp. 1-31), European Commision.

Micheli, M., Ponti, M., Craglia, M., and Suman A.B. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*, 1-15. DOI: 10.1177/2053951720948087

Morozov, E., & Bria, F. (2018). Rethinking the Smart City: Democratizing Urban Technology. New York, NY: Rosa Luxemburg Stiftung.

Ruppert, E., Isin, E., & Bigo, D. (2017). "Data Politics". *Big Data & Society*, 1–7.

Powell, A. (2014). Datafication', transparency, and good governance of the data city. In K. O'Hara, M.-H. C. Nguyen, & P. Haynes, Digital Enlightenment Yearbook 2014. Social Networks and Social Machines, Surveillance and Empowerment (pag. 215–224). Amsterdam: IOS Press.

Shkabatur, J. (2019). The global commons of data. Stanford Technology Law Review, 22, 1–46.

Silverstone, R. & Haddon, L. (1996). "Design the Domestication of Information and Communication Technologies: Technical Change and Everyday Life". In R. Mansell and R. Silverstone (Eds.), Communication by Design: The Politics of Information and Communication Technologies (pp. 44 -74). Oxford: Oxford University Press.