REUSERS OF OPEN PUBLIC DATA

The case of Montreal

Jérémy Diaz and Sandra Breux

INRS
Institut national de la recherche scientifique

Institut national de la recherche scientifique

REUSERS OF OPEN PUBLIC DATA

The case of Montreal

Jérémy Diaz and Sandra Breux

Research report submitted to the

Laboratoire d'innovation urbaine de Montréal

Institut national de la recherche scientifique
Centre Urbanisation Culture Société

January 2021

Scientific responsibility: Sandra Breux

sandra.breux@ucs.inrs.ca

Institut national de la recherche scientifique

Centre—Urbanisation Culture Société

Dissemination:

Institut national de la recherche scientifique

Centre—Urbanisation Culture Société

385, rue Sherbrooke Est

Montréal (Québec) H2X 1E3

Telephone: (514) 499-4000

Fax: (514) 499-4065

https://inrs.ca/

ISBN 978-2-89575-417-6

Legal deposit: Bibliothèque et Archives nationales du Québec, 2021

TABLE OF CONTENTS

List of tables

Executive summary

- The objective of this report is to provide a portrait of the reusers of the City of Montréal's open public data portal. In this report, open public data refers to data from a public institution made available to the public through a web portal.

- The writings identified five categories of reusers of this data: non-professional users; businesses, entrepreneurs and developers; data journalists, researchers, professors and students; community groups; and NGOs.

- Although responsive to the demand, the City of Montréal's open data portal—hereinafter ODP—is primarily a supply-led rather than a demand-led approach. An analysis of the requests submitted to the portal between 2016 and 2019 shows that a pool of 200 people gravitate around the portal each year and that the categories of reusers are the same as those identified in the literature. The intended use of the data is rarely indicated in these requests. The profile of the reusers and the impact or value created by the data is not known.

- The analysis of 19 semi-structured interviews with volunteer reusers of the open data allowed to remedy this lack of information, at least in part. Among the interviewees, two profiles emerged: explorers—those who use the data for personal purposes to develop their skills and see what can be done with the data—and exploiters—those who interact with the data to create a potentially marketable product.

- The reusers interviewed have all the necessary skills to read, understand, analyze and process data or are surrounded by people who can help them with these tasks. The impacts and use value created are not measured by the reusers; they are, at best, estimated. Reusers complained about a) the quality of the data and the lengthy clean-up and redesign work that follows; b) the lack of recognition and follow-up associated with the activities related to the products that result from their use of the data; and c) the variable interest of published data. At the same time, the reusers interviewed also highlighted the benefits of the ODP. For example, they see the publication of open data as an added value for their personal and professional practice (e.g., in that it provides an additional source of public information to enrich existing databases, offers the possibility of data serialization). These data also represent a gain in time (direct access to data) and a monetary gain (no need for an access to information request, rapid improvement of a service, etc.).

- Analysis of the messages received by the ODP team and interviews with volunteer reusers have revealed several courses of action: 1) *feasible* courses of action that could be realized by the team operating the portal, 2) *desirable* courses of action that would generate more human and financial resources, and 3) *ideal* courses of action that would require a redefinition of the ODP and the actors involved.

- The *feasible* courses of action aim to improve the reuser's context of use, increase the number of potential reusers and, in the longer term, develop a data quality charter. More specifically, this would involve having educational sections on the ODP, explaining the definition, purpose and construction of open data and its potential, while providing information on the types of files and formats provided and their objectives. Data traceability could also be offered, namely by publishing the different versions of a dataset and their possible errors. In the longer term, a data quality charter would be adopted that would be appropriate to the reality of the 19 boroughs of the municipality.

- The *desirable* courses of action are intended to go beyond the vision of the data reuser-consumer and to make the ODP a space for collaboration. This would involve making the ODP a springboard for a) creating interdependence and follow-up between reusers; b) creating added value that stems directly from these relationships; c) identifying other data suppliers that could be useful to reusers; d) offering tools and technologies for reusers to use this data and innovate. These courses of action, however, require broadening the number of actors involved in the management of the portal, and questioning the role and position of each as well as the associated means to achieve this new vision.

- The *ideal* courses of action aim to make the City of Montréal's data portal part of a broader political project to encourage the participation of a multitude of actors, by ensuring that reusers are a) committed over the long term, and b) are representative of a diversity of reusers; and that c) the portal is a platform for continuous learning in relation to the provision of the necessary tools and technologies; e) the City has the leadership to ensure that the interactions created flourish and are sustainable; and f) a space is created for the participatory production of data by and for citizens. This redefinition once again requires us to reflect on the involvement of various actors and their different missions.

Introduction

The purpose of this report is to provide a portrait of the reusers[1] of the City of Montréal's open public data portal.[2] This will involve, on the one hand, reviewing the main literature on open data in general and its use and, on the other hand, detailing the nature, content and profile of the various requests made in connection with the ODP between 2016 and 2019. It will also include semi-directed interviews with reusers, a more in-depth look at the type of impacts and use value of open data for each category of reuser, as well as an understanding of the data processing chain involved.

The City of Montréal implemented its ODP in 2011, followed in 2012 by a first open data policy (Dickner 2017), revised in 2015. The City's open data policy is an "open by default" policy (Ville de Montréal 2015, 5): all data is automatically open, a recurring choice in many municipal administrations internationally. Nevertheless, in the province of Quebec, only the City of Montréal, through the Laboratoire d'Innovation Urbaine de Montréal (LIUM) (or Montréal Urban Innovation Lab, in English), implements an "open by default." Since May 2018, the ODP has been administered by LIUM, who commit, among others, to "[e]nsure the transparency of the government apparatus both internally, most notably through the open and humane use of data" by supporting "data users and [encouraging] the community to leverage the city's data in order to generate the maximum amount of social and economic value" (LIUM website 2020).[3]

In 2020, the new ODP was launched. At the same time, a Digital Data Charter[4] was published. Independent of the open data policy, this charter establishes a number of general values and principles surrounding the use and dissemination of digital data. The charter is intended to guide the ethical use of data.

In this report, we proceed in four stages. First, using a database of English and French scientific literature, we conducted an overview of existing research on open data and its use and determined a methodological approach (sections 1 and 2). In the second step, we presented a portrait and an analysis of the requests received by the Urban Innovation Lab. This allowed us to reflect on the questions that would be relevant to ask reusers of the City of Montréal's ODP, which we did using semi-structured interviews (sections 3 and 4). Then, after analyzing the responses collected from reusers of open data, we identified the strengths and limitations of the ODP and its current use, while suggesting several courses of action (Conclusion).

There are limitations to this report. First, the analysis of received messages is likely to contain errors, as a person may correspond using different email addresses. Second, the interviews

[1] Following other writings (Dymytrova and Paquienséguy 2019), we will speak in this report of "reusers" insofar as the data exists, is used by the administration that created it, and is then published for reuse.

[2] For the purposes of this report and as specified in the summary, we refer to the data as "public" open data in the sense that it comes from a public institution. It is a question of distinguishing these data from those produced by a private actor and which may be "open."

[3] https://laburbain.montreal.ca/en/about

[4] https://laburbain.montreal.ca/sites/villeintelligente.montreal.ca/files/25817-charte_donnees_numeriques_ang.pdf

conducted reflect the reality of only a portion of the portal's reusers, since participation in the interviews was voluntary. Moreover, the majority of these volunteers are users of the former ODP, which was in effect until December 2020. Third, this analysis does not consider reusers who download open data without interacting with the ODP team. Therefore, this report should be viewed more as a snapshot of a reality at a given point in time.

1. Open public data: A look at the key literature

The literature on open data is relatively extensive and is part of the body of thinking on the smart city that has abounded since the mid-2000s. The definition of what open data can be remains relatively consensual depending on the context: "Open data is defined as data that can be 'freely used, reused and redistributed by anyone—subject only to a requirement of no more than identical attribution and sharing [...] Open data can, in short, describe all dimensions of public administration—except, of course, when public safety, privacy or government revenues are at stake" (Dickner 2017, 1).[5] Open data analyses can be divided into three broad categories. First, there is the literature that documents the emergence of ODPs and their evolution. Next is a series of papers that examine the nature, strengths and limitations of the portals that have been created. Finally, a final series of analyses, still quite rare, paints a portrait of the reusers of these portals. It is this last category that will be more specifically detailed here. In parallel with the scientific literature, a number of reflections from people involved in open data processes complement these analyses by offering insights that are often more practice-oriented, allowing the issues surrounding open data policies to be elaborated. Given this context, we begin this report with a presentation of the main principles that guide open data policies, followed by a summary of the evolution and the different forms these policies can take. Subsequently, we will detail the types of reusers identified, and conclude by addressing the political dimensions of data.

1.1 Open data policies: two major promises

From a theoretical point of view, data openness projects are based on two major promises: a democratic promise on the one hand and an economic promise on the other hand.[6] The democratic promise is built around two central concepts: transparency of government, and citizen participation (Silberzahn and Saujot 2018). In practice, "transparency is seen as a means of renewing the legitimacy of institutions and their representatives" (p. 4). It is a question of accountability of public authorities, in other words, "of accounting for the functioning of institutions" (p. 4), but also of regaining the population's trust in the institutions (Janssen, Charalabidis and Zuiderwijk 2012).

On the participation side, in theory, the opening up of data is associated with the idea of deepening democracy, allowing every individual to make his or her voice heard. It is about creating the "common good" and seeing new possibilities for management and governance (Meszaros et al. 2015). Alongside this democratic promise is an economic promise that is tied to two major concepts: innovation and the creation of services and businesses. The underlying idea is that the

[5] According to the Secrétariat du Conseil du trésor du Québec, "open data is raw, structured data that is distributed in open format as digital files, with a user license guaranteeing free access and specifying the conditions of use" (Secrétariat du Conseil du trésor 2019). We prefer Dickner's definition, which is based on the Open Knowledge Foundation and does not focus on so-called "raw" data, an element also found in the International Open Data Charter (2018). Dickner also points out that in the case of Montreal, the term *data* is used "as an equivalent of document" (Dickner 2017, 2).

[6] We do not address here the 'open source' issues from which open data is derived, nor the design of the information as such, nor the link with how reusers are designed. For an overview of the existing debates, visit: https://www.gnu.org/philosophy/open-source-misses-the-point.en.html

release of public data will lead to the creation of new services. Such an idea is akin to the notion of feedback already present in the design of the smart city (Janssen, Charalabidis and Zuiderwijk 2012, 259). The principle of "open by default" is a reflection of the "ideology of communication," whose central idea is that the free flow of information is the foundation of any transformation of society. Several authors have described the origin of this belief within the cybernetics and counter-culture movements in the United States (Turner 2006), right up to the recent alliance between liberals and libertarians, one of whose desires is to have free and open access to as much information as possible in all public and private domains (Barbook and Cameron 1996; Breton and Proulx 2006), and the project for a global information society at the end of the 1990s (Ouellet 2016). The "open by default" data, coupled with the promises of democracy on the one hand and economics on the other, subscribes to a market-oriented model of democracy based on three core beliefs: 1) the belief in the optimal distribution of information and the promotion of a market-like democracy in which everyone freely chooses their consumption of information; 2) the belief that social problems result from, among other things, a lack of communication and that these can be solved through the use of communication and information technologies; 3) the belief that each individual is a source of information that can add value to communities and the economy.

Open data: a certain vision of citizenship?

Open data policies tend to convey a certain vision of citizenship based on two poles: 1) the citizen must be active; 2) the citizen must be innovative: "Citizens no longer want to be passive recipients of legislation that is considered 'inflicted' upon them but, rather, seek constructive ways to engage-contribute-use the formation of public policy as a means to enhance their civic responsibilities" (Weerakkody et al. 2017, 285). This vision of citizenship is associated with a number of values expressed by Cardullo and Kitchin (2019, 11): "personal autonomy, consumer choice, entrepreneurship, participation, feedback, negotiation and creation."

These two promises are at the heart of the different policies adopted in various contexts, particularly at the municipal level.

1.2 The evolution of municipal open data policies: Between supply and demand

Open data: Return to libertarian utopia?

The opening of data echoes ideas already present in the early beginnings of cyberspace: 1) contribution, cooperation and sharing between reusers (each user can contribute and participate); 2) absence of hierarchy (anti-statism, in particular); 3) elimination of politics (Turner 2013; Flichy 2001, Musso 2010; Peugeot 2018).

A look at the various municipal policies adopted over the last ten years highlights the adoption of two specific strategies that are not mutually exclusive: open data policies focused on the supply side and massive data publication, on the one hand, and open data policies focused on the demand side, on the other.

1.2.1 Supply-side policies

In the 1980s, data was a luxury product: its production cost was high, as was the cost of storing and analyzing it. The various technical advances in the field of computing and telecommunications have made it possible to multiply data production, while drastically reducing storage and analysis costs, leading some to speak of a veritable "data deluge" (*The Economist* 2010).

This overabundance of data has contributed to changing the balance of power between the various actors: data control and mastery have become a challenge, particularly for public administrations. This is how the first municipal digital platform initiatives emerged: ODPs.

Open data:
Some key international dates

2001: Open Access Initiative
2002: Creative Commons
2004: Open Knowledge Foundation
2007–2009: Open Government Act
2010: Open Govermment Licence
2013: Open Data Charter
2018: Revision of the preceding charter

After Meszaros et al. (2015, 13–14)

These early supply-side policies had a number of limitations. Among these is the low quality of some published data. Researchers point out that data volume is not correlated with data quality: "More data does not necessarily mean more knowledge" (Silberzahn and Saujot 2018, 11; Janssen, Charalabidis and Zuiderwijk 2012). On municipal portals, it is not uncommon to find datasets on sites that are of poor quality, relevance or openness (Kubler et al. 2018). Some datasets may be obsolete, poorly documented or in formats that are difficult to reuse. Some point to the lack of reuser focus (Gascó-Hernández et al. 2017; Dymytrova and Paquienséguy 2017), while others note and rank the quality of the datasets (Tim Berners-Lee 2009; Martin et al. 2013) to encourage the upgrading of catalogs and the recruitment of new professional profiles within public administrations (network architect, data analyst, etc.). Indeed, public administrations face various obstacles that hinder the release of data in the form of production, formatting or dissemination (Conradie and Choaenni 2014). Quebec is no exception (Boudreau and Caron 2016; Dickner 2017). Added to this is the idea that published data is not very sensitive (Silberzahn and Saujot 2018), which does not really provide a counter force, undermining one of the promises underlying data openness projects.

These initial limitations will call into question a major presupposition, namely that there are necessarily people who are interested in the published data as well as specific "audiences." In other words, it would be "enough" to publish the data to create interest and benefits. This is one of the reasons why some policies have become more demand-driven than supply-driven.

1.2.2 Pro-demand policies

Since the mid-1990s, the demand for data has been steadily increasing (Chignard and Benyayer 2015). The private sector has developed business models that take advantage of the exploitation of open data to provide value-added services. At the same time, on the civil society side, new local and international institutions have emerged to demand greater government transparency, to fight corruption and to foster citizen engagement and access to information.

This context, however, raises the challenge of engaging audiences (Ubaldi 2013; Weerakkody et al. 2017). Others consider the demand to be real, but find that the open data movement suffers

from a mismatch between the supply (or availability) of data and the actual demand (and uses) (Verhulst and Young 2016). They attribute this, in part, to the lack of intermediaries and collaboration around these data (Verhulst and Young 2016). According to several authors (Johnson and Robinson 2014; Turki et al. 2017; Landry 2019), the organization of events and competitions (hackathons) nevertheless made it possible to bridge the gap between supply and demand by stimulating the interest of a variety of stakeholders in open data.

One of the limitations most often identified is the low level of audience participation. For some, this is as much a matter of information as it is a lack of capacity of reusers to reuse and analyze data (Gascó-Hernández et al. 2017). For Silberzahn and Saujot (2018), it is access that needs to be looked at more closely: "Access in its broadest sense implies: the availability of physical resources to access content (computer, smartphone); the ability to understand it and make content usable (data literacy); the availability of institutional or collective resources (to disseminate the message and engage as many people as possible)" (p. 6). A number of researchers thus emphasize the need to train citizens in data literacy (Goëta and Mabi 2014) and computer programming skills that will enable them to live better in the digital world (Proulx, Garcia and Heaton 2014).

However, several surveys point out that audiences remain poorly identified. Foulonneau et al. (2014) note that open data that prioritizes demand over supply may reduce the openness of data to a simple service removed from any socio-political dimension. Conversely, Johnson and Robinson (2014) argue that open data provided as part of hackathons may promote a form of outsourcing of government functions with potentially negative outcomes for society. Finally, the scientific literature raises the risks and drawbacks of open data such as the violation of privacy, misinterpretation of data, forms of exclusion (Gurstein 2011; Zuiderwijk et al. 2014; Peeters and Widlack 2018; Kobayashi et al. 2018, Brandusescu and Nwakanma 2019) and, more broadly, the right to information (Johnson 2014; Robinson 2014).

How is data to be valued?

"In fact, the opening of a certain specific dataset has little value in itself: it is in the possibilities of cross-referencing between different datasets for the production of new information that open data draws its strength. As a result, the importance of 'interpreters' (who sort, analyze, simplify and create visualizations) is growing." (Silberzahn and Saujot 2018, 7).

1.3 The reusers

A few analyses attempted to paint a picture of the actual reusers of open data. Rather than users, researchers tend to speak of reusers: "[...] reusers, whether professional or amateur, value open data through the applications they create by transforming data into services for users, customers, consumers, citizens and the territory" (Dymytrova and Paquienséguy 2019, 2). These analyses all converge and indicate primarily the existence of six main types of reusers: individuals, businesses, academia, developers, NGOs and journalists. Nevertheless, as Françoise Paquienséguy and Valentyna Dymytrova point out: "Our survey revealed a variety of profiles of open data reusers that is problematic because it groups together, under a theoretically unifying heading, different professions, skills and hierarchical positions" (Dymytrova and Paquienséguy 2017, 8). Secondly, another element mentioned in the literature is the absence of an "average user," particularly because motivations are diverse among the different groups but also, sometimes, within the groups

themselves (Worthy 2015, 794). Another clarification to be made is that some identified reusers are also categorized as "infomediaries." To put it briefly, "infomediaries, both human and nonhuman, negotiate the gap between open data providers and end-users, and can take the form of service providers, portals, and platform" (Fast and Rinner 2017, 7). In other words, an infomediary makes data more or less accessible depending on the business model, by transforming raw data (Ferro and Orsella 2013; Jannsen and Zuiderwijk 2014). According to Johnson and Greene (2017), journalists, community groups, NGOs, businesses—in addition to academia—can be considered infomediaries. It is in this sense that Dymytrova and Paquienséguy (2017) refer to "professional reusers" when discussing infomediaries.

1.3.1 The non-professional reuser

The name of this public is delicate in that it refers to people who make a request of a personal nature without any link to a company or organization. It is difficult to speak of "citizen," as this gives a political dimension to the designated actor, or "resident or inhabitant," since the place of residence does not, a priori, matter. The literature tends to speak of "publics" (Worthy 2015) or more simply of civil society. By "non-professional" user we mean a user who uses open data for non-professional purposes. The literature highlights the extreme diversity of this group. According to Worthy's survey, "The survey and the statistics above point to Open Data users as 'specific' individuals with 'specific interests.' It is likely that they are, as with FOI at local level, a core political activists, and a wider less committed group of the curious, often seeking micro or non-political information" (2015, p. 795). Little work has been done on this type of reuser, so our knowledge is still incomplete: What does the non-professional user do with the data obtained? What impacts have these data had? Moreover, there may be several types of non-professional reusers. One can think of the beginner or amateur user, in other words, one who is not very familiar with databases and computer tools, but also of the expert reuser, whose knowledge in these two fields can facilitate the understanding, processing and even dissemination of these data. Some expert users could also be infomediaries. To our knowledge, however, this type of user remains undocumented in the literature. Nor do we know how data is processed by this type of reuser.

1.3.2 Businesses, entrepreneurs and developers

Businesses, entrepreneurs and application developers are among the groups that have a specific interest in open data, in particular with regard to expenses and published tenders, or in accessible data developers (Worthy 2015; Magalhaes and Roseira 2018). According to Dymytrova, Larroche and Paquienséguy (2018), this type of reuser does the following data processing: "1) data retrieval; 2) integration into an internal database; 3) format transformation; 4) data verification; 5) integration into a web or mobile application" (p. 23[7]). This category of reusers can also include "data scientists" whose functions Dymytrova, Larroche and Paquienséguy define as follows: "Data scientists use statistical and mathematical models to produce information with a view to decision support in order to respond precisely to the demands of their clients, who are mainly large companies" (2018, p. 28). It should be noted, however, that such people may be a part of municipalities as well as the academic community. According to the same authors, the processing chain is as follows: "1) data recovery; 2) data extraction and filtering; 3) data cleansing and

[7] Note that Dymytrova et al. (2018) only talk about developers when presenting the data processing chain.

structuring; 4) data lake integration"[8]; 5) calculations and analysis; 6) restitution layer" (p. 28). Once again, although the data processing chain is known, information on the impact of the data, the audiences affected and other aspects is still missing.

1.3.3 Data journalists

As mentioned earlier, knowledge of data reusers is still young. Regarding journalists, considered as infomediaries, it would seem that open data is used by a specific type of journalist, the so-called data journalist (Worthy 2015; Dymytrova 2018). This is a new form of investigative journalism that consists of systematically analyzing huge databases to make sense of them (see, for example, WikiLeaks' collaboration with major international dailies). However, Rodrigo Zamith (2019) notes that data journalism relies mainly on institutional sources and offers little original data collection. Most of the time, open data is used by journalists to document a subject: "[...] data published in open data allows 'context' journalism such as the Les Décodeurs project, which aims to 'set the facts straight' by mapping and visualizing data. In the case of data journalism, open data hardly serves as a 'scandal machine,' and rather as a support for argumentation" (Goëta and Mabi 2014, 88). According to Dymytrova, Larroche and Paquienséguy (2018), the chain of data processing by this type of reuser is as follows: "1) recovery; 2) format cleaning and conversion; 3) analysis and calculations; 4) data verification; and 5) preparation of a final production" (p. 35). To our knowledge, however, there is no portrait of these data journalists and the types of media (print, television, Internet, etc.) that they work with.

1.3.4 Researchers, professors and students

Open data represents untapped educational resources. However, the academic community makes marginal use of these accessible resources in teaching or training: "Some teachers construct pedagogical sequences that implement this new learning process of using and manipulating open data" (Capelle, Jutand and Morandi 2018, 20). Similarly, it is not known whether the academic community uses open data in the context of their research. For others, the academic community, by publishing reports and articles based on open data, constitutes infomediaries. To our knowledge, the chain of data processing by these reusers has not been documented.

1.3.5 Community groups and NGOs

Civil society organizations and activists (NGOs, NPOs, associations) are closely involved in the production, distribution and use of open data (Church 2017). First, they regularly publish their own research data (e.g., the state of corruption in each country). Second, they incorporate publicly available data to make it more accessible to reusers (e.g., arms sales or environmental health) (Juanals and Pascual-Espuny 2016). Thirdly, they reveal and make public information of public utility (e.g., bank files of offshore accounts). In this sense, they can also act as infomediaries.

1.4 Data, a political issue

This brief portrait indicates the existence of many grey areas, particularly with regard to the value of the data and its political dimensions, all of which call into question the democratic and economic promises that underlie the projects to open up data.

[8] A data lake is a method of storing data.

1.4.1 The political issues

One invariant lies in the strategies of the various open data policies implemented: data is a political issue (Lehmans 2018). Several elements need to be considered. First, data tends to redistribute roles between the different actors it concerns. Data also raises questions about the role of public administrations, namely with regard to their decision to publish certain data or not, as well as with regard to their role as regulators. Second, many note the existence of limited transparency and accountability (Goëta and Mabi 2014). These often remain linked to data on government spending. Similarly, for Verdier, "the transparency process encloses citizens in the data we want to give them" (2016, online). Ruijer et al. (2020) go further and point out that "[t]his discussion revolves around the question whether transparency can exist if there is no one to make use of it. Effective transparency requires receptors capable of processing, digesting and using information" (p. 14). Third, there is geopolitics and data diplomacy: from the ongoing battle over the use of personal data to international movements for open data and open government, data is also a sovereignty issue. Given that there is a cost to collecting, maintaining and disseminating open (sometimes politically sensitive) data in formats that are easy for the general public to understand and use, some authors (Hess and Ostrom 2007) question the most appropriate and legitimate institution to manage this commons. Similarly, a third wave of open data policies appears to be underway.

1.4.2 Towards value-driven open data policies?

There are limits to open data policies, whether supply- or demand-side, or both. In 2018, the overhaul of the International Open Data Charter marked a turning point, building in part on the limitations of previous open data policies. The Charter proposes to move toward a targeted publication strategy to "solve specific policy problems," "prioritize actions," and "integrate institutional change." In their book *Datanomics: les nouveaux business models des données* (Datanomics: the new business models of data), Chignard and Benyayer (2015) propose an approach to data value to understand the opportunities and risks associated with the new data landscape. They point out that the three foundations of data value are: work (data is worth what it costs to produce and obtain), scarcity (from access to the source of the data) and utility (data is worth what we can do with it). This triptych echoes the findings of the OECD (2019, see box). These same authors specify that data also has an option value, "that is, a value attributed to a potential future use" (Chignard and Benyayer 2015, 45). In other words, the direct or indirect uses of open data do not appear immediately: either because quantitative and qualitative thresholds must be reached for emerging uses, or because collaboration between actors from different and very distant sectors is sometimes required to be able to identify and extract value from them.

1.4.3 A deficient knowledge base

Creating value from open data

"Creating value from Open Public Data (OPD) involves increasing the amount of data and learning how to better identify high-value data to increase reuse. It also means understanding the barriers to putting public data to work for the public service value chain." (OCDE 2019, 16).

These elements call for a better knowledge of existing reusers, including their uses and professional practices in relation to open data, as well as for an exploration of potential and indirect reusers of these available resources. As Pichot Damon (2019a) points out, knowledge of downloaded datasets alone is not sufficient to provide an accurate picture of the reusers of open data. In order to understand the value of the data used, for example, a more detailed understanding of their impacts is required (Pichot Damon 2019b). The table below summarizes the knowledge that is currently lacking.

TABLE 1. SYNTHESIS OF MISSING KNOWLEDGE ON REUSERS OF OPEN DATA FROM MUNICIPAL PORTALS AND THE IMPACT OF THESE USES

	WHO ?				
	Non-professional users	Businesses, entrepreneurs, developers	Journalists	Academic community	Community groups and NGOs
Type of request	Non-political information / specific information	Specific data base	Specific data base	Specific data base	Specific data base
Data processing chain	**Not documented in the literature**	Documented	Documented	**Little documented in the literature**	**Little documented in the literature**
Type of use	**Not documented in the literature**	Application by example	Support a context with an example	Document a reality with an example	**Little documented in the literature**
Type of impact	**Not documented in the literature**	Who uses the data products? **Partially documented in the literature**	Who uses the data products? **Partially documented in the literature**	Who uses the data products? **Partially documented in the literature**	Who uses the data products? **Partially documented in the literature**
Value of the created use	**Not documented in the literature**	**Not documented in the literature**	**Not documented in the literature**	**Not documented in the literature**	**Not documented in the literature**

2. Methodology

Our methodological approach was carried out in two stages. First, we conducted an overall analysis of the messages received by the Montréal Urban Innovation Lab from 2016 to 2019, comprising 4,980 messages.[9] The objective was not to perform a fine and detailed analysis of each message but rather to capture the type of reuser and the type of request. It should also be remembered that the main purpose of this summary analysis was to better target the relevant questions to be asked in the next phase, being the holding of semi-directed interviews with users of the City's ODP.

Some emails were removed, such as spam, out-of-office emails, emails consisting only of a word of thanks, the scheduling or postponing of meetings, and automatic emails regarding site usage and updates. The high number of automatic emails contributed to a subsequent reduction in the number of emails to be analyzed. This also applied to messages sent on Twitter or via the generic address of the ODP. In the same vein, we did not retain the emails subscribing to the newsletter. We estimate that approximately 2,000 messages were not retained for analysis.

Most of the messages received were already pre-classified into several categories: internal question, external question, question for another team, opening request. We therefore began by gathering the messages belonging to the same category.[10]

Then, in order to develop a portrait of reusers, preliminary analysis categories were developed from the literature review:

1) type of reuser (non-professional, journalist, academia, community, business)
2) purpose of the request (opening up of data, complaint, etc.)
3) type of request (according to the classification offered by the form)
4) nature of the request, with reference to open datasets

These categories are not static. The reading of the messages allowed for their precision and adjustment. We first processed the data by year. This processing allowed us to trace the history of the ODP. It is worth mentioning that one and the same person can write several messages. Our count is, therefore, aimed at the actors rather than the number of messages. Moreover, we can only suggest trends. Indeed, because some emails were forwarded to a number of people, or because the original message body was at times not included, we cannot guarantee that emails were not counted twice. Other messages, referring to face-to-face meetings, were sometimes encrypted.

This initial analysis allowed us to sketch out an interview grid for the people on the ODP mailing list. This interview grid is based on the work of Paquienséguy et al. (2016) and the Pichot-Damon impact pyramid (2019b). Its main purpose was to retrace the user's path and the value and usefulness of open data (see the grid in the Appendix). The call to participate in the interviews was circulated on the portal's mailing list, which includes some 300 subscribers, and at a seminar held

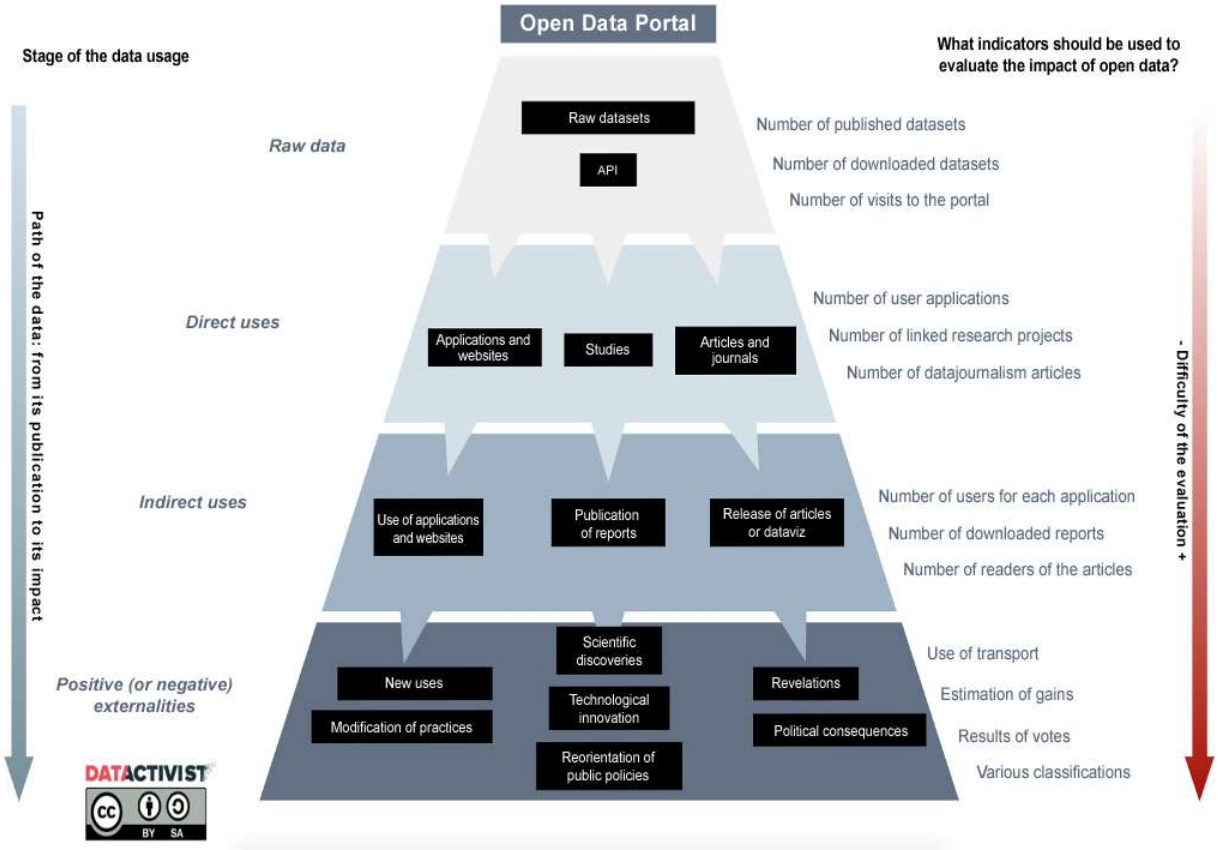
[9] Of these emails, the 33 emails received in 2020 have not been analyzed.

[10] On rare occasions, we have moved messages into other categories, considering that the initial classification was not adequate.

on June 17, 2020 presenting the new Montreal open data website, attended by some 60 people. This research was presented at this seminar.

Interested persons filled out an online form distributed via the ODP. This type of recruitment can therefore in no way be considered representative, since the people targeted are a priori people already interested in the ODP, having voluntarily registered on the portal's mailing list. Our analysis hence represents, instead, a one-time incursion into potential reusers of this portal.

IMPACT PYRAMID OF THE ODPs OF PICHOT-DAMON (2019B, ONLINE,TRANSLATED BY AUTHORS)



Nineteen people responded to this call and were interviewed between June and August 2020. Due to the context of the COVID-19 pandemic, these interviews were conducted remotely via a remote meeting tool. They lasted between 30 minutes and 1.5 hours, for an average of 45 minutes. These individuals were divided into different categories (see Table 2).

TABLE 2. CATEGORY AND NUMBER OF REUSERS WHO AGREED TO PARTICIPATE IN AN INTERVIEW

Category of reuser	**Number of reusers who accepted to participate in an interview**
Non-professional reusers	6
Academic community	4
Entrepreneurs	5
Civil servants and community organizations	2
Journalists	2
Total	19

The following section will present the analysis of the requests received by the ODP, followed by an analysis of the interviews.

3. Presentation of results

3.1. Presentation of received messages

The following discussion presents the main trends in this analysis. We will start with the evolution of the number of messages and the types of requests according to the reusers.

3.1.1 Evolution of the number of messages

The number of messages received was relatively stable from year to year, with a slight decrease in 2017 and a slight increase in 2019. These increases and decreases do not mean anything as such, as they do not reflect the number of messages that can actually be analyzed. As an example, the sender *Script_ckan_upload* generated 167 automatic messages in 2016. These automatic messages are not analyzable. In 2017, this same sender generated 210 messages and is used internally to exchange information. In 2018 and 2019, this sender disappeared. Similarly, the sender *Vue sur les contrats* generated 39 messages in 2016, 54 messages in 2017, 697 messages in 2018 and 854 in 2019. In sum, the number of emails received is not really significant, given the number of automatic emails present. From one year to the next, there were approximately 700 to 800 messages that were analyzable.

TABLE 3. NUMBER OF MESSAGES RECEIVED BY YEAR

	2016	**2017**	**2018**	**2019**
Messages received	1240	985	1216	1539

3.1.2 Internal request

The messages exchanged internally were organized around several themes, all of which bear witness to the history of the portal. In 2016, the elements addressed included:

a) the implementation of open data coordinators within the boroughs
b) the creation of data inventories in a range of sectors
c) the question of format
d) the question of reliability, the possibility of opening data
e) adjustment with other data, published by other organizations (CMM, Données Québec, etc.)

These different elements are evidence of the construction of the portal. In the subsequent years, we saw that the information exchanged internally was aimed at requests for data accuracy, contacts for opening up certain datasets, data updates and more comprehensive follow-ups. External requests were also exchanged internally. In general, the number of emails exchanged internally decreased between 2016 and 2019.

3.1.3 External request

Requests from outside the City were of several kinds. There are several types of actors:

1) non-professional reusers
2) developers, entrepreneurs and businesses
3) academic community (researchers, students, etc.)
4) data journalists
5) community groups and NGOs
6) civil servants of other organizations: Données Québec, City of Ottawa, etc.

It should be noted that it is sometimes difficult to identify with certainty the source of a request. We have cross-referenced the subject of the email with the nature of the email address used, although the reliability of this analysis cannot be guaranteed. Some people write to several different addresses, while others use their professional addresses to submit a personal request.

TABLE 4. NUMBER OF CATEGORIES OF ACTORS WHO SUBMITTED A REQUEST TO THE ODP FROM 2016 TO 2019

	2016	**2017**	**2018**	**2019**
Non-professional reusers	95	70	70	82
Developers/entrepreneurs/businesses	10	64	65	48
Academic community	21	39	38	37
Data journalists	8	5	11	11
Community groups and NGOs	9	9	9	4
Non-identified	19	92	49	0
Total	162	279	242	182

This table is likely to contain errors because the same person can write under different addresses/hats. For example, Sophie may write a first message using her Gmail address, signing with her name, Sophie X. She may then write a thank-you email using a business address of a magazine, likewise signing with her name, Sophie X. We then find her with yet another professional email address, this one associated with a recognized daily newspaper. In this case, it is extremely difficult for us to know if Sophie's first request was personal or job-related; similarly, it is difficult to know if there was a link between the other two messages and the person's professional function. Finally, how can we be sure that "Sophie X" designates only one and the same person?

Also, Table 3 lists the number of messages received year by year, hence not specifying the names of the senders. We subsequently identified the number of actors who sent at least one email between 2016 and 2019 (Table 4). This allowed us to see that 834 actors revolved around the ODP during this period, with an average of 200 users per year.

TABLE 5. NUMBER OF USERS WHO SENT AT LEAST ONE MESSAGE BETWEEN 2016 AND 2019

Number of reusers who sent at least one message	1109
Number of reusers who sent at least one message (not counting civil servants)	834

The requests, regardless of the actors, can be classified into different categories: 1) requests for data or for an opening up of data; 2) clarifications on specific data; 3) requests for information; 4) errors in a dataset; 5) contact requests; 6) requests for customized products; 7) requests that have nothing to do with the open data; 8) complaints (broken link, file not downloading). These requests have allowed us to identify several types of reusers.

3.2 The types of reusers

3.2.1 Requests from non-professional reusers

The requests from non-professional reusers are of several kinds. Some of them are not related to the ODP, and remain a very small minority. The portal is then used as a kind of single window. Such requests are related to everyday problems (truck on the street, traffic light duration too short, etc.). Other requests are of an extremely precise nature (e.g., a request for an archive, a request for an aerial photo of one's house).

In 2016, many requests questioned the existence of data rather than asking for an opening up of data. In subsequent years, by contrast, there were more requests for openings. There were also a few requests related to formatting, a non-functioning link or a complaint, although this remained rare. Among the complaints from citizens, a few questioned the reliability of the data and the level of transparency of the information transmitted.

Based on the writing of the emails sent by non-professional reusers, many are amateur reusers, even if the type of request or the frequency of emails sent at times indicate specific knowledge among these reusers, whereby they might qualify as expert users. One reuser was thus clearly identified as an expert and his action was even made public in the media. This person works in the field of information technology.

3.2.2 Requests from businesses, entrepreneurs and developers

For this type of public, one portion of the requests are related to API requests, mainly for snow removal operations.[11] The purpose of obtaining an API access token was not specified. Others concern basic data (cartography, number of vehicles, etc.) or details on the commercial exploitation of photographs. Many solicitations from entrepreneurs and businesses concern canvassing activities. It is rarely the same contractors or developers who write from one year to the next.

[11] Open data for snow removal operations is distributed via an API and requires a token from the Information Technology Department to access it. This allows for email interaction with reusers when there is a change in the API data attributes.

3.2.3 Requests from the academic community

Requests from the academic community often come from people presenting themselves as students, although some come from teachers or university librarians. These requests are for specific data (e.g., related to electoral data) or for access to certain data. There are also requests for interviews, most often related to international comparisons of data portals.

Other requests require clarification on the conditions under which the data were created. A few recent requests from the academic community indicate that open data are used as teaching resources in university courses. As in the case of businesses, it is rarely the same people who write from one year to the next.

3.2.4 Data journalists

It is often the same journalists who write from one year to the next, or at least the same media outlets. Their questions are aimed at requests for data or requests for clarification regarding certain datasets.

3.2.5 Requests from organizations (community-based organizations, NGOs, etc.)

Like the requests from businesses, those from associations and community organizations are extremely specific. We also identified Wikidata's contact with LIUM.

3.2.6 Requests from civil servants from other entities

Requests from civil servants from other entities are part of exchanges concerning datasets (with the Ministry of Transport for example, the CMM, etc.). Often, these requests are about adjusting data or continuing exchanges that have taken place verbally. Other requests come from cities or other institutions wishing to learn more about the Montreal experience or requiring assistance.

Messages from civil servants outside of the City of Montréal were counted as part of the internal messages, since requests between 2016 and 2020 came from the same set of organizations, who, based on the email exchanges, appear to have more regular contact with the City.

3.2.7 Recurring themes and datasets

It is often the same datasets that are discussed, namely those related to the environment, transportation, infrastructure, government and finance. The themes tend to be similar and revolve around the publication of the datasets (towing, parks, fire hydrants, parking, evaluation role, health inspection, vehicle, pedestrian, bicycle, skating rinks, trees, etc.).

3.2.8 Summary of the analysis of the messages

This summary analysis shows:
1) that there is a pool of approximately 200 people who each year contact the ODP to make a specific request;
2) that reading the messages did not allow us to find other types of reusers than those listed in the literature. Nevertheless, it might be interesting to look at the civil servants, other than those involved in operating the portal, of the municipality;

3) the requests from non-professional reusers, businesses/developers, organizations and journalists reflect what the literature has shown in other contexts;
4) that the immediate or future use of the downloaded open data is rarely explained (especially in the context of businesses/developers).
5) that the geographic location of the people writing is not disclosed.

These elements, in conjunction with Table 1, which concluded the review of the literature, enabled us to write an interview grid for reusers (see Appendix 1).

3.3. Presentation of interviews with reusers of the ODP

Over the course of the interviews, a number of common threads emerged in the reusers' comments. We will detail the responses obtained by presenting 1) the respondents (gender, profile and technical skills, workplace); 2) their expectations of the portal; 3) the type of requests made; and 4) the positive and negative points in relation to the themes addressed in our interview grid.

3.3.1 Non-professional reusers

A total of six non-professional reusers responded to our call. Their profiles have several similarities. They are men who, because of their profession or education, are able to read and analyze data. For example, there is a retired research officer, a web developer and a data analyst. The geographic origin of our respondents is mostly Montreal, with only one respondent residing in Quebec City.

The contact with the ODP was random for some, while others eagerly awaited it. Generally, this latter group people does not consult other open data sites. The type of requests they make varies and leads us to distinguish two profiles of non-professional users, who can be seen as the two poles of the same continuum: the passionate user and the aspiring professional.

For the passionate reuser, the motivations for using open data are personal and to varying degrees reflect a certain political involvement. For example, one reuser wanted to make available the number and quality of green alleyways in Montreal, and another wanted to document the number of dangerous rodents in a borough. Through this data mining work, these individuals wish to give back to the community what they are helping to create through open data. The aspiring professional is more of a data miner than a data explorer.[12] He works with open data to perfect his skills (a training that is largely self-taught) with the aim of increasing his employability. For many of these non-professional reusers, the boundary remains porous between personal, even political, motivations and the professional world.

These reusers come into contact with the ODP team when they start their projects, and then have very specific requests. They are satisfied with the answers they receive and resemble one another with regard to their path of data use. It begins with downloading, viewing, cleaning and restructuring. However, the degree of difficulty encountered by the users surveyed is not the same, according to the interviewee. For some, the cleaning phase is the most time consuming and for others it is the restructuring phase.

[12] March (1991)

The interviewees mentioned three negative points. One was the cleaning phase, which they often described as long and tedious. They did not find any file format to be ideal, depending each time on the nature of the project. Some users have created applications, a visualization of a phenomenon or a tool. Nevertheless, as they do not, in general, measure the impacts of these creations, these are—at best—estimated. Generally, these creations are used by less than a thousand people. One person who obtained a contract thanks to an application he had developed estimated that it was used by 3,000 people. Secondly, several respondents expressed being dissatisfied with the support received from public institutions when developing projects. According to them, such projects do not interest the institutions, in turn demotivating those involved in trying to launch them. Thirdly, all agree that the quality of the data is variable and that many data are missing.

Nevertheless, open data is an added value for their practices. In particular, it allows them to save a considerable amount of time. Some also pointed to the monetary gain, seeing that it allowed them to do without submitting an access to information request that is charged with a fee, depending on the nature of the data requested. None of the respondents download datasets for potential future use, although one respondent indicated that he routinely keeps a copy of the datasets he downloads.

In summary, the profile of the non-professional users surveyed can be summarized as follows:

- they are male;
- they have the skills to read and analyze data;
- they present a hybrid profile ranging from passionate users to aspiring professionals, meaning they are both data explorers and data exploiters;
- their expectations of the ODP are high;
- they identify three limitations to the current portal: 1) the data require too much time to clean up; 2) the lack of follow-up by public institutions on projects created from the data; 3) the variable quality of the data;
- they also identify three strong points: 1) added value for their practices, 2) time savings and 3) monetary gain.

3.3.2 Businesses, entrepreneurs and developers

Five businesses, entrepreneurs and developers agreed to participate in our interviews. Their profile was similar: men with a strong capacity for data analysis. They included two executives of companies specializing in data analysis, one consultant for a consulting firm, one self-employed programmer, and one civil engineer. Given their profiles, their motivations were professional, with the exception of the programmer, who used open data as a teaching resource to train his trainees. Four of them reside in the Greater Montreal area and one in Quebec City.

All of these reusers had been aware of the existence of the Montreal ODP since its inception. With the exception of two of them, they download datasets frequently (once a month) or very frequently (once a week). All members of this category are reusers of other datasets from other public institutions, community or civil society organizations or businesses. The use of open datasets by members of this category is used to enrich a database consisting of several datasets from a variety of backgrounds. Their data processing chain is similar to that of non-professional reusers. For the most part, these individuals produce deliverables to their clients illustrated with maps, spatial analyses and statistics. They are clearly data exploiters.

Three limitations were raised by these reusers. One reuser stopped using the City of Montréal datasets a few months after the service was made available to the population because the quality of the data was insufficient. The preparatory work required too much time and effort to obtain a usable dataset. Another reuser noted that the documentation provided for understanding how the data were created was insufficient. The presence of inconsistencies or errors in the dataset is not a problem for these users, who have the skills to overcome such difficulties, either by cleaning up the dataset or adding to other datasets collected in other databases to create the desired information. One of the developers expressed a kind of dissatisfaction with the lack of interest that public institutions seem to have when developing applications. He would like stronger support from public institutions for the creation of applications in an open innovation paradigm. A showcase of the developed applications, a single window or the opening of a direct channel would be appreciated. Two additional proposals for improving the service of open data were put forward: more visualizations and no more accents in datasets.

Open access to open datasets has several advantages for these users. One person estimated that the existence of these datasets represents time and cost savings of 5 to 10% in their business. The centralization of the information makes it possible to obtain the datasets more quickly without making requests to the various departments, which requires time and energy. Several months are sometimes needed to obtain the data. Two other people mentioned that quick access to this information allows them to improve the analyses by enriching the data. The reusers reduce the time spent obtaining the data and cleaning it, to increase the time spent on data analysis. They increase the quality of their services and deepen their knowledge of the situation under study. For still others, access to this data has a windfall effect. They use it to provide value-added service to a targeted clientele.

In short, the profile of these users is as follows:

- they are male;
- they have the skills to read and analyze data;
- they are data processors;
- they currently identify three limitations to the portal: 1) the quality of the data; 2) the lack of monitoring/support from public institutions when an application using open data is developed; and 3) the lack of information on how the data is created.
- they also identify three positive points: 1) time savings, 2) cost savings, and 3) improvement of services.

3.3.3 Researchers and students

Two researchers and two students agreed to answer our questions, two women and two men. These individuals have a high degree of technical autonomy in their use of the data as well as a great understanding of the issues related to the dissemination and use of open data for a wide audience. Both researchers work in health research structures. One of the graduate students holds a research chair on mobility while the other is studying computer science. They are all based in Montreal. The profiles of these individuals is very similar to that of other reusers with regard to the use made of the open data.

Their motivations for using open data are professional for researchers and academics and personal for students. Three of them use open data to characterize different urban environments in Montreal in order to understand the physical evolution of the urban environment and its impacts on population behaviours. One of the students, on the other hand, uses open data for personal reasons, to acquire competencies that are transversal to his field of study (computer science), in particular proficiency in spatial analysis. All the people use a very large number of open datasets. They are both exploiters and explorers. These datasets enrich a personal database fed by other data sources.

These individuals have identified several limitations to the current open data. For example, in the clean-up phase of the open data, the researchers indicated the need, at times, to ask the open data department to release certain columns whose information was restricted for privacy reasons. These actions are permitted through bilateral agreements between the research institution and the City. Sometimes, it is the nature of the data that poses a problem: one of the individuals regretted, in particular, that certain datasets were only available at the borough level of the City of Montréal and did not cover the entire island of Montreal. Another regretted the lack of a common format among all city portals in Canada, which prevents their comparison. However, these members were bothered by the lack of documentation related to each dataset and regretted that some older datasets are no longer available on the portal. Finally, two discussions with the students focused on the City's ability to enable a service to support the creation of collaborative data, understood as datasets created by and for citizens (crowdsourcing or participatory production).

Several positive points were also raised. Concretely, free access to open datasets allows researchers to enrich their databases. Such datasets support the production of knowledge (writing reports, scientific articles, public decision support) even if, according to the testimonies, these researchers would have been able to obtain these datasets without the existence of this portal. While open access to data is not the primary interest for researchers, the existence of the ODP has another advantage: the archiving and centralization of data by theme allows the identification of similar datasets (e.g., data around cycling or greening). It is more the serialization of datasets that is of great interest to researchers, insofar as it saves them time.

Summary

- They are two men and two women;
- they have the skills to read and analyze data;
- they are both explorers and data exploiters.
- they consider the limitations of the current open data to be: 1) the nature of the data; 2) the format of the data; 3) the lack of information on how the data is created;
- they consider the positive points of the current open data to be: 1) the enrichment of existing databases; 2) the possibility of data serialization.

3.3.4 Data journalists

Two journalist users agreed to participate in the interview, one man and one woman. This low number can be explained, as one person points out, by the low number of journalists specializing in data in Quebec and Canada. Their data analysis skills are largely self-taught and both rely on their colleagues. Their motivations for using open data are professional. Their data processing chain is exactly the same as for non-professional reusers. Sometimes, it is a matter of documenting

a specific topic using open data or using open data to create a "topic." Both are explorers as well as data exploiters. Their workplace is located in Montreal.

Their contact with open data was made as soon as the ODP was opened. The City of Montréal ODP is not the only portal they use, but they were looking forward to it. One of them said that he (or she) always checks the data offered, especially to capture the degree of transparency. That person also noted that he (or she) initially consulted it on a weekly basis, following which the frequency of his (or her) consultation declined due to the nature of the data presented.

Specifically, three main limitations were described. First, the data need a great deal of reworking before they were usable, which does not sit well with the limited time journalists have to write articles. Second, a lot of the data that is presented as being "not interesting" to write about is, according to one journalist at least, rather relevant. Still, neither of the journalists find that there is enough data that is "really interesting," albeit they consider this to apply to all of Quebec and not only the City of Montréal portal. Thirdly, the quality of the data is questioned, in particular because of its incompleteness.

Both individuals stress the importance of keeping this data open, even if it is not used clearly and immediately, as it could be useful in the longer term. The interviewees consider that in this sense the ODP is moving in the right direction. One of the two interviewees also noted that the data saves time, in that the ODP allows for quick access to information related to the data.

In short, the profile of these reusers is as follows:

- They are a man and a woman;
- they are self-taught in their training and improvement of data processing and analysis skills; they are surrounded by people who can make up for any lack of knowledge they may have;
- they are as much exploiters as they are data explorers;
- they consider the limitations of the current open data to be: 1) the degree of "interest" in the data; 2) the quality of the data; 3) the lengthy work involved in processing the data prior to exploitation;
- they consider the positive points from the current open data to be: 1) important data for the future; 2) time savings.

3.3.5 Other respondents (NGOs and community groups)

This last category includes three respondents who have a specific academic profile, although they work or have worked in public institutions as public servants or NPOs. Since these last interviews confirmed what the other categories of users said, we will quickly go over their observations.

The three respondents are two men and one woman. While all used open data in their work, the ways in which they initially engaged with the ODP differed. For one of the two, the contact with the portal is recent (May 2020) and did not provide direct access to the necessary data. Although this person seems to have data processing skills, he has a whole technical support team at his disposal. The other person was hired by a federal public institution because of his academic profile and skills. She has extensive knowledge of open datasets at all political levels in Quebec and

Canada. One of the respondents worked in Montreal and the other in Quebec City. Data processing is the same as for the other users.

The objective of all three respondents was to create new knowledge within the targeted institutions. They are mostly data exploiters. They identified the same limitations and strengths with respect to the ODP as the other respondents. However, we noted two differences in their comments compared to other users. In one case, the ODP made it possible to identify where the data sought was located. In another case, one of the two interviewees pointed out that, apart from the positive impacts of open data, there may be a negative impact consisting of the misinterpretation of the data, or the cross-referencing and meaning given to it.

In short, the profile of these reusers is as follows:

- They are two men and one woman;
- they have the skills to read and analyze data;
- they are data processors;
- their identified limitations of the current open data are identical with those of previous users;
- the identified strengths of the current open data are identical with those of previous users.

3.3.6 Summary of the analysis of the interviews

These results allow us to draw up an initial general portrait of all our respondents. In fact, despite some differences between the profiles, common points emerge between all respondents and their use of open-ended data (see Table 6).

TABLE 6. TYPES OF REUSERS OF THE ODP OF THE CITY OF MONTRÉAL WHO AGREED TO PARTICIPATE IN THE INTERVIEWS

	Non-professional reuser		Professional reuser			
	Passionate	Aspiring professional	Businesses, entrepreneurs, developers	Journalists	Researchers, professors, students	Community groups and NGOs
Type of request	Request for a personalized product / Specific database	Specific database	Specific database / Request for a personalized product	Specific database / Precisions about the datasets	Specific database / Precisions about the datasets	Specific database
Data processing chain	Wishes to learn by resolving the barriers.	Wishes to learn by resolving the barriers.	Without difficulty. Wishes to save time without cleaning the databases.	Learns by doing. Gives importance to the documentation.	Gives importance to the documentation.	Without difficulty. Wishes to save time without cleaning the databases.
Type of use	Exploratory approach	Exploratory approach	Exploitative approach	Exploitative and exploratory approach	Exploitative and exploratory approach	Exploitative approach
Type of impact	Takes pleasure. Self-guided learning. Development of personal competencies.	Development of personal competencies. Increase one's employability. Development of a service or application.	Development of a service or application. Assistance in decision-making.	Disclosure of information. Formation of opinion.	Knowledge creation. Assistance in decision-making.	Reinforcing one's knowledge of a sector of activity.

Generally speaking, the reusers who agreed to participate in our survey:

- are mostly men (15 out of 19 reusers);
- have the skills to read and analyze data or, failing that, are able to find others to perform these tasks;
- The businesses, entrepreneurs and developers, as well as community groups and NGOs, are predominantly data exploiters.
- Other reusers (non-professional users, journalists, researchers and students) are both explorers and exploiters of data.
- Three main limitations were identified by respondents:
  - data quality (includes nature, format, information)
  - monitoring of public institutions when a product is created from open data
  - the variable interest of published data

- Three main strengths of the ODP were identified:
  - an added value for their personal and professional practices (data that is useful later, improved services, etc.)
  - time savings
  - cost savings

A few respondents, such as entrepreneurs and students, also had some ideas about what an ideal ODP might look like:

- a showcase for applications and projects developed using open data
- a single window or hotline to all existing open data (including that of other organizations)
- collaborative data creation (created by and for citizens)

4. Main trends and findings on reusers and their approaches

The analysis of the messages received by the portal and the analysis of the responses of several reusers allows identifying certain trends regarding the data processing chain, the approaches undertaken and their objectives, and the resulting impacts. Three main trends can be identified:

- the existence of asymmetry among users in the data processing chain
- explorer- and exploratory-type approaches
- types of direct and indirect impacts

4.1 Asymmetry among reusers in the data processing chain

Data reuse consists of extracting data from their initial production context, in order to provide them with another framework for interpretation and processing within new social contexts. Between these two actions of extraction and reinterpretation, the data processing chain ensures data reuse. However, some users do not go beyond the consultation stage. Our results showed that some reusers directly consumed the information by consulting a map or downloading an aerial photograph. In the case of data reuse, as we have defined it, our results confirm that reuse through a similar data processing chain for all "professional" users is conditioned by the quality (regular

updating), reliability (verification of the data, presence of documentation explaining its production conditions and limitations) and interoperability of the data (Paquienséguy 2017).

We observed disparities in reusers' ability to assess the reuse of datasets. There is clearly a strong difference between non-professionals and professionals in their ability to identify a recoverable dataset and to recover it by producing understandable information from the City's datasets. Observation of differences in the extraction and processing of the data shows that users share great similarities in their usage paths with the data. They have the same digital tools and advanced technical skills. The differences are more related to organizational resources and a more refined understanding of the nature of a piece of information. In addition, the data we analyzed does not allow us to capture the number of people who "only" upload data to the portal without contacting the open data team.[13]

Finally, the majority of respondents who agreed to answer our questions were men. The small size of our sample does not allow us to draw more general conclusions about a possible gender imbalance in the use of open-ended data and the skills required to use it. Nevertheless, this observation invites us to pursue this aspect in greater depth with a greater volume of data (Brandusescu and Nwakanma 2019).

4.2 Exploration and exploitation

Understanding the disparities within the data processing chain leads us to believe that the improvement of the conditions for enhancing the value of open data is not only linked to the types of users but also to the consideration of the steps taken by the user. The use of open reuser data can be classified into two types (March 1991): exploratory and exploitational approaches.

Exploratory uses present hybrid profiles in terms of reuser motivations (sometimes personal, sometimes professional, sometimes both) and the goals to be achieved. Non-professional reusers, developers, students, entrepreneurs or even researchers emphasize the importance of learning in their use of open data. These approaches are related more to experimentation, creativity, but also to the acceptance of failure. What is sought is on-the-job learning and the development of knowledge and competencies that are transversal to a user's core expertise. For these reusers, the time spent processing datasets is of little importance. The process is more important than the objective.

[13] Note that the new data website that replaces the old portal displays the number of downloads and datasets uploaded to the City of Montréal's ODP. A more detailed analysis of the frequency of these downloads remains to be done: https://donnees.montreal.ca/ville-de-montreal/frequentation-du-portail-de-donnees-ouvertes

Exploitation-type uses are identifiable only to those who use certain data in the professional context of carrying out a paid project. The processing chain for data reuse should be as short as possible. Several respondents mentioned receiving a document in PDF format as a response to their request for information. The time spent reworking the information into a usable file discouraged many of them. Some then turned to other organizations for similar information.

4.3 Types of direct and indirect impacts

Whether they are professionals or not, reusers value open data by transforming data into services for themselves, customers, citizens or the community. For the most part, the direct and indirect impacts of these values have not been measured by users.

At the individual level, the identified impact is linked to a quest for individual performance and efficiency. The positive impacts are aimed at the acquisition of new skills in data analysis and statistical and cartographic processing. For some, this desire to learn on the job is part of a desire to increase their employability and become more productive. Some amateurs have published blog posts with their discoveries, have built websites to disseminate their knowledge or developed an application.

Conclusion

Far from the initial militant aspirations of proponents of free software, who criticized the open source movement for being nothing more than a mode of development without political vision (Stallman 2007),[14] the ODP is structured as a service to a population that places greater value on needs satisfaction than on the realization of a larger project. The user relationship is of a utilitarian nature, based mainly on the request for a product that is not available, the resolution of identified problems or the request for additional information on the origin of the data. On rare occasions, exchanges go beyond this framework of an operational and utilitarian relationship between a portal representative and an individual. Valuing open data does not seem to be at the heart of these exchanges. Under these conditions, the technical vision of digital technologies as a means of modernizing institutions and revitalizing democracy seems unlikely to be realized (Noveck 2009; 2012). As reported by Laëtitia Prido (2017), proponents of open data argued that it should allow citizens to access more and more information from governments in order to allow them to better understand the actions of the state. This openness would encourage them to collaborate and become involved in political decisions.

Currently, at the level of the reuser, the Montreal ODP does not yet seem capable of fulfilling all the conditions to promote transparency, collaboration and participation. However, this observation must be qualified. In the eyes of the users interviewed as well as in the eyes of those who send messages to the ODP team, the portal is, on the one hand, a forward-looking project, namely in that it raises a lot of expectations. As a result, its achievements often appear—in the minds of the reusers—to fall short of their expectations. Moreover, the majority of the reusers interviewed are users of the old version of the portal, and many see a continuous need for adjusting the new version of the web portal, launched in 2020.

Nevertheless, the portal's mandate does not always seem clear to those who turn to it. This imbalance between users' expectations and the ODP's ability to respond requires: 1) putting the ODP back into the context of the municipality and its competencies, and 2) thinking about courses of action in three stages: courses of action that can be carried out by the team supporting the portal; desirable courses of action that would require more substantial human and financial resources; and, finally, ideal courses of action that would make it possible to respond both to the limits identified and to develop the portal's strengths.

Actions that can be carried out by the team supporting the ODP and aimed at all reusers

Currently, the ODP is designed as a public service. However, the number of reusers remains quite low (based on the number of messages sent and interviews conducted) and is characterized by its ability to have some reading and data analysis skills. As a first step,

[14] https://www.gnu.org/philosophy/open-source-misses-the-point.en.html

therefore, it could be envisaged to work on improving the service offered while allowing a larger number of potential users to benefit from the service.

- **Action item 1**: In the short term, a page could be added on the new website presenting, in layman's terms, what open data is and its purposes. The way in which the opening up of the data is carried out, the chosen themes and other aspects could also be explained. This would complement the video currently available on the Urban Innovation Lab's website that explains what open government is all about.[15] Having pedagogical sections explaining the purpose of open data, the nature of the available formats and their use, and having training areas for both beginners and experts could help increase the number of users of the portal and better orient the messages intended for it.[16] This would better respond to the portal's aim to promote the accessibility of data and its use.
- **Action item 2**: In the short term, a web page could be created on the new portal site explaining the types of files and formats provided and their objectives. For example, videos presenting tutorials for using or exploring certain datasets could be made available. This would allow—again—to better respond to the portal's desire to promote data accessibility and valorization.
- **Action item 3**: In the short term, an annual survey could be conducted to determine the evolution of user profiles, user projects and the impact and value created by the data. Such a survey would offer the possibility of adjusting the open data service to the different types of reusers, be it those who consult the site to download/request a specific document, for example, or those who have larger projects.
- **Action item 4**: In the medium term, the definition of the data and an explanation of its construction could be added. When a dataset contains an error, it could be published to ensure the traceability of the datasets. The oldest datasets could also appear on the site.
- **Action item 5**: In the long term, develop a data quality charter that allows all of the City's boroughs to easily feed the portal with open data. The guidelines proposed by Données Québec could be a good starting point in this regard.[17]
- **Action item 6**: In the long term, in order to promote the enrichment and number of datasets available, a standardization of the construction of the data and of the governance surrounding it would be necessary within the boroughs. Given that the City of Montréal has 19 boroughs, each possibly differing with regard to their data management methods, this course of action will require a clear political commitment and adequate financial resources to achieve standardization in data construction.

These courses of action would make it possible to work on the quality of the data as well as to better identify what can be expected from a municipal ODP, in order to rebalance expectations and responses between users and the platform. This would make it possible to work on the transparency aspect of open data.

[15] https://laburbain.montreal.ca/documentation, useful links: https://www.youtube.com/watch?v=bln9ZGd26KE

[16] Note that the new ODP does explain what data is and how it is governed, albeit with insufficient emphasis.

[17] See the following site: https://www.donneesquebec.ca/lignes-directrices-sur-la-diffusion-de-donnees-ouvertes/

Desirable courses of actions

Desirable courses of action are those that would need to be put in place—especially to respond to reusers who wish to add value to a product made from open data, but not only. However, such courses of action require the involvement of actors and logistical support other than the team in charge of the portal. Currently, the City of Montréal's ODP qualifies the user as an individual "data consumer." This vision does not allow a follow-up of the products eventually created by users. Instead, it requires creating more opportunities for innovation by creating the foundations of a true ecosystem, involving the following action items:

- **Action item 7**: Allowing for networking between the different users to enable monitoring and creating interdependence between them. The added value of the portal consists, in addition to the publication of data, of generating relationships between actors. Some reusers want their products to be better showcased. The current portal has a page where the different applications created by reusers are presented.[18] The users surveyed did not mention this and stressed the desire for this enhancement to go beyond this page.
- **Action item 8:** Identify other data providers that could be useful to different users. Some reusers indicated that they would like to have a single window for all open data available in the province, rather than having different sites offering different data. Information about available datasets from other data providers could fill this need.
- **Action item 9:** Offer tools and technologies on the platform that users may need to innovate.

In other words, this would require an increase in the number of actors to manage the open data site, in order to enable it to meet the innovation objective underlying this type of approach.

Ideal courses of action

The ideal courses of action are those that require a redefinition of the ODP and therefore require the participation of a greater number of actors and the addition of significant human and financial resources to enable its realization. The main goal is to make the City of Montréal's ODP a broader political project—which could be independent of the municipality—aimed at encouraging the participation of a variety of stakeholders.

- **Action item 10**: Make sure that the actors are committed over the long term, to move away from a "one-time consumption" of data, without relationships with other actors. It is, therefore, a long-term commitment, allowing newcomers to exchange with the older ones.
- **Action item 11**: Ensure that a diverse range of actors around the ODP are represented. The Digital Data Charter speaks of "digital sovereignty," referring to data coming from a number of actors, a notion that could be extended to the actors involved.

[18] https://donnees.montreal.ca/applications

- **Action item 12**: Ensure that the platform is a continuous learning platform in relation to the availability of tools and interactions between the actors involved.
- **Action item 13**: Ensure that the City has the leadership to ensure the development and sustainability of the interactions created.
- **Action item 14**: Create a space for the participatory production of data by and for citizens.

These various courses of action invite us to rethink the definition and purpose of the City of Montréal's ODP. Within the scientific literature, there is no single, consensual definition of an ODP. However, most often, it considers that once open data, as generally defined, is put online by a government on a platform, it constitutes an ODP. Others emphasize the idea that an ODP is a set of services (Danneels et al. 2017), while still others refer more to stakeholders (Kassen 2018). There is, however, a common thread to the goal of an ODP, in addition to the promises discussed in the first part of this report: an ODP is intended to foster transparency, accountability, participation and innovation. Such objectives cannot be achieved without establishing ties between the actors revolving around such a system. Nor can it be achieved without redefining the notion of democratization, since access to data alone cannot guarantee the exercise of digital citizenship (Foth, Brynskov and Ojala 2015). Nevertheless, these different objectives should be put in relation to the specific context of the municipality and the competences and means at its disposal. Such an undertaking also implies clarifying everyone's roles: Should the municipality be responsible for this project? Does it fall under its mandate? Or, as is done in other cities, should such an undertaking be independent of the municipality?

References

Barbrook, R. and Cameron, A. (1996). The Californian ideology. *Science as Culture*, *6*(1), 44–72. doi: 10.1080/09505439609526455

Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography*, *3*(3), 274–279. doi: 10.1177/2043820613513390

Boudreau, C. and Caron, D. (2016). *Les données ouvertes dans l'administration publique québécoise : Utilités, freins et pistes de solution.* http://www.enap.ca/cerberus/files/nouvelles/documents/La_recherche/NetGouv_Les donn eesouvertesdansadministrationquebecoise_2016.pdf

Brandusescu, A. and Nwakanma, N. (2019). Gender equity. In B. Walker, T. Davies and F. Perini (Eds.), *The state of open data: Histories and horizons.* (p. 287–299). Cape Town and Ottawa: African Minds and International Development Research Centre.

Breton, P. and Proulx, S. (2006). *L'explosion de la communication: à l'aube du XXIe siècle*. Montreal, Quebec: Boréal: La Découverte.

Capelle, C., Jutand, M.-A. and Morandi, F. (2018) Stratégies pédagogiques pour diffuser la culture des données ouvertes, Revue Cossi, n°4, https://www.revue-cossi.lerass-ceric.com/images/article-pdf/RevueCOSSI-n4-2018.pdf#page=20

Chignard, S. and Benyayer, L.-D. (2015). *Datanomics: les nouveaux business models des données*. Limoges: FYP éditions.

Conradie, P. and Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, *31*, S10–S17. doi: 10.1016/j.giq.2014.01.003

Church, James. (2017) Civil society and the open data movement. *A Quarterly Journal of Government Information Practice and Perspective, 45*(3).

Danneels, L., Viaene, S. and Van den Bergh, J. (2017). Open data platforms: Discussing alternative knowledge epistemologies. *Government Information Quarterly, 3*(3): 365–78. https://doi.org/10.1016/j.giq.2017.08.007.

Dickner, N. (2017). *Comprendre et manipuler les données ouvertes de l'administration publique. La situation au Gouvernement du Québec et à la Ville de Montréal*. Montreal: École de bibliothéconomie et des sciences de l'information.

Dymytrova, V. and Paquienséguy, F. (2017). La réutilisation et les réréutilisateurs des données ouvertes en France : une approche centrée sur les usagers. *Revue Internationale des Gouvernements Ouverts*, 117–32.

———. (2019). Des données aux services dans les Métropoles : enquête auprès des réréutilisateurs professionnels de données ouvertes. CODATA, France, DATA VALUE CHAIN.

Dymytrova, V., Larroche, V. and Paquienséguy, F. (2018). *Cadres d'usages des données par des developers, des data scientists et des data journalistes*. ANR OpenSensingCity 14-CE24-0029 – Elico – Livrable 3. https://www.researchgate.net/publication/324976707_Livrable_3

Dymytrova, V. (2018). Data journalisme, entre pratique créative innovante et nouvelle médiation experte? Une analyse conjointe des discours et des productions journalistiques. Paper presented at the XXI Congrès of SFSIC Création, créativité et médiations, Paris. https://hal.archives-ouvertes.fr/hal-01817597/document

Ferro, E. and Orsella, M. (n. d.). *Eight Business Model Archetypes for PSI Re-Us*. Paper presented at the Open Data on the Web workshop, Google Campus, London. https://www.w3.org/2013/04/odw/odw13_submission_27.pdf

Figer, J.-P. (2005). L'influence de l'évolution de la technologie sur les métiers des SSII ou 40 ans de surf sur la loi de Moore. *Entreprises et histoire*, *40*(3), 15. doi: 10.3917/eh.040.0015

Foth, M., Brynskov, M., and Ojala, T., (Eds.) 2015. *Citizen's right to the digital city: Urban interfaces, activism, and placemaking*. Singapore Heidelberg New York: Springer.

Gascó-Hernández, M., Martin, E. G., Reggi, L., Pyo, S. and Luna-Reyes, L. F. (2017). Citizen co-production through open data: cases of citizen training and engagement. In *Proceedings of the 18th Annual International Conference on Digital Government Research—Dg.o '17*, 562‑63. Staten Island, NY, USA: ACM Press. https://doi.org/10.1145/3085228.3085252.

Goëta, S. and Mabi, C. (2014). L'Open Data peut-il (encore) servir les citoyens ? *Mouvements*, 2014, sect. 3, n°79.

Gurstein, M. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, *16*(2). https://firstmonday.org/ojs/index.php/fm/article/download/3316/2764

Hess, C. and Ostrom, E. (Eds). (2007). *Understanding knowledge as a commons: From theory to practice*. Cambridge, Mass: MIT Press.

Janssen, M, Charalabidis, Y. and Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 2012, sect. 29.

Janssen, M. and Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review*, *32*(5), 694–711. doi: 10.1177/0894439314525902

Johnson, J. A. (2014). From open data to information justice. *Ethics and Information Technology*, *16*(4), 263–274. doi: 10.1007/s10676-014-9351-8

Johnson, P. and Robinson, P. (2014). Civic hackathons: Innovation, procurement, or civic engagement?: Civic hackathon: procurement or civic engagement? *Review of Policy Research*, *31*(4), 349–357. doi: 10.1111/ropr.12074

Juanals, B. and Pascual-Espuny, C. (2016). La santé environnementale à l'ère des données ouvertes: le cas de la qualité de l'air urbain en PACA. *Colloque international TICEMED, Xe édition*, laboratoires I3M (Universités Nice Sophia Antipolis, Toulon), CREM (Université de Lorraine, Haute-Alsace), IRSIC (Aix-Marseille Université), Oct 2016, Marseille, France ⟨halshs-01569942⟩

Kassen, M. (2018). Open data and its institutional ecosystems: A comparative cross-jurisdictional analysis of open data platforms. *Canadian Public Administration, 61*(1): 109–29. https://doi.org/10.1111/capa.12251.

Kobayashi, S., Kane, T. B. and Paton, C. (2018). The privacy and security implications of open data in healthcare. *Yearbook of Medical Informatics*, *27*(1), 41–47. doi: 10.1055/s-0038-1641201

Kubler, S., Robert, J., Neumaier, S., Umbrich, J. and Le Traon, Y. (2018). Comparison of metadata quality in ODPs using the analytic hierarchy process. *Government Information Quarterly, 35*(1): 13–29. https://doi.org/10.1016/j.giq.2017.11.003.

Landry, J.-N. (2019). Urban development. In B. Walker, T. Davies and F. Perini (Eds.), *The state of open data: Histories and horizons* (p. 226–236). Cape Town and Ottawa: African Minds and International Development Research Centre (n. p.: n. p.).

Marcon, C., Grosjean, S. and Mallowan, M. (2018). Introduction: Méthodes et stratégies de gestion de l'information par les organisations: des "big data" aux "thick data", *Revue Cossi*, n°4, https://www.revue-cossi.lerass-ceric.com/images/article-pdf/RevueCOSSI-n4-2018.pdf#page=20

Magalhaes, G. and Roseira, C. (2017). Open government data and the private sector: An empirical view on business models and value creation. *Government Information Quarterly*. doi: 10.1016/j.giq.2017.08.004

March, J. (1991). Exploration and exploitation in organizational learning. *Organization science*, *2*(1), 71–87.

Martin, S., Foulonneau, M. and Turki, S. (2013). 1–5 stars: Metadata on the openness level of open datasets in Europe. In E. Garoufallou and J. Greenberg (Eds.), *Metadata and semantics research* (vol. 390, p. 234–245). Cham: Springer International Publishing. doi: 10.1007/978-3-319-03437-9_24

Meszaros, B., Samath, S., Guérin-Hamdi, S. and Faure, C. (2015). *Livre blanc sur les données ouvertes*. ISH.

Noveck, B. S. (2009). *Wiki government: How technology can make government better, democracy stronger, and citizens more powerful*. United States: Brookings Institution Press.

Noveck, B. S. (2012). *Demand a more open source government* [Web conference]. Technology, Entertainment and Design. https://www. ted.corn/talks beth _ noveck _ demand _a_ more_ open_ source _government#t- 1025995

Ouellet, M. (2016). *La révolution culturelle du capital: le capitalisme cybernétique dans la société globale de l'information*. Montreal: Éditions Écosociété.

OECD (2019). *Rapport sur les données ouvertes publiques: Encourager la maturité des politiques de données ouvertes pour un impact durable*, Paris: Éditions OCDE.

Paquienséguy, F., Larroche, V., Peyrelong, M.-F., Vila-Raimondi, M., Dymytrova, V. (2016) *Synthèse des résultats de l'enquête auprès de ré-réutilisateurs de données ouvertes: livrable n°1.* [Research report] SciencesPo Lyon; Enssib; Lyon3. 2016. hal-01432124

Peeters, R. and Widlak, A. (2018). The digital cage: Administrative exclusion through information architecture – The case of the Dutch civil registry's master data management system. *Government Information Quarterly*, *35*(2), 175–183. doi: 10.1016/j.giq.2018.02.003

Pichot-Damon, Étienne. (2019a) L'open Data, pas assez sexy pour la smart city? Datactivist, https://medium.com/datactivist/lopen-data-pas-assez-sexy-pour-la-smart-city-b7c0c5633905

Pichot-Damon, É. (2019b). Des analytics à l'impact, pourquoi évaluer une politique open data ?, *Datactivist*, https://medium.com/datactivist/des-analytics-%C3%A0-limpact-pourquoi-%C3%A9valuer-une-politique-open-data-461504a34eca

Prido, L. (2017). *Données ouvertes, gouvernement ouvert: fondements épistémologiques et discours.* Dissertation. Montreal: Université du Québec à Montréal, Master's in communication.

Proulx, S., Garcia, J. L. and Heaton, L. (Eds.). (2014). *La contribution en ligne: pratiques participatives à l'ère du capitalisme informationnel*. Quebec: Presses de l'Université du Québec.

Ruijer, E., Grimmelikhuijsen, S., van den Berg, J. and Meijer, A. (2020). Open data work: Understanding open data usage from a practice lens. *International Review of Administrative Sciences, 86*(1): 3–19. https://doi.org/10.1177/0020852317753068.

Secrétariat du Conseil du trésor. (2019). *Défi A- Redonner les données publiques aux citoyens.* https://www.tresor.gouv.qc.ca/information-resources/plan-daction-pour-laccessibilite-et-le-partage-des-donnees-ouvertes/defis/defi-a-redonner-les-donnees-publiques-aux-citoyens/?L=2.

Silberzahn, L. and Saujot, M. (2018). *Stigmatisation ou statactivisme : quelle utilisation de l'ouverture des données dans la ville numérique réelle ? Pourquoi "ouvrir" les données ne suffit pas*. Audacities.

The Economist. (2010). The data deluge. February 2010.

Berners-Lee, T. (n. d.). Linked data—design issues. In *Linked Data*. https://www.w3.org/DesignIssues/LinkedData.html

Turki, S., Martin, S. and Renault, S. (2017). How open data ecosystems are stimulated? In *Proceedings of the International Conference on Electronic Governance and Open Society Challenges in Eurasia—eGose '17* (p. 179–187). St. Petersburg, Russia: ACM Press. doi: 10.1145/3129757.3129787

Ubaldi, B. (2013). *Open government data: Towards empirical analysis of open government data initiatives*. OECD Working Papers on Public Governance 22. Vol. 22. https://doi.org/10.1787/5k46bj4f03s7-en.

Verhulst, S., and Young, A. (2016). Open data impact when demand and supply meet. Key findings of the open datat impact case studies. https://www.thegovlab.org/static/files/publications/open-data-impact-key-findings.pdf.

Vetrò, A., Canova, L., Torchiano, M., Orozco Minotas, C., Iemma, R., and Morando, F. (2016). Open data quality measurement framework: Definition and application to open government data. *Government Information Quarterly, 33*(2): 325–37. https://doi.org/10.1016/j.giq.2016.02.001.

Ville de Montréal. (2015). *Politique de données ouvertes.* http://donnees.ville.montreal.qc.ca/portail/wp-content/uploads/2015/10/Politique-de-donn%C3%A9es-ouvertes.pdf.

———. (2019). *Politique des données numériques, version préliminaire.* http://ville.montreal.qc.ca/pls/portal/docs/page/prt_vdm_fr/media/documents/politique_de_donnees_ville_de_montreral.pdf.

Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U. and Dwivedi, Y. K. (2017). Open data and its usability: An empirical view from the citizen's perspective. *Information Systems Frontiers, 19*(2): 285–300. https://doi.org/10.1007/s10796-016-9679-1.

Worthy, B. (2015). The impact of open data in the UK : Complex, unpredictable and Political. *Public Administration*, *93*(3).

Zamith, R. (2019). Transparency, interactivity, diversity, and information provenance in everyday data journalism. *Digital Journalism*, *7*(4), 470–489. doi: 10.1080/21670811.2018.1554409

Zuiderwijk, A., Helbig, N., Gil-Garcia, J. R. and Janssen, M. (2014). Special issue on innovation through open data—A review of the state-of-the-art and an emerging research agenda: Guest editors' introduction. *Journal of theoretical and applied electronic commerce research*, *9*(2), 1–2. doi: 10.4067/S0718-18762014000200001

Annex 1

Interview grid for reusers on the LIUM mailing list

	What we want to know	Sample questions
Types of users	**Who?** Ordinary reusers, businesses, entrepreneurs, developers, academia, organizations, journalists, others **Location**	- What is your function? In what capacity do you use the ODP? - What are the first three characters of your postal code?
Type of request	**Nature and frequency of the request** 1) Requests for data or opening; 2) Clarification of specific data; 3) Requests for information; 4) Errors in a dataset; 5) Contacts; 6) Requests for customized products; 7) Requests that have nothing to do with the opened data; 8) Complaints (broken link, file that does not download); 9) Other? **Object of the request** The issue(s) concerned	- What is your general request? (see list on the left) - How often? - How long has it been going on? - Which datasets are involved? - Do you use data other than the portal data? - How did you learn about the city portal? - What is the most suitable format for you?
Data processing chain	**Nature of data processing carried out** Description of the path taken to find and exploit the data	- Can you describe how you process the data (downloading, cleaning, verification, validation, etc.)?
Type of use	**Direct use:** - Number and projects related to the data - Nature of the project	- When you have the requested data, what do you do with it and for what purpose? - What are the resulting projects, and what are they? - Do you develop applications, which ones? How did the project come about? What is the objective of the project? - Are you sometimes missing data?
	Indirect use: - Use of the project resulting from data opened by people other than you (application, research reports, etc.) - Type and nature of audiences affected	- Who are the people who benefit or are targeted by your project? (policy makers, residents, organizations, etc.) - What is their profile? How do you measure the impacts of your project? Are you in contact with potential users?
Type of impact	**Nature and number of potential impacts**	- In your opinion, what are the impacts of your project based on open data? - Example: new uses, use of a new technology, changes in practices, improvement of knowledge, scientific discoveries, innovation, reorientation of public policies, revelation, cost reduction, job creation/loss, better

		dissemination of information, etc. - In your opinion, are there any negative impacts?
Value of created usage	**Interest, utility value of the data for the user?**	- How well did the downloaded dataset meet your request? - How would you characterize your interest as a user in this dataset? - How would you rate the usefulness of the dataset to you? - Do you ever download or request datasets and tell yourself that it is not useful now but that it may be useful later?