Towards transparent municipal open data: risks, illusions and opportunities in a growing field

I. Steenhout[1,2] and L. Volinz[1]

[1] Crime & Society Research Group, Vrije Universiteit Brussels
[2] Square, Vrije Universiteit Brussels

DOI: 10.3217/978-3-85125-932-2-17

Abstract. In the last years, many municipalities started to embrace the potential of open data, who envision open data as a mean to substantially enhance transparency and accountability towards citizens, restore trust in public services and increase citizens' participation and engagement. However, the flip side of open data rarely surfaces when municipalities report on their open data successes. This paper attends to the dark side of open data by examining open data extracted using the FixMyStreet API, to report incidents of urban disorder, nuisance and minor crimes in the Brussels' streets, by visualizing the aggregated data through a dashboard. Our approach illustrates that open data has a malleable character and breaks with several of the eight 'Sebastopol principles' of open data: (1) Only a fragment of reported data is available; (2) The data is not primary, since information is sometimes added by third parties; (3) Interventions are not always added in a timely fashion, with little to no information on the handling of the incident. This results in inaccurate data; (4) Reasonable privacy restrictions are lacking. The open data pose a risk to citizens' and municipal employees' privacy, and is non-compliant with GDPR regulations. Given these limitations, we argue that there are risks of misinterpreting or misusing this fragmented and inaccurate data, leading to misinformed citizens and policymakers. These risks underpin the need for a transparent data policy that allows open data initiatives to deliver on their promises and enables citizens to meaningfully engage with the urban environment and its complexities.

1 Introduction

In this research, we focus on the quality of municipal open data and its unintended consequences. While there are several definitions for open data available, most descriptions have in common that data is considered 'open' when it can be collected and shared with others, to use as they wish, without restrictions on copyright or usage (a.o. Ayre and Craner, 2017, Máchová et al., 2018). Open data (OD) is often used exchangeable with big data (BD), but it is important to distinguish both concepts. BD is defined by its size: it is data that no longer be handled by traditional tools or

databases. OD, on the other hand, is characterized by its availability: it is available to anyone and can be used and re-used without restrictions (Janssen et al., 2015). OD can be 'big', but this is not necessarily the case. OD fosters new opportunities for both public servants and citizens but introduces some important caveats at the same time. We elaborate the risks, illusions and the opportunities of this field that is evolving continuously.

Under the impetus of technological advancements cities have digitized most information in recent years. Fostering a drive to elevate data to better information, city administrations globally started publishing parts of this data as OD. Cities historically always had a central role in the collection and generation of data on their territory and inhabitants. Data on housing, income levels, labor and road infrastructure are just a few examples. In a recent past, however, cities and municipalities simply put out summarizing reports and citizens had to request specific information (Ashlock, 2013). This turn to OD can be considered an extension of regular communication channels and traditional reports (Lauriault and Francoli, 2017). It has the potential to offer more and better information faster and on a larger scale.

The trend towards open government data (OGD), or OD commissioned by government-controlled entities, is visible globally but the stimuli for opening city databases to the general public is diverse. Some motivations can be traced to the ideology of freedom of information and democratic values (Janssen, 2011). Rationale is that revealing this data could increase transparency, citizen engagement and empowerment, and even co-creation of public services and policy. At the same time, openness of data is thought to decrease corruption and bring accountability. Information is considered a prerequisite for a properly functioning democracy and OGD can play a significant role in this process (Harrison and Sayago, 2014). It allows citizens to grasp and accept decisions that affect them directly (Meijer et al., 2012). Another motivation is rather economic in nature. Publishing this data and creating opportunities to link datasets might improve government efficiency and therefore reduce organizational costs. Furthermore, OGD can create significant additional economic value due to the development of innovative tools that build on this data (a.o. Janssen, 2011, Styrin et al., 2017).

The 'techno-fetishist' perspective (Jefferson, 2020), that focuses on the upside of OGD, has been criticized by several scholars. Habib et al. (2022) argue, for example, that OGD differs significantly from government reports. These traditional reports usually contain extensive analysis of context-bound data, collected through well-defined methods, and were converted to human-readable format. OGD, however, is 'raw data'. It is often void of contextual information and should therefore be handled as a 'black box'. Furthermore, government institutions tend to operate in a closed culture

that might prevent the disclosure of data (Huijboom and Broek, 2011). Managers can fear that publishing data might illustrate a lack of operational effectiveness. In fact, there often seems to be a lack of understanding of the value of OGD as an advantage that can increase public services' effectiveness (Ruijer et al., 2020). Other scholars warn that OGD is rarely checked and lacks options to assess the quality of the data, which can lead to decisions based on data of poor quality or even data that is non-compliant with legislation (a.o. Zuiderwijk and Janssen, 2014).

As argued by Janssen et al. (2012), OD itself has little intrinsic value; the actual value is created by its use. A prerequisite for optimal value creation, however, is having quality data to start from (Lessig, 2017). In 2007, thirty prominent open data advocates gathered in Sebastopol, California to formulate the fundamental principles leading to qualitative OGD (Malamud et al., 2007). Today, these eight 'Sebastopol principles' are still considered the corner stone for efficient OGD and are the foundation of most guidelines for publishing OGD, allowing to reap the full potential of OGD and, at the same time, providing the necessary conditions for data usability (Attard et al., 2015). The principles state that open government data should be:

(1) *Complete*: All public data is made available. Public data is not subject to privacy, security or privilege limitations.
(2) *Primary*: Data is collected at the source, with the highest possible level of granularity, not in aggregate or modified forms.
(3) *Timely*: Data is available as quickly as possible to preserve the value of the data.
(4) *Accessible*: Data is available to the widest range of users for the widest range of purposes.
(5) *Machine processible*: The data should be reasonably structured to allow automated procession.
(6) *Non-discriminatory*: Data is available to anyone, with no requirement of registration.
(7) *Non-proprietary*: Data should be available in a format over which no entity has exclusive control.
(8) *License-free*: Data is not subject to any copyright, patent, trademark, or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed.

Despite the benefits of OGD, and the concrete guidelines to achieve qualitative data that can be reused in order to create added value for internal and external users, we hypothesize that the importance of qualitative data is still highly undervalued and associated risks are neglected. Therefore, we focus in this paper on the quality assessment of OGD based on the Sebastopol principles. If issues are found, the second objective is to reflect on the (un)intended consequences of identified issues, and how these might be solved. FixMyStreet Brussels (FMS) is used as a case study (CIBG, 2013). FMS is a popular state-citizen app that allows citizens to report incidents

of urban disorder, nuisance and minor crimes in the streets to the municipality. Incidents and (case based) follow-up reports by the municipality are made available to the public as open data (CIBG, 2018b).

2 Methodology

This study uses an instrumental case study design (Stake, 1995), based on the FMS OGD. A novel dashboard is proposed that allows to assess the data quality of the FMS OGD and evaluates this data using the eight Sebastopol principles. The construction of this dashboard involved a series of procedures, starting from the data collection, data structuring to the development of the dashboard.

2.1 Data gathering

Three distinct Application Programming Interfaces (APIs) were used. An API works according to a set of predefined rules that explain how a third-party application can retrieve data from a host application. It is an intermediate layer that processes data transfers without offering insight in the business logic behind this process. This set-up ensures that, while data is exchanged, internal procedures remain shielded.

To leverage the FMS OGD, the platform's API was used. While the API was not made available to the public until 2018, historical data since the introduction of FMS is accessible. 171.185 incident reports were collected from Feb. 12, 2013 until Dec. 31, 2021. These incidents were associated with 560.587 interactions between citizens and public servants.

Since the FMS data contains Belgian Lambert coordinates for the localization of incidents, the Google Geocoding API was used to convert addresses to latitude and longitude (Google, 2022a). This simplified the plotting of incidents on interactive maps via the Google Maps API (Google, 2022b).

2.2 Data structuring

The JSON data, provided by the FMS API, is already structured and machine processible. Therefore, parsing and storing the data resulting from the FMS and Google GeoCoding API in a normalized MySQL database was straightforward. This was done via a Python 3.10 script. Since the parsed data did not contain the original language of the incident report, langdetect was used to detect and add the language when the incident was reported in textual format (Shuyo, 2022). Additionally, a Python script was used to obtain the district associated with each incident report via the Brussels' 'Wijkmonitor' form (BISA, 2022).

2.3 Data analysis

To explore the data in this case study, an extensible digital dashboard was designed. The dashboard supports all elements to apply the Sebasopol principles to the FMS OGD but is constructed to allow additional features for data analysis. The application was developed using the Flask web framework (Ronacher, 2010), following an Model-View-Controller (MVC) design that offers extensibility through modularity, and separates the internal logic from the way information is presented to and accepted from a user. This set-up allows on-the-fly exploration of the machine processed data and presents it in a human readable format.

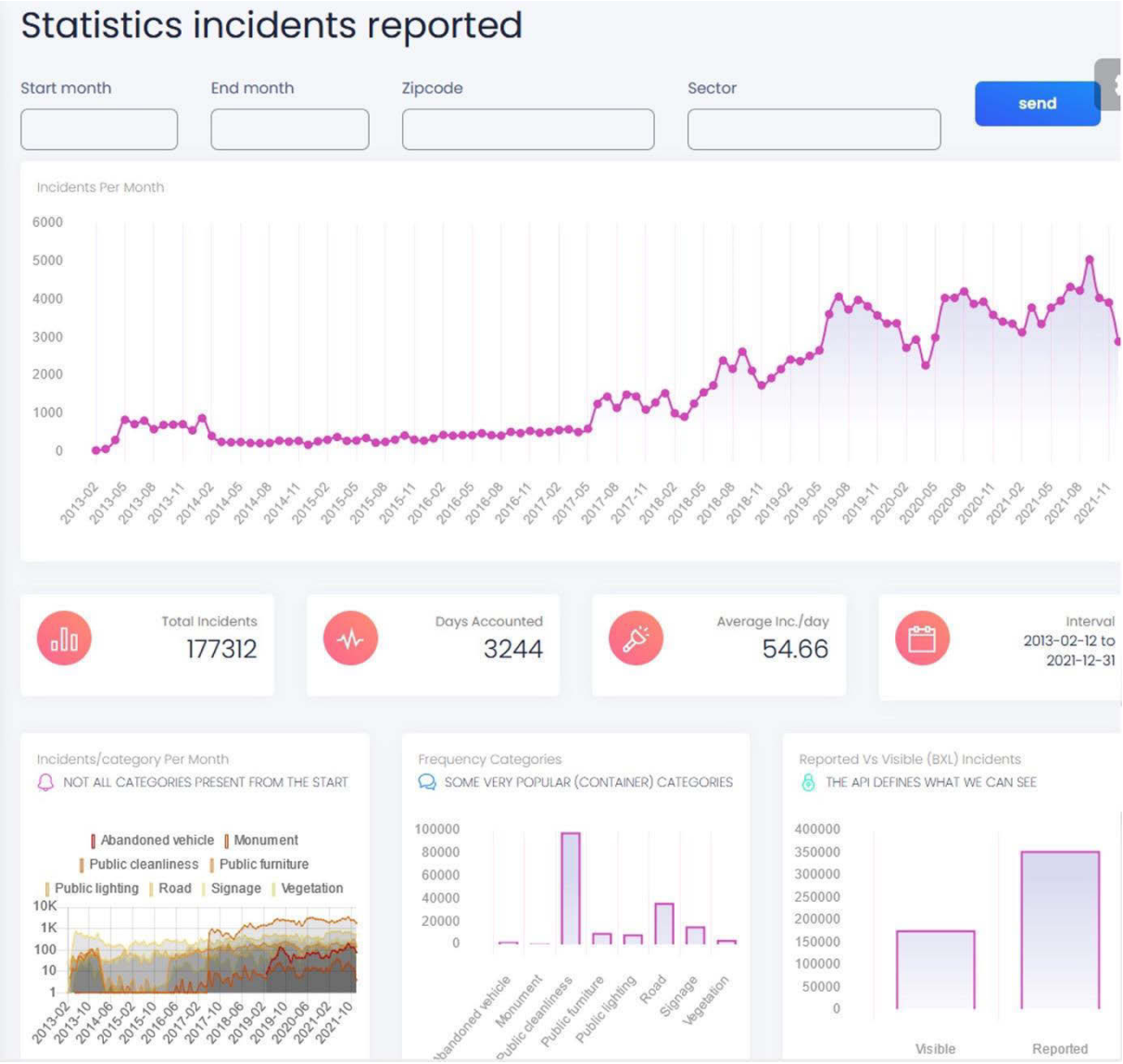

Figure 9: Overview of general data in the dashboard based on FMS OGD

The prototype dashboard has four main sections: the first part presents an overview of summarized data in FMS (Fig1). This includes the number of incidents per year, the distribution of incidents over the years, the average number of incidents, etc. It is

possible to filter based on district, time interval or zip code. A second page presents the localization of incidents (Fig. 2). It is possible to filter per incident and set the granularity of the representation (heat maps, per sector, or at street level). The third page offers insight in the handling of incidents by municipal services. It is possible to check what incidents are handled by which public service, to compare the handling time per public service and/or district or to get a summary of the open and closed incidents, as well of the incidents being processed. Filters can be applied based on year, zip code and district. Finally, the fourth page allows users to search the open data. Text search is possible, as well as filtering on zip code, district, main and/or sub-category of the reported incident.

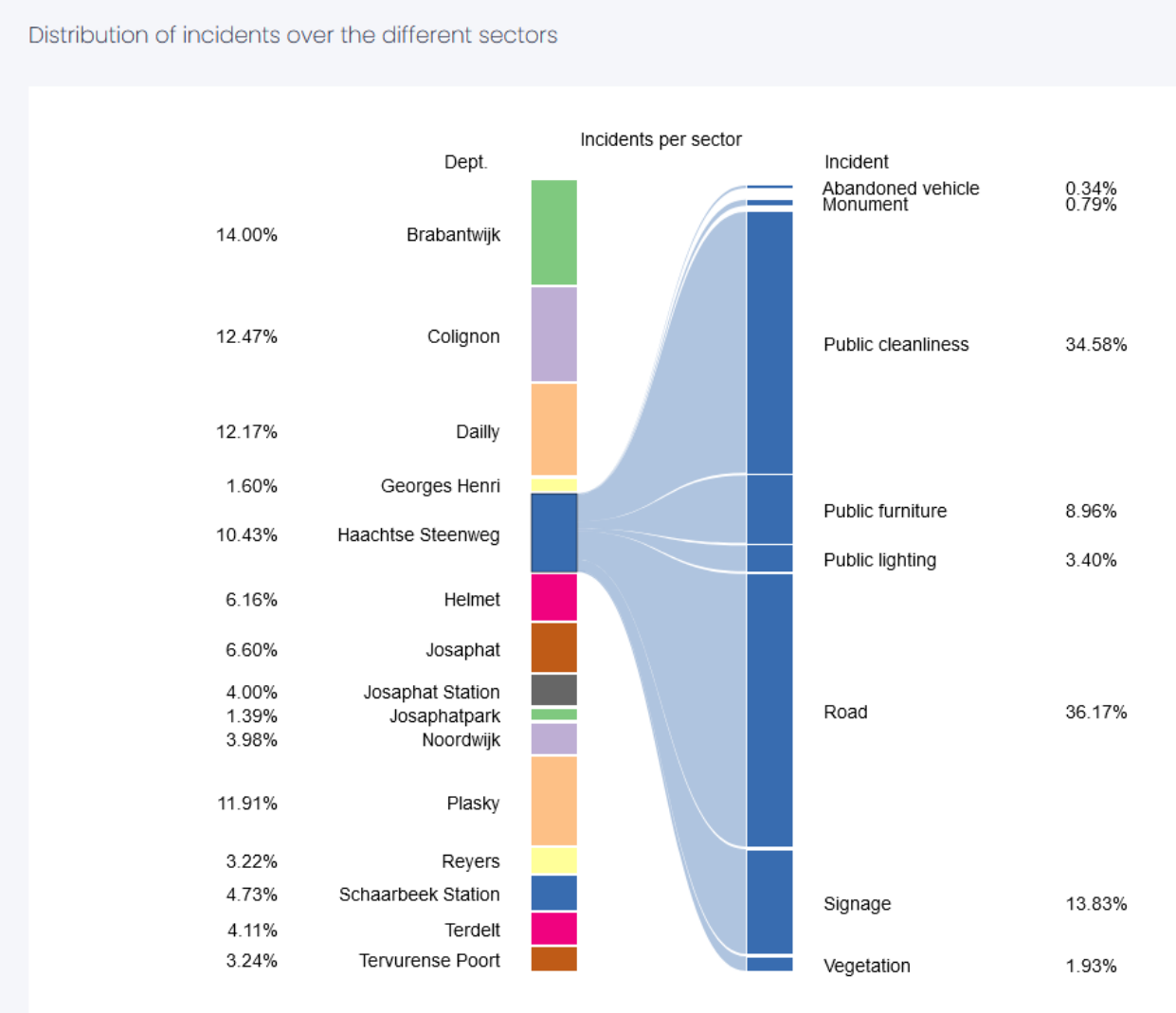

Figure 10: Detailed view of incidents per district in a municipality

3 Results and analysis

3.1 Complete

The completeness principle states that data should be open by default. However, data can be concealed, for example for privacy or security reasons. In these cases, the rationale behind hiding specific elements in the data should be transparent. Transparency is what allows citizens to see what and why actions are taken by the government. This principle breaks with the traditional way that a government interacts with its citizens. OGD builds on the presumption of publication for all, whereas traditionally a citizen had to ask public servants for the desired information. When data is not published, governments should justify why this data was obscured.
When reviewing the FMF OGD, a total of 171.185 incident reports were collected through the FMS API. All reports are uniquely identified by an incremental key. However, verifying the key of the last incident reveals that by Dec. 31, 2021, a total of 354.868 incidents had been reported since the activation of FMS in 2013. Hence, only 48,24% of all reported incidents are visible.

It is obvious that some type of selection is taking place and selection always equals data bias. An API allows a data provider to restrict access to the data, without delivering insight into the rationale behind this decision. As documented by the 'Centrum voor informatica van het Brussels Gewest' (CIBG), the data wasn't open to the public until 2018 (CIBG, 2018b). 'Openness' was conceived as an 'afterthought', a step taken five years after the application was implemented. Whereas a 'transparency-by-design' approach, that starts with 'outlining the objectives and what should be accomplished using transparency' (Janssen et al., 2017), this 'transparency-as-afterthought' seldom results in the desired level of transparency and understanding by the public since data is often 'patched' together without a clear vision and conceptualization of 'transparency'.

There can be good reasons for hiding data. OGD should be, for example, GDPR compliant. A general statement is found under the FAQ section on the FMS website claiming that incidents might be refused when (1) the incident does not concern a problem that has to be verified, (2) the incident was reported already, (3) the incident falls out of the scope of FMS, or (4) the content is abusive or illegal (Brussels Mobility, 2022). However, a closer inspection of the data through the dashboard's search module reveals that the data contains quite some citizens' private information. Search strings starting with '0475', '0486' (prefixes of Belgian mobile numbers) return mobile numbers of citizens with associated private information (name, address, etc.). Furthermore, the search string 'immatric' ('immatriculé' and 'immatriculation' refer to license plates) unveils several incident reports of littering where a citizen mentions the license plate of the offender. Other incidents are visible but were closed stating that

the incident falls out of the FMS scope or state that an incident was reported already. Additionally, duplicate notifications of incidents were found. These reports seem to contradict the FAQ and since a clear data policy is lacking, it is difficult to pinpoint why specific incidents were hidden. The disclaimer relaxes the ambition to offer complete data even further:

"[…]it does not guarantee the adequacy, accuracy or completeness of such information or warrant that the above-mentioned website shall be continually complete and up to date in every respect. The information on this web portal may include inaccuracies of content, technical inaccuracies or typographical errors"

Without a clear data policy, it is opaque who is responsible and accountable for the data. Furthermore, there is no guarantee of the accuracy or the quality of the data. It is however apparent that only a fragment of the gathered data is available as OGD, and that this data is not GDPR compliant.

3.2 Primary

OGD should be primary. It is collected at the source, and unmodified. A primary data source offers a first-hand account of an event. Secondary sources are already one step removed from the initial observer and can introduce bias or distort the original observation. Primary data is therefore considered more reliable and authentic.

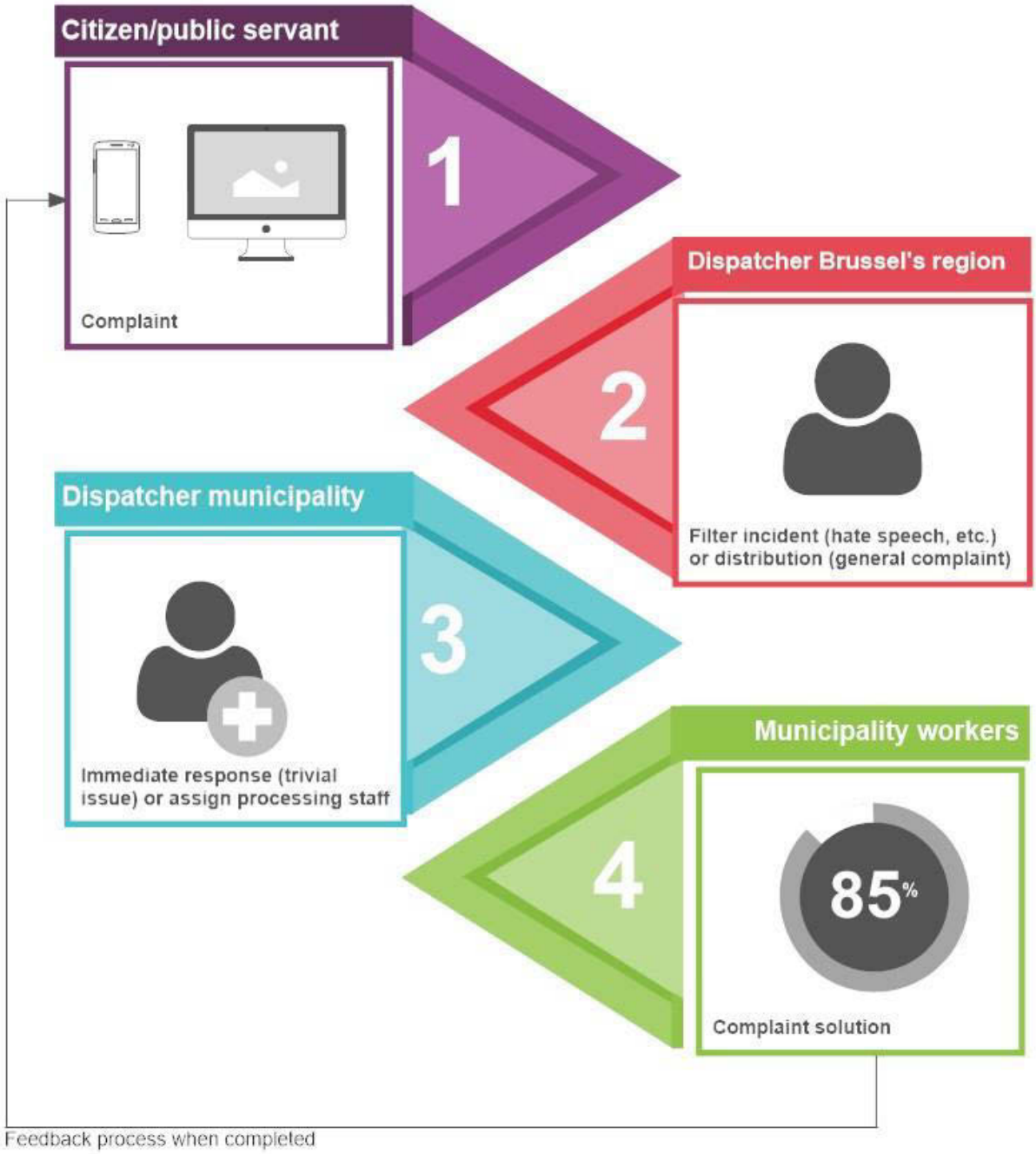

Figure 11: Communication flow triggered by an incident report in FixMyStreet (Steenhout et al., 2020)

To understand why the data is not primary, it is important to grasp the communication flow triggered by an incident report. Events can be an incident reported by a FMS user, or an action undertaken by a municipal service to solve an issue. The process from reporting an issue to resolving it, takes four steps (Fig. 3). First, a reporter can locate the problem on a map and add a text and/or picture describing the situation. The messenger does not need to worry which service is responsible. Based on the location and categorization of the incident, the incident will be filtered or directed to the correct municipality by the dispatcher at the Brussels' Region. In the next step, there is a back-end dispatcher at the municipality handling the incoming messages. This dispatcher decides what incidents will be handled and dispatched further. Some incidents are trivial in nature and can be answered and closed directly. If this is not the case, the

incident will be forwarded to the department responsible, where the department manager will assign staff to the issue to solve the problem. And finally, when the issue is solved, the user is informed on the progress.

Table 5: Handling time per service for a Brussels' municipality. Only services that handled more than five incidents over the observed interval were included. (*) Total of all closed incidents in the municipality.

					95% Confidence Interval for Mean	
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound
MIVB	87	416.80	602.903	64.638	288.31	545.30
Regional Dispatching	1679	147.31	392.349	9.575	128.53	166.09
Mobiris	383	102.69	211.116	10.788	81.48	123.90
Sibelga (lighting)	372	108.00	308.518	15.996	76.55	139.45
Net Brussels Group	**1276**	**.00**	**.000**	**.000**	**.00**	**.00**
Service Public Lighting	63	176.89	508.382	64.050	48.85	304.92
Municipal Roads - Dispatching	5223	49.44	198.430	2.746	44.05	54.82
Total*	9096	69.64	255.636	2.680	64.39	74.89

There is a potential conflict between the several steps to handle an incident and the principle of primary data. This becomes apparent when the actual handling time of incidents per service is analyzed. To illustrate, the closed incidents within a specific municipality in Brussels were analyzed for the interval 2013-2021.[39] Table 1 demonstrates an expected pattern for public services: a significant mean handling time for most services since outliers can be expected as several incidents will take considerable time to solve (e.g. fixing a pothole), whereas others can be fixed rather quickly (e.g. removing litter). This is also illustrated by the high standard deviations and, subsequently, the wide 95% CI for the mean. There is one exception, however. The data suggests that Net Brussels Group (cleanliness service for Brussels) is a highly performant service that manages to handle all 1276 incidents the same day of the incident report. The municipality was requested to provide insight into this phenomenon. The handling time visible in the FMS OGD is not the actual handling time of incidents by Net Brussels Group. Instead, this service does not work with FMS

[39] We opted not to name the municipality since this is not relevant as this text is not meant as a performance measure of municipalities, but instead wants to contribute to a better service of all municipalities using FMS.

and requested to close the incidents immediately. It was suggested that the service is highly performant and is usually aware of the problem before it is signaled in FMS. This strategy was not used in all municipalities. In several other municipalities, for example, the incidents regarding Net Brussels Group were closed by other services after a while. This introduction of automated handling of incidents, or closure by third parties, does introduce significant bias and no longer provides correct information to the citizen that reported these incidents. When a report for this service is filed, the signaler will be informed that the situation was handled while there is no guarantee of a (prompt) intervention at all.

Other types of automated messages were found that obscured the nature of the intervention, such as the notification that a report was being communicated to a specific service and consequently closed. In these cases, no detailed follow-up of the incident was possible. Incidents for several services were closed on behalf of a third, and hence non-primary, party, without guarantee that the closure time reflected the actual time to handle the incident or without detailed information on the interventions.

3.3 Timely

OGD has the potential to inform citizens and bridge the gap between citizen and, in this case, municipality by providing quick and correct feedback. This aligns well with the third Sebastopol principle, stating that data should be available as quickly as possible to preserve the value of the data. Ensuring that information is published fast increases the success of the application.

Apart from the previously described observation that not all reported incidents are visible in the FMS OGD, visible incidents are published without significant latency. We retrieved the data at 5 moments in time and each time incidents from the same day were already present. Incidents contain a timestamp for each intervention. This does not only inform citizens on a specific incident's status via the FixMyStreet data, but also allows third-party developments, based on the FMS OGD, to filter relevant data. If an application, for example, wants to offer an overview of the current incidents in a specific area it can focus on relatively recent incidents.

Inspecting the incidents that were being processed in a specific Brussels' municipality, however, triggered a red flag. While the vast majority of visible incidents were being handled by the services involved, none of the incidents that should be handled by the

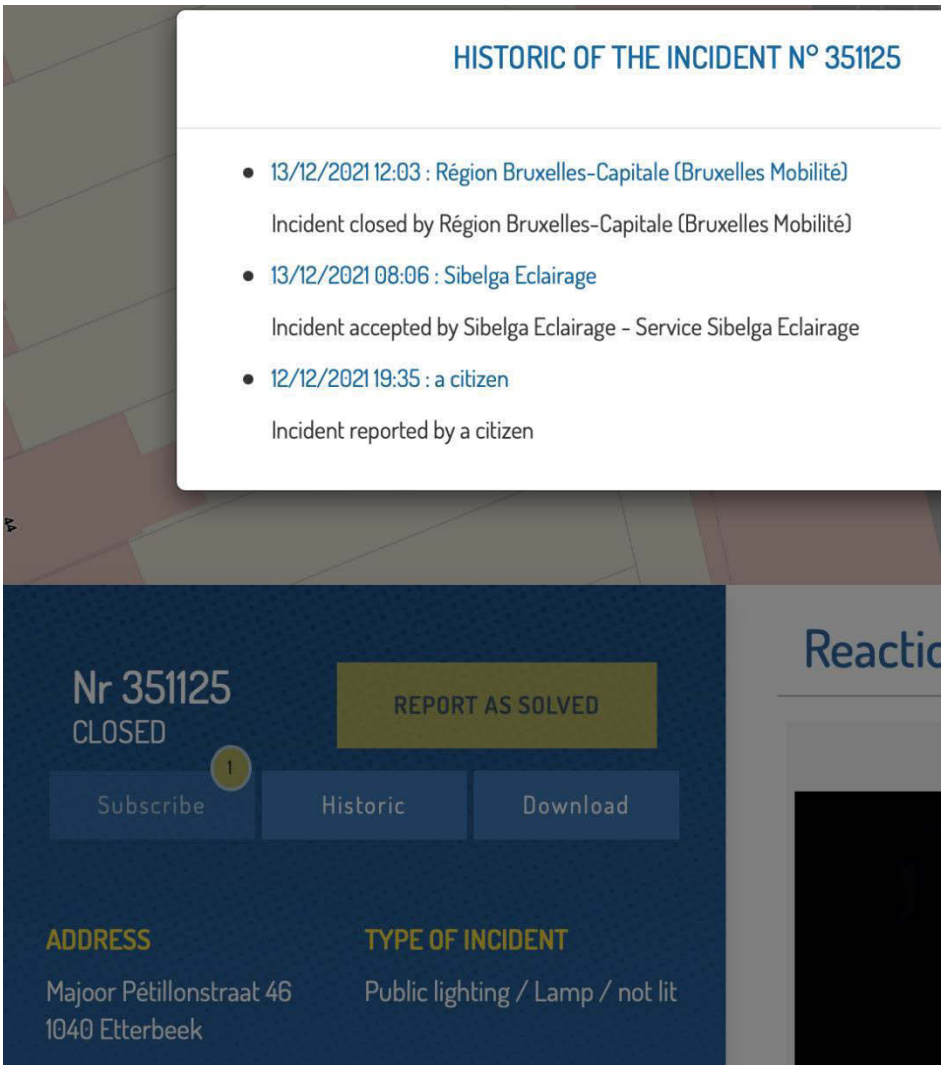

Figure 12: Citizen's report with 'track record' of interventions

service for community guards, nor the service Green Spaces, were closed between 2013 and 2022. After inquiry at the municipality, it became clear that the services intervened but worked with parallel systems to trace their interventions. These systems did not synchronize with the FMS application, leaving citizens in the dark. Reports kept the status 'processing' while, in fact, the issue might have been solved already.

Another issue relates to the use of standardized messages. Quite a lot of incidents mention that the 'incident has been closed by' a specific service. However, no information is given on the nature of the intervention. Fig. 4 illustrates the problem: a citizen reports that an area is not being lit. The next day, this issue is accepted by the service that handles lighting issues (Sibelga) and within 4 hours, another service closes the incident. The issue is flagged as 'solved', without mentioning the intervention made. This introduces ambiguity and leaves room for interpretation: Was this considered a non-issue and therefore closed after only 4 hours by another service, was it a matter of replacing light bolbs, was there a power failure in the street, or did Sibelga install new light lamp posts? As argued by Attard et al. (2015), success of an OGD initiative should "not only be evaluated on the amount of data published, but also on the usability of this data". The mere mentioning of an intervention, void of context, does not

guarantee transparency, nor does it bring accountability. The strength of OGD is in providing real content of value to citizens.

3.4 Accessible

Everyone should be able to use, reuse and redistribute the OGD. There can be no discrimination against private persons, nor public or commercial entities. Allowing reuse and redistribution for academic or non-commercial use only is prohibited. Furthermore, the data can be used for the widest range of purposes.

The FMS OGD is, in principle, available to the widest range of users. While this looks like an appealing realization, in practice obtaining 'access' is more challenging. Access to data is governed by a 'digital divide', or the gap between citizens with access to ICT's and those who do not. Additionally, it is also driven by a 'data divide' between citizens that have access to data and those who do not (Gurstein, 2011). Unlike bulk data that is easy to download, the use of the FMS API is limited to the rather tech savvy citizens. Furthermore, when the data can be retrieved, capitalizing on FMS OGD through additional analysis or linking the data with other data sources is quite complex. Developing new innovative technologies on the backbone of this data stays the privilege of software developers and will therefore be dependent on the interest and envisioned solutions of these developers. These solutions might not match the needs of citizens, nor complement services of the municipalities. Additionally, since the previous Sebastopol principles were not fulfilled, there is no guarantee on the quality of the data. The data is clearly incomplete and lacks important contextual information on the interventions. This limits the range of 'purposes', since inaccurate data will inevitably lead to inapt analysis or innovations ('garbage in, garbage out' principle).

3.5 Machine processible

The FMS OGD complies with the 5th Sebastopol principle: all incident related info is provided in JavaScript Object Notation (JSON) format by the FMS API. JSON is a lightweight data-interchange format that is easy for machines to generate and parse. It allows conversion to different data formats that allow in-depth exploration of the data. Furthermore, JSON promotes interoperability: it enables the linkage with other data sources and therefore increases the potential of developing new and better services on top of the FMS OGD.

3.6 Non-discriminatory

The 6th Sebastopol principle states that data should be available to anyone, with no requirement of registration. Considering that an API can introduce selection bias, the data that is offered by the API can be retrieved by everyone. The use of API is also documented and does not require a login, nor registration (CIBG, 2018a).

3.7 Non-proprietary

According to the 7th Sebastopol principle, data should be available in a format over which no entity has exclusive control. The FMS API serves data in a JSON format that is non-proprietary and released under an open license granting the use by everyone without any restrictions and at no cost.

The syntax is straightforward and does not require a specific program to explore its contents. JSON files can be opened in any text editor, without conversion, and the majority of programming languages support JSON. This setup guarantees that the format will not become out of date over time, nor that costs will be charged in the future to read these JSON files.

3.8 License free

The FMS OGD is available free of charge and under an open license. Therefore, it complies with the last Sebastopol principle. The data can be used to improve municipal performance but may also be used for commercial goals. It is thus allowed to make money out of the freely available FMS OGD by developing additional services.

4 Discussion

The Sebastopol principles are considered the corner stone for adequate access to OGD. In this paper the FMS OGD was scrutinized and juxtaposed with the Sebastopol principles. The data satisfies 3 principles: (1) it is machine processible, (2) the access is non-discriminatory since it does not require login, nor access fees, (3) the data is presented in a non-proprietary format and (4) is offered license free.

Other issues encountered related to the remaining four Sebastopol principles: (1) With less than 50% of the incidents visible, the data was obviously not complete. (2) Third parties sometimes close incidents on behalf of the responsible department or vague and ambiguous automated messages are used in response to incidents. This leaves the reporting and consulting citizens in a state of uncertainty. It is often unclear what type of, if any, intervention took place. Additionally, (3) since third parties can close incidents (sometimes several months later), the actual time of the intervention is unclear. This introduces additional bias in the data. Furthermore, the principle of accessibility is not satisfied. The quality of the data is unknown and can therefore lead to inapt solutions: poor data to start from, results in poor analysis that might lead to wrong prioritizations. It is also worth mentioning that providing OGD solely through an API limits access to the rather tech savvy sample within a population. In addition to the well-documented digital divide, this tends to create an additional data divide.

The ‘open’ character of OGD is remarkably often framed as translating the existing data and/or communications to a digital format. In the case of FMS, for example, there already was a ‘hotline’ available to signal incidents by phone. These notifications

triggered interventions as well and left a paper trail. Today, this hotline still co-exists with FMS and is illustrative of the interpretation of 'openness'. Notifications in FMS leave a digital trace and, unless the incident is obscured by public services, can be consulted via the platform or the API. Openness can also be interpreted more broadly however: it shouldn't be limited to providing data, but should be about "claiming access to knowledge and information" hidden behind gateways created by public resources (Shah, 2013). Transparency does not only depend on 'technology solutionism', or the incorrect belief that technology itself promotes transparency (Morozov, 2013). Effective use of OGD is ensuring that this data is translated into outcomes available to the widest range of citizens and ascertain that those who need it most are reached as well. Rationale behind the principle of accessibility is that the data should empower everyone. It could be argued, however, that access is mostly restricted to those who have access to the technology and the necessary technological and analytical skills to optimally extract information from the provided data. As argued by Gurstein (2011) "this would then suggest that a primary impact of 'open data' may be to further empower and enrich the already empowered and the well provided for. On the other hand, those most in need of the benefits of such new developments may find themselves out of luck".

While it is quite possible that incidents with abusive content, redundant notifications or privacy issues are – as stated in the terms of use - withheld from the data. Janssen et al. (2012) argue that public services only tend to publish data that are not sensitive, not very complex, or can do no harm. Upon inspection of the data, however, we did find quite some content that was not GDPR compliant. Without a clear data policy, it is impossible to assess the quality of the data.

The lack of a policy leaves room for tinkering with the data. In extreme cases, this might even lead to the opposite of what open data is trying to achieve. Instead of creating more transparency and a more just and equitable society, the use of APIs introduces the possibility of gatekeeping. Without a clear data policy, citizens are oblivious to the selection processes that take place behind the scenes. In extremis, only incidents and interventions that support a certain agenda might be published, whereas incidents that contradict this agenda might be obscured. It also leaves room for discrepancies between municipalities. Some municipalities might decide to open up more incidents to the public than others. A perverse effect might be the 'mislabeling' of certain areas as 'problem areas' simply because less incident reports are obscured. Shah (2013) warns that we cannot "ignore the politics of data themselves, what the data reveals, or how they are used and for whose interests". APIs can be extremely useful, but also put a lot of power in the hands of the data publisher. This, in turn, can – in the worst case - open the door for abuse and corruption. Fragmented data results in misinformed citizens and policymakers and thus introduces serious risks in the supposedly data-enriched citizen/government relation. Inaccurate or incomplete data might lead to misguided policy recommendations, misplaced funding and a complete misunderstanding of the situation.

Furthermore, once citizens realize that data can be tinkered with, this can backfire and create further mistrust in the public sector. Full comprehension requires a clear data policy that foresees in detailed context on actions taken by public servants, and if reports are withheld from the public, this should be justified. After all, openness is not limited to opening up fragments of data, it should also create clarity on the removal of content.

5 Conclusion

While there are no indications of malicious intent, this study reveals some disturbing facts about the FMS OGD. Despite the numerous efforts to invest in quality OGD by the Brussels Capital Region, some conditions and good practices to provide quality data are not met. Furthermore, the open data proved to be non-GDPR compliant. These concerns should be addressed. Citizens should be able to trust that the open data is accurate, fair and at the same time will not compromise their privacy. A clear data policy can deliver insight in the data that is made available to the public and should justify why some parts are kept closed, while better moderation – assisted by automated detection of private information - should guarantee GDPR compliance. A coordinated approach and clear strategy will be needed to fix these issues. We do recognize that OGD is still at its infancy and will continue to evolve and improve.

This paper identified the gaps in the existing FMS OGD and therefore contributes to a better understanding of steps that should be taken to provide a better service to citizens and to create more opportunities to build on the available OGD. These findings allow the Brussels Capital Region to address these issues and deliver improved OGD in the future.

As proposal for improvements and future work, we aim to further develop the dashboard and extend it with needs-based views that – through a user centered – approach will translate the data to meaningful output for the different types of users. Additionally, further research is needed to address the needs of citizens that do not have access to digital infrastructure. This should prevent that the exciting outcomes expected from OGD - once it reaches its state of maturity - are available for the widest range of citizens.

References

ASHLOCK, P. 2013. The biggest failure of open data in government [Online]. Open Knowledge foundation. Available: https://blog.okfn.org/2013/03/15/the-biggest-failure-of-open-data-in-government/ [Accessed May, 8 2022].

ATTARD, J., ORLANDI, F., SCERRI, S. & AUER, S. 2015. A systematic review of open government data initiatives. Government Information Quarterly, 32, 399-418.

AYRE, L. B. & CRANER, J. 2017. Open Data: What It Is and Why You Should Care. Public Library Quarterly, 36, 173-184.

BISA. 2022. De Wijkmonitoring van het Brussels Hoofdstedelijk Gewest [Online]. Available: http://http://wijkmonitoring.brussels/locate-sector/ [Accessed May, 8 2022].

BRUSSELS MOBILITY. 2022. FixMyStreet.Brussels: Frequently asked questions (FAQ) [Online]. Available: https://fixmystreet.brussels/faq [Accessed May, 8 2022].

CIBG. 2013. FixMyStreet.Brussels: Meld problemen in de Brusselse openbare ruimten [Online]. Available: https://fixmystreet.brussels [Accessed May, 8 2022].

CIBG. 2018a. API for developers with Smart City [Online]. Brussels: Brussels Smart City. Available: https://api.brussels/store/apis/info?name=fixmystreet&version=1.0.0&provider=admin [Accessed May, 8 2022].

CIBG. 2018b. FixMyStreet now available as Open Data [Online]. Bussels: CIRB. Available: https://smartcity.brussels/news-570-fix-my-street-now-available-as-open-data [Accessed May, 9 2022].

GOOGLE. 2022a. Google Geocoding API [Online]. Available: https://developers.google.com/maps/documentation/geocoding [Accessed May, 8 2022].

GOOGLE. 2022b. Google Maps API [Online]. Available: https://developers.google.com/maps [Accessed May, 8 2022].

GURSTEIN, M. 2011. Open data: Empowering the empowered or effective data use for everyone? First Monday, 16.

HABIB, A., PRYBUTOK, V. & PHILPOT, D. 2022. Assessing and building municipal open data capability. Information Systems and e-Business Management, 20, 1-25.

HARRISON, T. & SAYAGO, D. 2014. Transparency, participation and accountability practices in open government: A comparative study. Government Information Quarterly, 31, 513-525.

HUIJBOOM, N. & BROEK, T. V. D. 2011. Open data: An international comparison of strategies. European Journal of EPractice, 12, 4-16.

JANSSEN, K. 2011. The influence of the PSI directive on open government data: An overview of recent developments. Government Information Quarterly, 28, 446-456.

JANSSEN, M., CHARALABIDIS, Y. & ZUIDERWIJK, A. 2012. Benefits, Adoption Barriers and Myths of Open Data and Open Government. Information Systems Management, 29, 258 - 268.

JANSSEN, M., MATHEUS, R., LONGO, J. & WEERAKKODY, V. 2017. Transparency-by-design as a foundation for open government. Transforming Government: People, Process and Policy, 11, 2-8.

JANSSEN, M., MATHEUS, R. & ZUIDERWIJK, A. Big and Open Linked Data (BOLD) to Create Smart Cities and Citizens: Insights from Smart Energy and Mobility Cases. In: TAMBOURIS, E., JANSSEN, M., SCHOLL, H. J., WIMMER, M. A., TARABANIS, K., GASCÓ, M., KLIEVINK, B., LINDGREN, I. & PARYCEK, P., eds. Electronic Government, 2015// 2015 Cham. Springer International Publishing, 79-90.

JEFFERSON, B. 2020. Digitize and punish: racial criminalization in the digital age, Minneapolis, University of Minesota Press.

LAURIAULT, T. & FRANCOLI, M. 2017. Openess, transparency, participation. In: KITCHIN, R., MCARDLE, G. & WILSON, M. (eds.) Understanding spatial Media. London: Sage.

LESSIG, L. 2017. Ten principles for opening up government information [Online]. Sunlight Foundation. Available: https://sunlightfoundation.com/policy/documents/ten-open-data-principles/ [Accessed April, 16 2022].

MÁCHOVÁ, R., HUB, M. & LNENICKA, M. 2018. Usability evaluation of open data portals. Aslib Journal of Information Management, 70, 252-268.

MALAMUD, C., O'REILLY, T., ELIN, G., SIFRY, M., HOLOVATY, A., O'NEIL, D. X., MIGURSKI, M., ALLEN, S., TAUBERER, J., LESSIG, L., NEWMAN, D., GERACI, J., BENDER, E., STEINBERG, T., MOORE, D., SHAW, D., NEEDHAM, J., HARDI, J., ZUCKERMAN, E., PALMER, G., TAYLOR, J., HOROWITZ, B., EXLEY, Z., FOGEL, K., DALE, M., HALL, J. L., HOFMANN, M., ORBAN, D., FITZPATRICK, W. & SWARTZ, A. 2007. 8 Principles of Open Govenment Data [Online]. Sebastopol: Open Government Working Group. Available: https://public.resource.org/8_principles.html [Accessed May, 8 2022].

MEIJER, A., CURTIN, D. & HILDEBRANDT, M. 2012. Open government: connecting vision and voice. International Review of Administrative Sciences, 78, 10-29.

MOROZOV, E. 2013. To save everything, click here: technology, solutionism and the urge to fix problems that don’t exist, London, Pinguin Books.

RONACHER, A. 2010. Flask: Web development, one drop at a time [Online]. Available: https://flask.palletsprojects.com/en/2.1.x/ [Accessed May, 8 2022].

RUIJER, E., GRIMMELIKHUIJSEN, S., BERG, K. C. D. & MEIJER, A. 2020. Open data work: understanding open data usage from a practice lens. International Review of Administrative Sciences, 86, 3-19.

SHAH, N. 2013. Big data, people's lives, and the importance of openess [Online]. Connected Learning Alliance. Available: https://clalliance.org/blog/big-data-people-s-lives-and-the-importance-of-openness/ [Accessed May, 8 2022].

SHUYO, N. 2022. Langdetect: a language-detection library for Python [Online]. Available: https://pypi.org/project/langdetect/ [Accessed May, 8 2022].

STAKE, S. E. 1995. The art of case study research, London, Sage Publications Ltd.

STEENHOUT, I., VOLINZ, L., BEYENS, K. & MELGAÇO, L. 2022. The long road to Municipality 2.0: mobile city apps as catalyst for change? Brussels: Vrije Universiteit Brussels.

STYRIN, E., LUNA-REYES, L. & HARRISON, T. 2017. Open data ecosystems: an international comparison. Transforming Government: People, Process and Policy, 11, 132-156.

ZUIDERWIJK, A. & JANSSEN, M. The negative effects of open data: investigating the dark side of open data. 15th Annual International Conference on Digital Government Research, 2014 Aguascalientes Mexico. Association for Computing Machinery, 147-152.