Open Data Index for cities 2018: challenges for open data at the local level

Wagner Oliveira[1]
Davi Ventura[2]
Natália Mazotte[3]
Ariel Kogan[4]

Abstract

The Open Data Index for cities was created by FGV DAPP (Department of Public Policy Analysis of FGV) and OKBr (Open Knowledge Brazil) to assess the level of data openness in the local level. The evaluation was conducted in 2018 with data collected in 2017 from 8 selected cities in Brazil: Belo Horizonte – MG, Brasília – DF, Natal – RN, Porto Alegre – RS, Rio de Janeiro – RJ, Salvador – BA, São Paulo – SP and Uberlândia – MG. The study evaluated the degree of openness for 17 datasets for each city through a web-survey filled by local groups based on the Open Definition criteria and then reviewed by FGV DAPP and OKBr with the collaboration of the Open Knowledge Forum. The results show that only a quarter of the datasets evaluated were in full accordance with those criteria, while the main problems were insufficiency of metadata, unavailability of bulk download, datasets with incomplete and outdated information and absence of data in open formats. Those problems are also very common in the federal sphere, based on earlier assessments. The highest frequency of bottlenecks were found in Land Ownership, Company Register, Air Quality and Water Quality data. We also highlighted that little change was perceived from the last assessment (2016) for São Paulo – SP and Rio de Janeiro – RJ, even though some good practices were also highlighted. Lastly, we proposed some guidelines for open data policies in the local level.

[1] Researcher at Fundação Getulio Vargas, Department of Public Policy Analysis (FGV DAPP) and PhD student in Economics at the Brazilian School of Economics and Finance (FGV EPGE).

[2] Economics Undergraduate student at Universidade Federal Fluminense (UFF).

[3] Executive Director at Open Knowledge Brazil.

[4] Director at Instituto Tecnologia e Equidade (IT&E).

1. Introduction

"Data is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness" (OKFN, 2006). The Open Knowledge Foundation developed the global Open Data Index (ODI) to compare countries in terms of their ability to open data to citizens, to the press and to the civil society. The information used to compose the index is gathered through a crowd-sourced platform, that is, through the contribution of local agents connected to the OKFN network in each country, and then reviewed by specialists in open data in each country, generating a ranking of countries as the result.

The goal of the index is to evaluate the state of open data policies in each country, taking into account all of its characteristics: the type of data released, the formats, how easy it is to access data, the transformation of data into information, and more. The index evaluates several dimensions, such as public finance, socioeconomic, legislative and electoral data, public services, geolocated information and environmental indicators. Therefore, the index offers a parameter of reference about the ability to provide open data for all countries, presenting this information in a clear, easy-to-understand and easy-to-use way.

The global index, which had already been built for 2013, 2014 and 2015, had its data collection for Brazil in 2016 (OKFN, 2016) carried out by Open Knowledge Brazil (OKBr) and by the Department of Public Policy Analysis of Fundação Getulio Vargas (FGV DAPP), leading to the elaboration of the report "Open Data Index for Brazil" (RUEDIGER et al, 2017a). This publication pointed out the main bottlenecks found by Brazilian open data initiatives and highlighted good practices that may be replicated by the public administration. The main finding of the study was that Brazil can be considered a country with advanced openness and publication of information, but it still needs to develop in some aspects regarding easy access and the ability to promote understanding from data.

In addition to the global index, which compares countries, the index was replicated in 2016 for the subnational sphere: cities. In its Brazilian iteration, the methodology was adapted to be compatible with the reality in Brazilian cities. In this process, we highlight the existing dimensions of the index for cities and the inclusion of new, important dimensions for the situation of Brazilian municipalities: transportation, crime and education.

A pilot version of the project was carried out in São Paulo and in Rio de Janeiro, generating two reports (RUEDIGER et al, 2017b; 2017c), which were presented in events in each city, mobilizing the community and local authorities around the theme. The practical results of the ODI evaluation could be felt, for example, in statements by public authorities in which they revealed their interest in understanding how to improve the quality of their open databases.

This process led to an expansion into other Brazilian cities in 2017. To that end, a mapping was conducted of people and organizations interested in gathering data for Open Data Index in other municipalities. An initial set of ten cities was chosen by OKBr and by FGV DAPP to closely monitor the process of data collection and review, in addition to the cities of São Paulo and Rio de Janeiro, whose evaluation parameters would be updated. The criteria used were mainly the ability of the local teams to carry out the required tasks and the regional distribution.

In this process, it is natural that some cities do not go through to the end, which led the results presented in this report to contain a total universe of eight cities: Belo Horizonte-MG, Brasília-DF, Natal-RN, Porto Alegre-RS, Rio de Janeiro-RJ, Salvador-BA, São Paulo-SP and Uberlândia-MG.

Therefore, this study presents the results of the ODI evaluation for cities in 2017/2018. The following section will discuss aspects of the methodology used for this evaluation, and the section after that will discuss the consolidated results of the data collections carried out by the local groups and reviewed by the FGV DAPP and OKBr teams. This input may be useful in the elaboration of guidelines for open data policies inside the scope of each city. Each dimension of the index will be detailed in order to highlight the good practices found and the main points requiring attention.

2. Methodology

The Open Data Index (ODI) is an independent assessment of open government data publication from a civic perspective. The ODI allows different parties interested in open data to monitor the government's progress in releasing it, and enables governments to receive direct feedback from the users.

The detailed methodology can be found in Ruediger and Mazotte (2018). This section will be limited to some of the main aspects related mostly to the references for building the questionnaires, index scoring and interpretation of the results.

2.1. General aspects

As any other comparative evaluation tool, the ODI tries to answer a question. In our case, the question is this: *How do governments around the world disclose open data?* Other important questions arise from this one, such as:

- Which governments readily disclose open data? Which governments still need to improve their publication of open data?
- Which dataset is the most open? Which dataset is the least open?
- Which aspects of open data are easier or more difficult to implement?

By design, the ODI limits its assessment to the publication of government data. It does not assess other aspects of the common evaluation structure for open data, such as context, use, or impact. This focus allows the index to provide a standardized, robust and comparable open data evaluation for the Brazilian cities. The ODI observes only the characteristics of access and usability, without advancing, as of yet, towards the issue of data measurement quality.

We defined open data according to OKFN (2006), a set of principles which describes the openness of data and content. It is also simple and easy to operationalize. The only part of the methodology used in this study which is not completely in line with it is our assessment of "open, machine processable" formats. We attribute the maximum score to machine processable formats even if their source code is not open. In the Open Definition, the formats must be usable with at least one free, open source software. This means that the index favors the easy opening of data over the real opening of a format.

We understand that not all cities have the same political structure; therefore, it is possible that not all subnational governments produce the same data, since they are potentially subject to different laws and procedures. Therefore, the local ODI can evaluate the publication of city data that is not necessarily provided by the city's administration.

It is important to highlight that governments frequently disclose data in various websites and in many files and formats. If a given piece of data is doubly disclosed, we choose the one presenting the most complete database possible as the entity to evaluate, according to the requirements of what to evaluate in each dimension. In some cases, the reviewers cannot find a reference dataset because the data are divided into several files, formats and locations. In this case, they refer to the research for different files. It is important that the sum of these files contains all the data characteristics required.

2.2. The process

There is a standard process supporting the data collectors and reviewers in order to reduce biased evaluation and personal judgment. The process includes data collection by a local team followed by a review that is jointly carried out by FGV DAPP and OKBr. After that, the results from the review come to the knowledge of the data providers and eventually go through a second review and reach a final result.

The index obtains its data through mass collaboration. For the cities index, we formed local groups in the interested cities, which underwent a selection process based on the criteria of regionalization and the team's competence. The selection led to a sample of 12 cities. However, some of them were not able to complete the data collection along the process, leaving a final sample of eight municipalities.

Each team was composed of contributors interested in open government data who could assess the availability and quality of the open datasets in their respective locations. The teams learned about the data collection through publications on social networks, ongoing communications in the Open Government Data Forums (OKFN, 2018), and actively in the network, in conferences or events.

It is important to highlight the contextual nature of the ODI evaluation. The data was collected by the local groups between November and December 2017, and then reviewed in 2018. We took into account the state of open data in 2017 in order to keep consistency for comparisons between the cities. For that reason, databases which were created or released in 2018 were not considered for this evaluation.

To provide reliable and valid results, each data collection must be reviewed. To that end, OKBr, together with FGV DAPP, carried out an extensive review of the contents submitted, consulting with specialists in each area whenever necessary to ask questions and with specialists in open data through the discussion forum. The review included conversations with the local teams of data collectors, identifying potential inconsistencies. This process complied with the Open Definition guidelines and with the methodology of the index. The data collections which were finally accepted into the platform encompass the result of a systematic review and feedback from the local collectors.

We took the following steps to ensure review quality: First, we periodically visited the forums in order to incorporate and answer comments from the Open Knowledge community, including the data collection teams. Secondly, whenever possible, we carried out a comparison with the results from last year, checking whether the same origin URLs were used, whether there was any new open data initiative or whether something from last year was discontinued, identifying the reasons behind each change. This was done for São Paulo and Rio de Janeiro, which were part of the pilot. Finally, we carried out a verification together with the data collectors in order to refine some of the information presented. In this process, answers were changed and databases were resubmitted.

Once the results are released, the civil society and the government are invited to provide their opinion about the results and comment on how they believe we could improve the evaluation. This conversation stage stays open for a month after releasing the results, and the local authorities can submit change requests based on data available in the year of evaluation; in this case, 2017.

In general, policy makers and those responsible for the publication of open data make references to the results and publicly highlight their advancement in the index. In addition, the results indicate how each government can improve in their publication of open data.

2.3. Dimensions evaluated

The ODI measures the openness of certain categories of data previously defined over the years by the international Open Knowledge community. The ODI evaluates open government data that has been proven useful for the public. Other types of open data that do not fit these categories were not considered in our evaluation.

The categories of data reflect relevant information for the civil society in general. The categories were developed collaboratively over the years together with specialists on each theme, including organizations which support open data in their respective fields. In some bases, we based our definition on international standards of data production and reports used by governments around the world. Next, we will present the sources used for each dimension. The specific characteristics mapped in each dimension and why we look at them can be viewed in Ruediger and Mazotte (2018).

The index for Government Budget and Government Spending was based on the Open Spending (OKFN, 2019) initiative. For Procurement, we based the analysis on the Open Contracting Partnership (2016). As for Election Results, the index referred to the Open Election Initiative by the National Democratic Institute (NDI, 2015). The category of Company Register is based on the Open Corporates (2019) initiative. The evaluation criteria for Land Ownership were jointly developed with Cadasta Foundation (CADASTA, 2019). The index developed the City Maps category based on a reference report by the United Nations Committee of Experts on Global Geospatial Information Management (UNGGIM, 2015) and the Administrative Boundaries dimension on FAO Global Administrative Unit Layers (GAUL) project (FAO, 2015), as well as on the United Nations Geographic Information Working Group (UNGIWG, 2006). For Locations, we used the Universal Postal Union (UPU, 2009) work. The research for City Statistics was based on the Open Data Watch (ODW, 2013) initiative. The Draft Legislation and City Laws categories were based on the National Democratic Institute (NDI, 2019) and the Declaration of Parliamentary Openness initiatives (OPENING PARLIAMENT, 2012). The definitions and standards from the World Health Organization were the main source for Air Quality (WHO, 2005) and Water Quality (WHO, 2017).

The local index has three new datasets compared to the global one, all of them based on previous work from FGV DAPP. Ruediger et al. (2016c) was the main source for Public Schools, elaborated from public data on education. The Public Transportation dimension was based on Ruediger et al. (2016b), elaborated from public data on urban mobility and the Crime Statistics dataset was based on FGV DAPP's Security and Citizenship line of applied research (RUEDIGER et al, 2016a; 2018a; 2018b).

2.4. Evaluated attributes and score

Each dataset in each location is evaluated using a series of questions that assess the openness of the datasets based on the Open Definition (OKFN, 2006) and on the Open Data Charter (2015). Each research question measures a key legal, technical, or practical aspect of data openness. With that approach, we seek to reduce potential biased evaluations regarding single aspects of openness. The full questionnaire can be found in Ruediger and Mazotte (2018).

The structure of the questionnaire contains an initial section about the data collector, which is not scored; its goal is to develop an understanding about the profile of the local teams. After that, it contains questions about the evaluated datasets. In addition to answering multiple-choice questions, the data collector can make comments to explain aspects related to the process as well as clarify details, which is very useful when disclosing the results.

Table 1 below shows, on the left column, the original weight distribution of the global ODI, created by OKFN and used by Ruediger et al. (2017b, 2017c). The middle column shows the weights used this year, while the right column contains an explanation for the change. We must highlight that the sum of the scores in the old methodology was 90 and not 100. We understand that it would be more intuitive to use scores that summed up to 100; to reach this value, we assigned additional points to certain questions as described.

The changes were made after implementing the pilot, in which we detected some incompatibilities with the reality in Brazilian cities. For that reason, time comparison exercises must be carried out with caution. Each time this was done in this document, we recalculated the index from last year using the weights from this year, in order to have some notion of the evolution of the status of open data in the two cities which were part of the pilot.

An important note for this evaluation is that, in general, a lot of ambiguity is generated around the presentation of an open data license. While some datasets have explicit licenses (e.g. Creative Commons), others have textual declarations about the public domain or potential uses for the data; some of them do not have anything specific, but somehow indicate that the publication of the data complies with some legal system.

Table 1 - Weight changes applied to the 2017/2018 City ODI

Question	Original methodology	2017/2018 City ODI methodology	Reason
B2. Available with no register or access request	15	15	No change. We kept a lower score for considering that there are more limiting aspects in the open definition.
B4. Available free of charge	15	20	We assigned additional points to this category for understanding that a restriction in this aspect would be very limiting to data openness (more than in other criteria).
B5. Available for bulk download	15	20	We assigned additional points to this category due to the issue of usability, which is very limited if the download is not available in bulk.
B6. Available in an up-to-date form	15	15	No change. We kept a lower score for considering that there are more limiting aspects in the open definition.
B7. Available with an open license or declaration of public domain	20	10	The existence of certain legal provisions generates ambiguities about the need to disclose a license for public data in Brazil. For that reason, we understand that this aspect should receive a higher weight.
B8. Available in an open/machine processable format	10	20	The open format is a very important condition for defining openness. For that reason, we understand that this should be among the attributes with the highest score. A limitation here represents a strong restriction to the usability and applicability of data.

Source: elaborated by the authors.

Because Brazil has provisions such as the Law of Access to Information (Law 12.257/2011), the Fiscal Responsibility Law (Supplementary Law 101/2000), and others, many information providers exempt themselves from clarifying the potential uses of their data, since they understand that this is already done through the legal provisions. For that reason, the evaluation focused on cases in which there was a very clear use restriction (patent protection or declaration of data republication limitations) to give a negative answer to this question, even though this was not a very frequent issue.

2.5. Interpretation of the results

As explained in the sections above, the index evaluates specific data using specific research questions. The result is a final score that must be interpreted with caution. First of all, it refers exclusively to data with obligatory characteristics. If no dataset can be found online combining these characteristics, the data will not be considered available (leading to a score of 0%). In addition, the research question checks different aspects of data access and usability (see Table 2). This means that, behind very high scores, we often do not find open data, but data with controlled access or public data in formats that are not well structured or machine readable. Therefore, the score does not show a linear increase in openness. Instead, it highlights areas in which the government could improve its publication of open data.

Table 2 - Interpretation of the 2017/2018 City ODI results

Type	Description	Score
Open data	Data considered open can be freely used, modified and shared with any person, for any purpose. Main criteria: machine readable, open formats, full access, available in a complete form, at no cost and legally in a public domain.	100%
The data is public but not fully open.	The data is public but with some limitation that make it not fully open. The data is public if it can be seen online by the public, with no restrictions. They may be available only in a non-editable format (e.g. PDF) or not up-to-date, with some access limitation (registration, patent protection or similar limitation), only partially available, or cannot be downloaded.	Intermediate scores between 0 and 100%.
Lack of data	Lack of data means the governments do not produce data about a certain phenomenon, or they do produce it but do not disclose it. This shows that some governments still have a long way to go before they are ready to produce data.	0%

Source: elaborated by the authors.

For example, we may evaluate budget data in PDF format, which might be in public domain and available online for free, but in a format that is virtually unusable. This data is presented as 80% open. The score suggests a very high level of openness, but, in fact, the data is not open. We considered that the data is actually open if the dimension scores 100%. With this approach, assigning intermediate scores to databases that are not fully open, the goal of the index is to demonstrate which part of the data is already available, and how this can be improved. Therefore, it is important to carefully interpret how the data is disclosed.

3. Findings

In this section, we will present the results from the evaluation regarding the state of open data in 2017 for the eight cities evaluated. More details about the results from each location can be seen in Ruediger and Mazotte (2018). These results allow us to understand the current state of the open data policy and transparency in the cities evaluated, as well as diagnose priority measures seeking maximum transparency.

Next, we will present analyses under different aspects of evaluation of the results obtained during the data collection. After presenting the general results in section 3.1, we will present, in section 3.2, the results per dimension, analyzing which datasets require the most effort from the government authorities. Then, section 3.3 will clarify how we categorized the issues observed in the datasets, presenting an analysis of the main bottlenecks. Sections 3.4 and 3.4 present more detailed analyses per city and per dimension, respectively.

3.1. Overview

The Open Data Index score can be viewed in the score column, which evaluates the compliance of the data provided by the government with the transparency criteria used in many countries in the world, and in the "%Open" column, which calculates the percentage of evaluated datasets which comply with all the criteria in the methodology. Table 3 shows the results from the two metrics for the set of cities in the study.

The general average among the cities was 65%, with individual results ranging from 43% and 84%. Out of the 136 databases evaluated (17 per city, 9 cities), 25% (or 34) were fully open.

Of the eight cities evaluated, São Paulo showed the highest score and the highest %Open. This means that it was the most successful city both in disclosing 100% open databases and in getting their public databases closer to meeting the criteria of the Open Definition. On the other hand, that does not mean that all challenges were overcome, as less than half of the databases evaluated fully met the ODI criteria. Eight of the 17 datasets evaluated for São Paulo meet all the criteria of the methodology, but do not necessarily have all the characteristics[5].

---

[5] In crime statistics, there is not detailed information about gun seizure. In land ownership, the property boundaries are not available.

Table 3 - City scores in the ODI according to score and %Open

Cities	Score	Cities	%Open
São Paulo	84%	São Paulo	47%
Rio de Janeiro	75%	Belo Horizonte	35%
Belo Horizonte	73%	Rio de Janeiro	29%
Porto Alegre	68%	Brasília	29%
Brasília	68%	Porto Alegre	23%
Salvador	55%	Uberlândia	17%
Uberlândia	53%	Natal	11%
Natal	43%	Salvador	5%

Source: elaborated by the authors.

For São Paulo and Rio de Janeiro, it is possible to compare these results with the evaluation from last year[6]. In São Paulo, there was a small qualitative improvement compared to the previous evaluation: more characteristics evaluated by the index were made available through the portal GeoSampa (e.g. coordinates for schools), which is a great transparency and clarity practice in the availability of information for the common citizen who is not a professional in that type of data or a researcher.

However, regarding the general transparency parameters, the difference between the pilot and this evaluation was not very significant: the city obtained a positive difference of 2 percentage points in the score[7]; considering we have one less dimension this year (Weather Forecast), that is a very similar result. The databases with maximum openness score were a different case. Compared to last year, the Crime Statistics dimension now met all methodological criteria, and Draft Legislation lost some points due to the complete database not being fully downloadable.

The situation is similar in Rio. There was an increase of 1% in the score, which can be explained by the fact that the Locations dimension obtained the maximum score in 2017 and 80% in 2016, while the City Maps dimension decreased from 100% to 85%[8]. In Rio de Janeiro, like São Paulo, the amount of databases which obtained full openness in the two years was the same (that is, 5 out of 17, or 29%).

---

[6] We replicated the methodological changes implemented this year in the results from the 2016 data collection to allow for comparability. This was done only for São Paulo and Rio de Janeiro, which were evaluated in the pilot project carried out in 2017.

[7] The score was 82% in 2016 and 84% in 2017. The %Open did not change in practice, since, considering 17 dimensions (removing weather forecast from last year), São Paulo still has eight 100% open dimensions, or 47% in the indicator.

[8] The city maps used were not updated in the period between the two evaluations, which resulted in a decrease in that dimension's score.

A high score in the %Open criterion demonstrates the government's intention to provide quality open data to the population, meeting the needs of researchers and developers who wish to work with these databases as well as the needs of common citizens who want information about their government, public policies, laws, the location where they live, etc.

We can obtain more interesting insights comparing each city's score to its %Open instead of simply comparing the performance of the cities. For instance, we can see that the rankings are not symmetrical. The city of Rio de Janeiro, for example, has the second highest score, which means that each particular dataset has specific challenges in order to reach full openness, but its %Open is lower than that of, for example, Belo Horizonte, which presents a slightly higher number of 100% open databases.

There is no judgment on whether a certain metric is better or worse than the next, but the comparison between results suggests different ways to increase data openness: while Rio de Janeiro can generally focus its measures on identifying particular issues in its databases which prevent them from being 100% open, Belo Horizon could try to look at its 100% open databases and find ways to replicate them for the others.

The city of Salvador was an interesting case in terms of database openness, as it obtained a score of 55% and ranked sixth, but had a %Open of 5% (only one of the databases fully meets the criteria), ranking eighth in that measure. This type of result shows that there was an effort to make information relevant to the population public and accessible online, yet it complied very little with the Open Definition criteria.

For the city of Natal, which ranked last in the score, we found public data for only ten of the 17 datasets evaluated. Out of those, only two obtained the maximum openness evaluation. None of the databases presented by the Natal government met all the transparency criteria. These results show that a greater effort by the government entities is needed for providing more relevant information for the public life of its citizens in order to comply with the good practices of transparency and access to information adopted in other cities in Brazil and the world.

The results indicate that there is an intention of the government mechanisms to bring relevant information to the citizens regarding their public life, with the creation of transparency portals that comply with the Law of Access to Information, making data public. However, there is still a lot to improve in terms of these public data becoming actually open, and challenges regarding the quality of data, the formats available and the stability of the government websites. During the research process, many websites presented failures or were unavailable for a few hours or even days.

3.2. Results per dimension

In this subsection, we will present the results consolidated per dimension in order to identify the priority measures for the municipalities evaluated and the main challenges for advancing in terms of open data policies.

As shown in Table 4, the four datasets with the most bottlenecks were: Land Ownership, Company Register, Air Quality and Water Quality. The data on Land Ownership, which are regarded as a bottleneck internationally, were only disclosed by São Paulo and Brasília. The data on Air Quality and Company Register were only available for half the cities evaluated. The data on water monitoring, although available in all cities, had many usability issues but not many difficulties for obtaining information. That is, the information was easy to obtain but of low quality.

In the %Open evaluation, we noticed that five dimensions did not have any 100% open database: Company Register, Water Quality, Public Transportation, Procurement and Draft Legislation. Comparing these two pieces of information, we noticed that the Company Register and Land Ownership dimensions are in a particularly critical state in the cities.

During the research, we observed that some of the city data are available in Brazil in a consolidated form in the federal level, such as data on Electoral Results (which explains the fact that the entire universe of cities received maximum score). In city statistics, that is also often the case. Regarding this evaluation, we observed that this data is mostly compliant with international transparency parameters, except for the GDP information made available by the IBGE, which are not up-to-date when compared to the methodology requirements.

The results of the study highlight the relevance of structuring databases in the shape of an online map that shows information related to a city's geographic and urban structure. Many datasets were often extracted from centralized platforms (congregating different datasets integrated into a single portal) such as administrative boundaries, city maps and locations. Initiatives such as these facilitate the access to information by citizens and are considered good transparency practices.

Table 4 - Average evaluation of the ODI dimensions

Datasets	Score	Datasets	%Open
Election Results	100%	Election Results	100%
Public Schools	99%	Public Schools	87%
City Statistics	87%	Crime Statistics	75%
Crime Statistics	85%	Government Budget	37%
Government Spending	83%	Government Spending	25%
Government Budget	81%	Administrative Boundaries	25%
Administrative Boundaries	76%	City Statistics	12%
City Laws	73%	City Laws	12%
Draft Legislation	70%	City Maps	12%
City Maps	68%	Locations	12%
Procurement	67%	Air Quality	12%
Public Transportation	57%	Land Ownership	12%
Locations	53%	Draft Legislation	0%
Water Quality	53%	Procurement	0%
Air Quality	35%	Public Transportation	0%
Company Register	30%	Water Quality	0%
Land Ownership	21%	Company Register	0%
Median	**70%**	*Average*	25%

Source: elaborated by the authors.

The public transportation datasets had a recurrent flaw in that none of the evaluated cities had adequate information about connections between different types of transportation. Another frequent issue was the lack of up-to-date data. Private sources such as Google release information of better quality and clarity on this topic, and it requires the attention of the municipalities evaluated.

Regarding data on draft legislation, we found issues in all cities. Information about the votes during council sections are hardly ever available together with the bills, and none of the cities associated the transcription of the debates with a law vote. Both types of information must be manually obtained by searching for the laws sorted by the data in which they were debated in a council section, or bill number separately from the laws. This practice makes it harder to monitor the legislative debates, so none of the municipalities obtained the maximum score in the evaluation of this dataset.

3.3. Bottleneck analysis

Here, we counted the issues found in each dimension and categorized them into two types: usability and process. In the usability category are listed issues related to difficulties using the data and the ability to transform this data into information, while the process category contains difficulties accessing the data and omission/lack of a license or declaration that the data are located in a public domain.

Table 5 - Categorization of the issues into usability and process

Usability	Process
Incomplete dataset	Restricted access
Outdated	Difficulty locating data
Open format unavailable	Bulk download is unavailable
Difficulty working with the data	Non-transparent license

Source: elaborated by the authors.

Each of these bottlenecks is related to some stage of filling in the surveys or collecting the information (as described in the methodology stage). In the usability category, we have:

1. *Incomplete dataset* refers to the lack of some of the characteristics required for the evaluation, which does not necessarily invalidates the dimension as a whole[9].

2. A specific update frequency is required for each dataset evaluated. The questionnaire item related to this type of issue assumes update intervals ranging from one day (such as air quality) to one year (such as city maps). When the dataset is not updated according to this parameter, we consider it an *outdated dataset* issue.

---

[9] As a convention, we adopted the idea that the dataset was not available (it received a 0% score), which means it has all the possible issues (8).

3. A list of possible open formats is made available, where we choose the ones found. If none are selected, we consider it an *open format unavailable* issue.

4. *Difficulty working the data* is not directly related with a questionnaire item, but with characteristics that make it difficult to transform the raw data into information. When a lot of work is necessary in order to gather the data, there is difficulty viewing the information, or there are issues with the metadata (insufficient documentation, unclear coding, etc.), we consider it a bottleneck in that category.

In the process category, we have:

5. Some government data are available only after creating an account or providing a tax number (CPF), e-mail, etc. When it is necessary to register or request access to the data, we consider it a *restricted access* issue.

6. The data collector answers a question based on his or her experience finding the data, assigning a number from 1 to 4, with 1 being the most difficult. We understand that there is an issue categorized as *difficulty locating data* when the answer is 1 or 2.

7. The questionnaire asks whether it is possible to download bulk data; when the answer to this is negative, this is an issue categorized as *bulk download is unavailable*.

8. If the data is not clearly licensed through an open license or declared as part of the public domain, we understand that there is a bottleneck categorized as *non-transparent license*.

We detected 429 issues based on this methodology for identifying bottlenecks. Out of those, 267 (62%) are usability issues and 162 (38%) are process issues, as shown in Table 6. This means that usability issues are more frequent, which corroborates the conclusion from the pilot evaluation (RUEDIGER et al., 2017a, 2017b, 2017c) that the dimension in which we advanced the most in Brazil in terms of transparency is the publication of data, but not necessarily its understanding and its transformation into something useful.

**Table 6 - Frequency of bottlenecks found per type
(usability and process)**

Usability	No. of Bottlenecks and % of the total	Process	No. of Bottlenecks and % of the total
Difficulty working with the data	72 (16.8%)	Bulk download is unavailable	70 (16.3%)
Incomplete dataset	67 (15.6%)	Non-transparent license	34 (7.9%)
Open format unavailable	65 (15.2%)	Difficulty locating data	32 (7.4%)
Outdated	63 (14.7%)	Restricted access	26 (6.1%)
Total	**267 (62.2%)**	**Total**	**162 (37.8%)**

Source: elaborated by the authors.

The most frequent issue was difficulty working with the data. This issue was also found to be the most frequent during the 2016 federal evaluation (RUEDIGER et al, 2017a), and it happens because of the following factors: difficulty using/viewing the information, problems with the metadata (insufficient documentation or unclear coding), bulk download unavailable for a complete database or large datasets, or open format unavailability. These issues lead to difficulties dealing with the data and creates obstacles, for instance, for statistical and scientific analyses from the data.

The second most frequent issue happens when the bulk download of a database is unavailable, which was also detected as the second most frequent bottleneck in Brazil during the Global Open Data Index 2016 (RUEDIGER et al, 2017a). This indicates that the bodies responsible for providing information still do not understand how to facilitate the process of editing and dealing with the data, which is important for promoting actual transparency and also an essential part of complying with the international transparency criteria.

Another frequent issue is the lack of characteristics regarding what the index looks for in each dimension, which is categorized as incomplete dataset. Take, for instance, the amount of arsenic in water quality reports, a characteristic which is measured in only one of the cities evaluated. The datasets on water quality had, on average, only 4 of the 7 characteristics looked for by the index, except for the Rio de Janeiro dataset, which encompasses all the characteristics in this dimension.

For half of the cities evaluated, the most frequent issue is open format unavailability, which is one of the main limiting factors for reaching full data openness.

3.4. Detailed analysis per city

In Table 7, we present a matrix of the main bottlenecks in each city for the datasets evaluated, enabling a clearer analysis about where each municipality must focus its efforts.

Table 7 - Frequency of dimensions with bottlenecks per type of issue and city

Bottlenecks / Cities	SP	RJ	BH	BSB	POA	UDI	SSA	NAT	Total
Difficulty working with the data	7	9	6	5	7	12	12	14	**72**
Bulk download is unavailable	5	7	9	9	7	11	11	11	**70**
Incomplete dataset	2	5	7	10	10	12	10	11	**67**
Outdated	4	5	6	9	10	8	10	11	**63**
Open format unavailable	2	8	4	7	6	13	12	13	**65**
Non-transparent license	2	1	3	2	6	6	7	7	**34**
Difficulty locating data	2	4	4	2	4	4	5	7	**32**
Restricted access	3	1	3	2	2	4	3	8	**26**
Total Bottlenecks	**27**	**40**	**42**	**46**	**52**	**70**	**70**	**82**	**429**

Source: elaborated by the authors.

The city of São Paulo, in general, performed well according to the desired criteria of openness and transparency. The municipality showed good practices in multiple data sets through its geoportal, which contains several segments of information evaluated. The platform is easy to use and allows bulk downloading of databases in an open format. The most frequent issues observed were the lack of a complete database: six of the 17 datasets failed in this regard, in addition to presenting outdated data. The most problematic dimensions were Air Quality and Company Register, both with four bottlenecks each.

In Rio de Janeiro, the city's geographical data portal provides several datasets for the dimensions evaluated which are easy to view and use, and good practices for disclosing information, concentrating several groups of data on the same platform. Another highlight in the city was the data on Crime Statistics; it was the only municipality with no issues in that dimension. Most of the problems observed were related to open format unavailability, a flaw observed in almost half of the databases evaluated. It was also found that in 41% of cases it was not possible to download a complete database. The most problematic dimensions were Air Quality, Company

Register and Public Transportation, all with four bottlenecks each. No database was found that qualifies for the desired criteria in the Land Ownership dimension.

Belo Horizonte, the capital of Minas Gerais was the only one to present up-to-date data for all characteristics in City Statistics and to obtain the maximum score in Air Quality. A good practice adopted by the city government is the concentration of several types of relevant data about public life in its portal, facilitating the access to information. However, more than half of the datasets in the municipality presented issues regarding the download of a complete database. It was also detected that seven of these datasets were incomplete, not presenting all the desired characteristics. The most problematic dimension was Water Quality, in which we detected six out of eight possible issues. No databases have been found that qualify for the desired criteria for the dimensions of Land Ownership and Company Register.

In Porto Alegre, we observed that it was easy to obtain cartographic data in an open format for various characteristics of the city through the *observaPOA* website. In addition, the city's transparency portal demonstrates the government's intention to show its citizens how public money is being used in a simple and easy way, meeting all the desired requirements for data openness and transparency. The most problematic dimension was Procurement, which showed six of the eight possible issues categorized. No databases were found that qualify for the desired criteria for the Land Ownership and Air Quality dimensions.

Brasília presented a medium performance, ranking fourth together with Porto Alegre. The Federal District's transparency portal provides several datasets that are looked for by the index, meeting criteria such as open format and download of the complete database, a good transparency practice. Unavailable download of the complete database, outdated dataset and incomplete dataset were the three most frequent issues, present in 53% of the datasets evaluated. The dimension with the highest number of bottlenecks was Water Quality, in which we detected five bottlenecks. No databases were found that qualify for the desired criteria for the Company Register and Locations dimensions.

In Uberlândia, the only database that obtained the maximum score, with the exception of the databases disclosed by federal entities, was Crime Statistics, which is made available by a state body. The Uberlândia government has made efforts to become transparent in terms of government spending, disclosing this database in a way that is easy to view and find. A higher compliance with the current transparency criteria is recommended, making data available in an open format and with the ability to download the complete database. The most frequent issues found were open format unavailability and incomplete database, which appeared in 76% and 70% of the datasets evaluated, respectively. The most problematic dimension was Water Quality, which had six bottlenecks. No databases were found that qualify for the desired criteria for the Locations, Air Quality, Company Register and Land Ownership dimensions.

A good practice adopted by the city of Salvador is the implementation of a transparency portal containing several types of useful information for citizens, which shows the government's intention to be transparent with the administration of public funds. On the other hand, with the exception of the Electoral Results dimension, which is disclosed by the TSE, Salvador did not obtain the maximum score in any of the datasets. The most frequent issues were open format unavailability, which happened in 70% of the datasets, and download of the complete database unavailable, in 64%. The most problematic dimension was Public Transportation, which showed seven of the eight bottlenecks categorized. City Maps, Company Register and Procurement showed six bottlenecks, which also makes them priority areas for data openness measures. No databases were found that qualify for the desired criteria for the Air Quality and Land Ownership dimensions.

Natal has good data openness practices with the implementation of the government's transparency portal, which enables easy and accessible viewing of how the public funds have been spent. Another positive aspect is the procurement portal, where it is possible to obtain information about ongoing and past tenders. However, the city obtained a score of 43%, the lowest among the cities evaluated. None of the municipal or state databases obtained a score of 100%. We found databases qualifying for the data collection for only ten of the 17 dimensions evaluated. The most frequent issues were open format unavailability, outdated datasets, incomplete datasets, and download of the complete database unavailable. All the most frequent problems were detected in 60% or more of the databases evaluated. The dimensions with the most bottlenecks were Water Quality, City Laws and Administrative Boundaries. No databases were found that qualify for the desired criteria for the Crime Statistics, Draft Legislation, Public Transportation, Locations, Company Register, Air Quality and Land Ownership dimensions.

3.5. Detailed analysis per dimension

Table 8 below inverts the logic adopted so far. We present the frequency of cities per dimension in the index and categories of issues diagnosed in order to assess which issues are the most relevant in each dimension[10]. We can observe that the least problematic datasets are Electoral Results, Public Schools and City Statistics. Except for a few specific cases[11], the sources of data for these datasets in all cities evaluated were federal entities, which had good practices of transparency and data openness.

---

[10] A more detailed analysis of this issue per bottleneck can be seen in Ruediger and Mazotte (2018).
[11] Belo Horizonte in Electoral Results, Porto Alegre in City Statistics and Salvador and Brasília in Public Schools.

Table 8 - Frequency of cities with bottlenecks per category and dimension

	Usability	Process	Total issues found
Election Results	0	0	**0**
City Statistics	7	0	**7**
Public Schools	9	1	**10**
Government Spending	5	7	**12**
Government Budget	11	3	**14**
Crime Statistics	12	5	**17**
Administrative Boundaries	13	5	**18**
City Laws	13	8	**21**
City Maps	17	7	**23**
Draft Legislation	15	12	**27**
Procurement	20	9	**29**
Locations	18	17	**36**
Public Transportation	21	15	**36**
Water Quality	30	6	**36**
Air Quality	22	20	**42**
Company Register	27	22	**49**
Land Ownership	27	25	**52**
Total	**267**	**162**	**429**

Source: elaborated by the authors.

Regarding City Statistics, the most common problem was the lack of updates, as the data provided by the IBGE are out of date for municipal GDP.

Public Schools had frequently incomplete databases, particularly because the INEP does not disclose school addresses and coordinates. In many cases, this information was also not available from the municipal or state bodies responsible for the school data.

The databases on Government Spending do not provide the download of a complete database in 62.5% of cases.

Open format unavailability is a frequent problem in Government Budget datasets, often made available in non-machine-readable formats such as PDF.

Crime statistics had incomplete datasets for almost all cities. In seven of the eight evaluations, no detailed data was found on gun seizure.

The Administrative Boundaries and City Maps databases were often made available on the same platform, presenting identical issues in most cases. In six of the eight cities, the datasets evaluated had update issues: their last updates had happened more than a year prior to when the data was collected. The main distinction between the two dimensions is due to the fact that several cities do not provide some of the characteristics looked for by the index for the city maps.

Most of the City Laws databases were evaluated as simple to find and view. However, when trying to obtain a large amount of laws simultaneously or data in open format, this dimension presents some issues. All cities, except for São Paulo, had open format issues or lack of a complete database available, and in most cases both issues were detected.

Regarding Draft Legislation, all databases evaluated were diagnosed with issues regarding the download of a complete database, since none of the sources provides the votes or transcription of the debates together with the bills. Therefore, it is necessary to obtain that information separately by searching by the date when the project was discussed.

Regarding Procurement, we noticed that, in almost all cases (seven of eight cities), it was not possible to download a complete database, and in six of the eight municipalities evaluated, an open format was not available.

In all cities, except for Rio de Janeiro, the datasets referring to the Locations dimension received scores of less than 100%. We also observed that, although all municipalities have this database, only five of them make it available, of which only two present information regarding the postal codes of the locations[12].

None of the Public Transportation datasets met the criteria for downloading a complete database. For information regarding the connections between different types of transportation, the only city who did not fail these criteria (Uberlândia-MG) did not have this possibility, as it has only one type of transportation available.

The Water Quality dataset presented the most frequent usability issues, even though all cities presented public information about this topic. The eight datasets had updating issues – updates should be at least weekly – and did not provide an open format. We

---

[12] There is a controversy about the commitment to disclosing this data by the company Correios (for the entire national territory). For more information, see the forum (OKFN, 2018) at the link https://discuss.okfn.org/t/entry-for-locations-brazil/4533. There is also a paper analyzing the appropriation of the postal codes by the company Correios (RIBEIRO e OLIVEIRA, 2017).

also observed that none of the cities evaluated, except for Rio de Janeiro, presented all the water evaluation parameters desired by the ODI methodology.

Regarding the databases on Air Quality, out of the eight cities evaluated, only four provided data on this topic. The city of Belo Horizonte was the only one to obtain a score of 100%, while the other cities evaluated had important flaws in their data presentation. In Brasília, the data was very outdated compared to what the ODI looks for: the last available information was from four months before the data collection, while daily updates are expected. In Porto Alegre, according to the air quality bulletin, the city's stations were undergoing maintenance and there was no sufficient data to conduct the survey. For the city of Salvador, the government body INEMA (Institute of the Environment and Water Resources) reports that the data is collected by the CETREL (Center for Liquid Effluent Treatment) in technical cooperation with the government of the state of Bahia, but there was no information available about the capital city.

Company Register data is also collected by all municipalities evaluated, but the data is public is only four of them. In most of the cities evaluated, there was the possibility to buy this database from the respective state trading boards, which is a great obstacle to the transparency of these datasets. In Porto Alegre, the disclosure of this information is done through the Municipal Department of Industry and Commerce, which provides a complete database in an open format, demonstrating its intention of being transparent with the records of companies in the city. In Rio de Janeiro and Salvador, it is only possible to obtain the data through prior knowledge of the company's tax identification number (CNPJ) or company name, which limits access to the complete data. More specifically in the case of Rio, it is possible to obtain a list of companies registered in the Rio de Janeiro invoice system, which is not the case for all companies registered in the municipality, since that registration is only necessary for issuing electronic invoices.

Land Ownership records are a type of data which presents issues everywhere in the world (OKFN, 2016). It is not different in Brazil, even though the data is collected – the city administration is responsible for the collection of taxes on urban property (IPTU), for example. However, nothing has changed regarding the transparency of this information, at least in the comparable cases (Rio de Janeiro and São Paulo). Regarding the other six cities, only one of them (Brasília) presented a qualifiable database. In São Paulo, which was a highlight since its database was 100% compliant, the data disclosed did not include the property boundaries, only identification, property type and value.

4. Final Remarks

The Brazilian Open Data Index for Cities evaluated the state of open data, based on data from 2017, in 136 databases: 17 dimensions were evaluated in eight participating

cities, a considerable increase compared to 2016, when 36 databases were evaluated in two participating cities. This alone reveals a positive fact: the number of people interested in the subject of transparency and open data in Brazilian cities is growing.

The process of data collection, review and consolidation is a pioneer in the country and included the participation of local groups in the eight cities evaluated: Belo Horizonte-MG, Brasília-DF, Natal-RN, Porto Alegre-RS, Rio de Janeiro-RJ, Salvador-BA, São Paulo-SP and Uberlândia-MG.

The purpose of this paper is to outline ways for participating cities to improve their open data policies. The main method for analyzing the results was the identification of bottlenecks by dimension and by city.

Regarding the year of 2017 in the case of the two cities in which the evaluation was made in two moments, there was no visible evolution. The number of 100% open databases is the same and the score improvement was small. This means that there is still the same room for improvement as last year for these cities.

Among the evaluated dimensions, four stood out because they presented the highest amounts of bottlenecks: Land Ownership, Company Register, Air Quality and Water Quality. Of all the bottleneck types detected, 62% are usability issues and 38% are process issues, corroborating the conclusion from last year's ODI that advances are more visible in the disclosure of data than in facilitating understanding and usability. There is also a latent need for more databases that can be adequately used and transformed into information.

The average score for the dimensions in the cities evaluated was 65%, similar to the national average in the 2016 Global Open Data Index[13]. This result shows that there is still a lot of room for improvement in all municipalities evaluated. Only 25% of the databases analyzed in the universe of the eight cities obtained the maximum score, with two dimensions presenting data which is made available to all cities through federal agencies.

Five of the 17 dimensions evaluated did not obtain the maximum score in any of the evaluated cities (Company Register, Public Transportation, Water Quality, Procurement and Draft Legislation). For three of the datasets evaluated – Air Quality, Company Register and Land Ownership – there was no public information for half or more of the cities assessed, showing that many municipalities still lack data availability for their population.

Many improvements can be made in the way data is made available from by city governments or bodies, and different dimensions require distinct forms of

---

[13] Brazil obtained a score of 68% in the evaluation based on OKFN (2016).

improvement. This paper outlines which datasets need priority attention and where the most common issues are found, with guidelines for improving the open data policies in the cities.

Replicable good practices should focus their efforts on centralizing information into a single portal that is easy to understand, prioritizing user experience. We can observe that the efforts made by the cities in this direction are, in general, incomplete. For example, the city of São Paulo is a good example of geographic information consolidation (GeoSampa), but the portal does not present other types of data. Efforts for providing APIs, for example, which are common in many of the cities evaluated, can also improve to include greater range of dimensions and use good documentation practices to help in the process of using the data.

With a more powerful assessment than last year, we came to a similar conclusion, which is also present in the result of the federal evaluation: Brazil is an important actor in the issue of transparency and this is reflected to some extent in its cities, but there is still a focus on the idea that disclosing information is enough in terms of the commitment to transparency. However, more needs to be done: open, user-friendly databases with appropriate metadata, clearly intended for the public domain, which are easy to use and understand for a growing population. With this, the transparency and open data agenda can effectively contribute to the advancement of democracy.

References

CADASTA (2019). *Overview of Property Rights Data.* Available at: https://cadasta.org/open-data/overview-of-property-rights-data/. Accessed on May 30th, 2019.

FAO - FOOD AND AGRICULTURE ORGANIZATION (2015). *Global Administrative Unit Layers.* Available at: http://www.fao.org/geonetwork/srv/en/metadata.show?id=12691. Accessed on May 30th, 2019.

NDI - NATIONAL DEMOCRATIC INSTITUTE (2019). *About NDI.* Available at: https://www.ndi.org/about-ndi . Accessed on May 30th, 2019.

_____ (2015). *Open Election Data Initiative:* Election Data Guide (2015). Available at: https://openelectiondata.net/en/guide/. Accessed on May 30th, 2019.

ODW - OPEN DATA WATCH (2013). *Open Data Watch: About.* Available at: https://opendatawatch.com/about/. Accessed on May 30th, 2019.

OKFN - OPEN KNOWLEDGE FOUNDATION (2019). *Open Spending.* Available at: https://openspending.org/. Accessed on May 31th, 2019.

_____ (2018). *Open Knowledge Forums.* Available at: https://discuss.okfn.org/ . Accessed on May 30th, 2019.

_____ (2016). *Global Open Data Index*: Survey. Available at: http://global.survey.okfn.org/. Accessed on May 29th, 2019.

_____ (2006). *Open Definition.* Available at: https://okfn.org/. Accessed on May 29th, 2019.

OPEN CONTRACTING PARTNERSHIP (2016). *Open Contracting Data Standard: Documentation.* Available at: http://standard.open-contracting.org/latest/en/. Accessed on May 30th, 2019.

OPEN CORPORATES (2010). *About OpenCorporates*. Available at: https://opencorporates.com/info/about . Accessed on May 30th, 2019.

OPEN DATA CHARTER (2015). *Open Data Principles.* Available at: http://opendatacharter.net/principles/. Accessed on May 31th, 2019.

OPENING PARLIAMENT (2012). *Declaration of Parliamentary Openness initiatives.* Available at: https://openingparliament.org/declaration/. Accessed on May 30th, 2019.

RIBEIRO, A. M. A.; OLIVEIRA, E. B. (2017). Código de endereçamento Postal (CEP), um obstáculo aos Dados abertos no Brasil. *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde*, v.11, nov.

RUEDIGER, M. A. et al. (2018a). *Os deslocamentos do crime:* relações entre espaço e práticas ilegais. https://bibliotecadigital.fgv.br/dspace/handle/10438/20258. Accessed on May 30th, 2019.

_____ (2018b). *Violência contra mulher:* um outro olhar sobre a violência a mulher: o que dizem as denúncias no Estado do Rio de Janeiro. Available at: https://bibliotecadigital.fgv.br/dspace/handle/10438/24196. Accessed on May 30th, 2019.

_____ (2017a). *Índice de dados abertos para o Brasil*. Available at: http://bibliotecadigital.fgv.br/dspace/handle/10438/18299. Accessed on May 29th, 2019.

_____ (2017b). *Índice de dados abertos para cidades:* São Paulo. Available at: http://bibliotecadigital.fgv.br/dspace/handle/10438/18295. Accessed on May 29th, 2019.

_____ (2017c). *Índice de dados abertos para cidades:* Rio de Janeiro. Available at: http://bibliotecadigital.fgv.br/dspace/handle/10438/18296. Accessed on May 29th, 2019.

_____ (2016a). *Denúncia, crime e castigo:* o ciclo da violência na cidade do Rio de Janeiro. Available at: https://bibliotecadigital.fgv.br/dspace/handle/10438/18914. Accessed on May 30th, 2019.

_____ (2016b). *Desigualdades sobre trilhos e rodas: um estudo sobre acesso à cidade.* Available at: https://bibliotecadigital.fgv.br/dspace/handle/10438/18237 . Accessed on May 30th, 2019.

_____ (2016c). *Universo escolar:* o estado da educação no Brasil. Available at: https://bibliotecadigital.fgv.br/dspace/handle/10438/18238. Accessed on May 30th, 2019.

RUEDIGER, M. A.; MAZZOTTE, N. (Ed.). *Índice de dados abertos para cidades*. Rio de Janeiro: FGV DAPP, 2018.

UNGGIM - UNITED NATIONS COMMITTEE OF EXPERTS ON GLOBAL GEOSPATIAL INFORMATION MANAGEMENT (2015). *The Status of Topographic Mapping in the World.* Available at: https://www.isprs.org/documents/reports/The_Status_of_Topographic_Mapping_in_the_World.pdf. Accessed on May 30th, 2019.

UNGIWG - UNITED NATIONS GEOGRAPHIC INFORMATION WORKING GROUP (2006). *Core Geo-Database.* Available at: http://ungiwg.org/coreDB. Accessed on May 30th, 2019.

UPU - UNIVERSAL POSTAL UNION (2009). *Addressing and Postcode Manual.* Available at: http://www.upu.int/fileadmin/documentsFiles/activities/addressingAssistance/manualAddressingAddressingAndPostcodeManualEn.pdf. Accessed on May 30th, 2019.

WHO - WORLD HEALTH ORGANIZATION (2005). *Air quality guidelines:* global update 2005. Available at: https://www.who.int/phe/health_topics/outdoorair/outdoorair_aqg/en/. Accessed on May 30th, 2019.

_____ (2017). *Guidelines for drinking-water quality:* 4th edition, incorporating the 1st addendum. Available at: https://www.who.int/water_sanitation_health/publications/drinking-water-quality-guidelines-4-including-1st-addendum/en/. Accessed on May 30th, 2019.