

A Proposal for Viewing Government Open Data Information, Based on User Profiles

César Alencar Assumpção(✉) and José Remo Ferreira Brega

Computer Science, Paulista State University "Júlio de Mesquita Filho" – UNESP, Bauru, SP 17033-360, Brazil
{cesar.assumpcao,remo.brega}@unesp.br

Abstract. There is a global movement for the democratization of access to open government data, for the free use of society. This movement should enable greater participation by society in the development of a more efficient State, with the provision of public utility services, making wide use of technologies. To make this understanding clearer, a Systematic Literature Review (SLR) was carried out, which sought to reveal strategies for the use of open data, tools and techniques of Information Visualization. In Brazil, the municipalities are inspected by the State Audit Courts, and the São Paulo State Court of Auditors provides an open database of more than 600 municipalities. From this open database, complemented by other data sources and performance indicators, based on SLR results, the work presented in this article shows that with the correct application of Information Visualization techniques, it facilitates the understanding of public management. The public service is an important source of data, as it contains information that can have a direct impact on each citizen's life and compromise essential services such as health, safety and education. What is intended again is to structure specialized presentations according to the user's profile. The proposal is presented in the form of specific dashboards according to the user's profile, which proved to be essential to add value to the information and improve the understanding of the interested parties.

Keywords: Visualization · Information visualization techniques · Open data · OGD - Open Government Data

1 Introduction

Public Transparency in Brazil is provided by law, which guarantees citizens access to public data generated and maintained by any sphere of government, including the legislative and judicial powers.

Law 12,527 of 2011 [1] is one of the legal provisions that regulate access to information provided for in the Constitution [2], it is known as the Law of Access to Information (LAI).

Article 5 of the LAI states that it is the duty of the State to guarantee the right of access to information in an objective, agile manner and using clear and accessible language.

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12952, pp. 163–182, 2021.
https://doi.org/10.1007/978-3-030-86973-1_12

With this, LAI and the Public Transparency Law complement each other as important instruments to fight corruption and control State activities, making it easier to monitor the acts of the Legislative and the Executive.

Brazil is a member and co-founder of the Open Government Partneship (OGP), created in 2011 and which includes 60 countries. This initiative is proof of the interest that open government is increasing worldwide [3]. It is an open, permanent and bidirectional form of communication between the administration and its managers, citizens, companies and civil society. The three pillars that define it are participation, collaboration and responsibility (publication of information).

The Brazilian Government has a data portal open since 2012, where it centralizes the search and access to data and public information. The creation of the portal is part of the actions determined in the First Open Government Action Plan, launched in 2011.

The publication of open data provides us with some benefits, such as increasing transparency and reducing corruption by political agents, empowering citizens who can better interpret information and encouraging innovation, promoting economic development. The development of applications based on open data benefits citizens in the form of providing services, accessing public services and assessing problems of public interest [4].

It also contributes to the measurement of the impact of public policies and the creation of new knowledge, based on the combination of data and information sources.

David Weinberger, one of the co-authors of "The Cluetrain Manifesto" [5], notes that "there is an inverse relationship between control and trust". Governments may want to negotiate a loss of control, offering more transparency and openness, hoping to receive in return an increase in public confidence.

The São Paulo State Court of Auditors [6] receives detailed data from 645 municipalities on a monthly basis, the vast majority of which are small municipalities. The data collected from the municipalities in CSV format are available on its portal, starting in 2014.

In Brazil, public transparency refers to fighting corruption, but the interest in government data must go further, where we can identify some classes of stakeholders with different views on the same mass of data.

In addition to the citizen, who is primarily interested in government actions, social organizations may be interested in government resources for social assistance, while businesspeople in the region want to know consumption data and government investment plans.

Members of the legislature and inspection auditors need to know whether the government is investing resources in accordance with legal limits and parameters. The municipal manager himself and his advisers are often unaware of fundamental information about his management, due to the absence of an information visualization tool, which could be useful in supporting decision making.

The main objective of this article is precisely to present a proposal for structuring the visualization of government data by user profile. In view of the large volume of data available and the variety and complexity of this data, it is very difficult to create a single dashboard with visualizations covering all the information.

The motivation for conducting the research and development of a visualization tool for government information is due to the large amount of published open data and the lack of a specific visualization application.

Society's role is to use the data made available in order to increase its participation and interaction with public policies, obtaining a better understanding of government actions [7].

The execution of the work was concentrated on the publication of open data for the municipalities of the State of São Paulo, related to budget execution, however, the legal requirements in relation to budget execution, public accounting and transparency are the same for all Brazilian municipalities. The result of the work, therefore, can be useful to any Brazilian municipality.

The article is structured in eight sections: after this introduction, Sect. 2 deals with concepts related to Open Data, followed by Sect. 3 which deals with Information Visualization concepts, complemented by Sect. 4 which describes the construction process of the information visualization. Subsequently, Sect. 5 presents the methodology and results of the Systematic Literature Review carried out. In Sect. 6 the structure of the proposed solution is demonstrated, followed by Sect. 7 which presents the application's functionalities, the identification of user profiles and some visualization techniques used. Finally, Sect. 8 contains final considerations and future work.

2 Open Data

The Brazilian Open Data Portal [4] defines and characterizes what Open Data represents.

According to the definition of Open Knowledge International [8], data are open when anyone can freely access, use, modify and share them for any purpose, being subject to, at most, requirements that aim to preserve its provenance and its openness. Open Knowledge International is an international non-profit organization that promotes information sharing and the creation of free knowledge.

This is generally satisfied by publishing the data in an open format and under an open license.

Open data is also governed by three laws and eight principles.

The three laws were proposed by David Eaves [9], a public policy expert, open data activist and public policy speaker at the Harvard Kennedy School of Government. Are they:

- If the data cannot be found and indexed on the Web, it does not exist;
- If it is not open and available in a machine-understandable format, it cannot be reused; and
- If a legal device does not allow it to be replicated, it is not useful.

The laws were proposed for open government data, but it can be said that they apply to open data in general, even outside government environments.

According to OpenGovData [10], a working group of 30 people met in California, United States of America, to define the principles of Government Open Data.

They reached consensus on the following eight principles for open government data:

- **Complete**: All public data in electronic format are made available;
- **Primary**: Data are published in the form collected at the source;
- **Current**: Available quickly;
- **Accessible**: Available to the widest possible audience and varied purposes;
- **Processable by machine**: The data is reasonably structured;
- **Non-discriminatory access**: No identification or registration is required;
- **Non-proprietary formats**: No entity can have exclusive control over the data; and
- **Free licenses**: Free from restrictions by copyright regulations.

The next section deals with Information Visualization concepts, another important foundation for this work.

3 Information Views

To access open data and understand it, in the format that is published, it is necessary to have technical knowledge. The concept of open data is to share it as widely as possible. With the help of tools and applications for data analysis and information visualization, innovative solutions for society's problems that would never have been imagined by the data owner emerge.

Even so, it is important to share the perspectives brought by these visualizations, applications and infographics so that everyone can access the data, without having to have the technical knowledge to use the data directly.

When information is made available on the Web using appropriate open standards, it can be used over and over again in new, unforeseen and imaginative ways, capable of greatly increasing the value of the data through its reuse and combination, with greater automation and better interoperability. As the Many Minds Principle document, "The Many Minds Principle," says [11]: the coolest thing to do with your data will be imagined by someone else.

Information visualization is the presentation of quantitative data in a graphical form. In other words, information visualizations turn large and small data sets into visuals that are easier for the human brain to understand and process.

According to Yamaguchi [12], one of the factors to be considered for choosing the most appropriate information visualization technique is the type of data to be manipulated.

As an example, Yamaguchi [12] points out that visualizations of information can be used to discover unknown facts and trends. You can have presentations in the form of line charts to show changes over time. Bar and column charts are useful when looking at relationships and making comparisons. Pie charts are a great way to show parts-of-a-whole. And maps are the best way to visually share geographic data.

Taking into account that 90% of the information that passes through our brain is visual, according to Dagostin [13], Information Visualization through graphics ends up becoming the fastest and most direct way to disseminate information. In addition, the graphical visualization is shared much faster than any written content.

Organizations are making use of visualizations of information to support decision making. With the aid of data analysis software and visual statement generators, it is

easier to understand the reality of the business, offering better conditions for decisions based on data.

The next section complements the Information Visualization concept, presenting the information visualization construction process.

4 Information Visualization Construction Process

Card and others [14] define Information Visualization as "the use of visual representations of abstract data supported by computer and interactive to increase cognition". It is important to highlight the inclusion of the computational factor as a support for visual representations, which allows for possibilities of interaction and increases the power of cognition.

The objective is to facilitate the observer to see a certain phenomenon, through visualizations. Visualizations amplify cognition because almost half of the human brain is focused on the visual sense - we are good at interpreting graphic patterns and discovering trends, according to Ware [15]. This cognition capacity is enhanced by the advances in new technologies, which allow the creation of interactive, dynamic, animated and real-time visualizations [15].

To create visualizations that amplify cognition, it is necessary to think about the data to be presented, how to extract them, organize them and how to represent them to the user, in a way that helps him to perform one or more tasks. CARD and others [14] proposed a model that details the steps for creating visualizations.

Figure 1 illustrates a reference model for the information visualization process, presented by Card and others [14], which is formed by three stages.

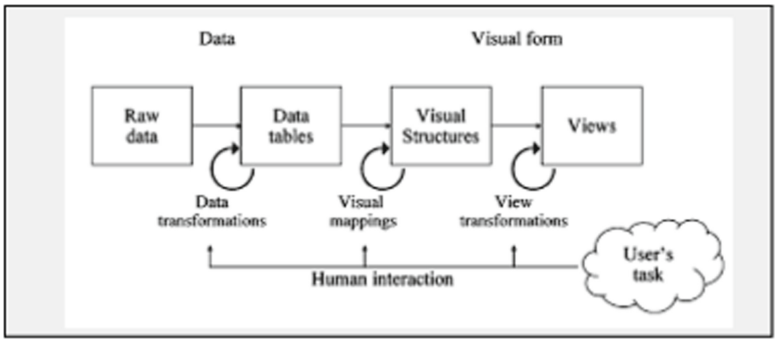

Fig. 1. Card et al. reference visualization model (Source: adapted from Card et al. 1999.)

The first step is called Data Transformations. In this step, a set of raw data is processed and structured, usually in the form of one or more tables. The processing seeks to perform data cleaning, elimination of redundancies, filtering, groupings and totalizations in categories and other classifications.

The second step is Visual Mapping, which involves building a structure that visually represents the data in the table. Usually represented by axes, such as the X and Y axes of the Cartesian plane.

The third stage of the visualization process is that of Visual Transformations, in which it is possible to modify and extend visual structures with the insertion of interaction resources. The interaction mechanisms allow the user to explore different scenarios for a better understanding of the visualized data.

According to Kirk [16] the data classification methods for viewing the information are: comparing categories; evaluation of hierarchies and relation parts of a whole, showing changes over time; plotting connections and relationships; and mapping of geospatial data.

According to Kirk [16], the criteria for building a consistent visualization are: data with statistical precision; precision in visualization; functional precision; visual inference; precision formatting; and accuracy in annotation. It is necessary to be rigorous with all the details, accurate and true in the representation of the data to obtain visualizations that are functional, effective and do not deceive the reader.

The next section presents the methodology and results of the Systematic Literature Review. Government areas of interest, tools, visual techniques and purpose of the applications present in the studies.

5 Systematic Literature Review (SLR)

SLR is a type of scientific research that aims to gather, critically evaluate and conduct a synthesis of the results of multiple primary studies [17]. It also aims to answer a clearly formulated question, using systematic and explicit methods to identify, select and evaluate relevant research, collect and analyze data from studies included in the SLR [18].

SLR is a technique that seeks to carry out comprehensive reviews of the literature on a topic, evaluating the results in a non-biased way, always explaining its selection criteria, so that the researcher who will use the SLR can evaluate the quality of it and execute it again. This technique was initially developed for use in the medical field, however, it is increasingly being used in other areas of knowledge [19].

The purpose of this review is to obtain an overview of the existence of applications for viewing information from open government data. Applications that allow the user to understand in a simple way the meaning of the data, without the need for technical knowledge, and that offer the user the possibility of interaction to express their criticism or opinion.

With the review, we intend to understand in depth the research field of visualization of information applied to open government data, especially budget execution data (public revenues and expenditures).

After a careful evaluation based on previous research, the following main research question was defined:

What information visualization tools and techniques are being applied to Government Open Data, seeking to facilitate the understanding of information to interested users, and what is the main purpose of these applications?

This main question can be expanded on to the following questions:

1) What is the governmental area, of open government data available?
2) What technologies are used for accessing and manipulating open government data, generating graphical visualizations and developing applications?
3) What information visualization techniques are explored by these applications?
4) What is the specific purpose of the application and what type of user is it intended for?

The searches were carried out in the ACM Digital Library, IEEE Xplore Digital Library and Scopus databases.

76 (seventy-six) studies were identified in the researched databases, and after the selection process, 38 (thirty-eight) studies for the extraction phase were included.

Several trends could be investigated during the execution of this review, among them, can be cited:

- **Governmental area**: the public area involves a multitude of services and administrative functions, and its management produces a huge amount of information with very different contents. In the analysis of the studies, it can be seen that the areas of greatest interest, referenced in the studies, are related to Finance/Public Budget and Public Health. In a second block, we have data on Education, Transport and Socioeconomic;
- **Tools for manipulating open data**: what has been observed as a trend is the use of data analysis tools, which allow importing data of different formats, defining data structures, establishing relationships between data sets, including those collected from different sources, being able to apply selections, filters and summaries;
- **Tools for generating graphical visualizations**: the observed trend is that the choice of a specific graphic resource is related to the type of visualization of the desired information;
- **Visual techniques**: the selected studies point to a clear trend of use in the application of visualization techniques on maps, followed by visualizations in graphs. Applications aimed at the area of public health and urban services, explore geolocation as the main information visualization technique; and
- **Main purpose of applications**: the most common trend is the classic purpose of demonstrating information, of giving publicity, also called public transparency. The goal is to make use of graphical visualizations to make it easier for users to understand the information. There is also a tendency to create applications that are useful or offer some service to the user. Other applications seek to provide the user with the power to inspect public administration and even interact, being able to participate in debates and decision-making.

In Fig. 2, it is possible to observe the classification of studies according to the purpose of the application.

In general, there is a trend in the development of common applications to improve the understanding of information using graphics, and applications that offer urban or public utility services using maps. It is noticed an increasing number of data sets published by

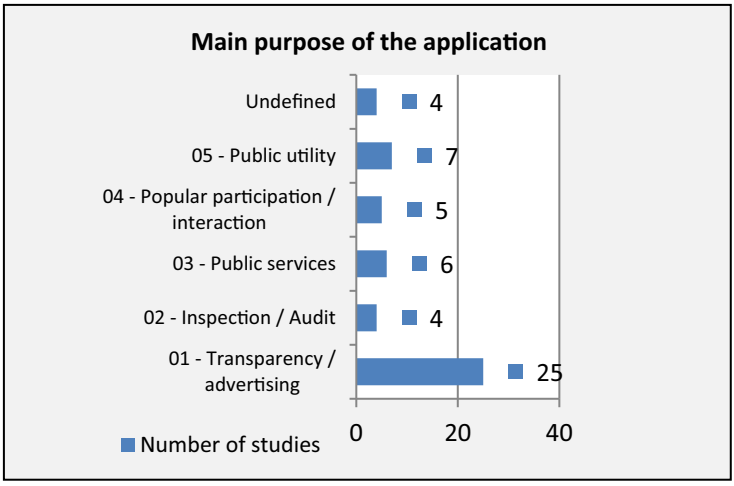

Fig. 2. Main purpose of the application (Source: produced by the author)

public agencies and the concern of many studies in finding a more adequate model for the process of accessing the data, preparation and visualization of the information of the same.

Among the studies, it is possible to highlight some that seek to create visualizations based on open government data:

1. In their study, Joohee Choi and Yla Tausczik [20], conclude that open government data analysis is expected to encourage citizens to participate in government, as well as to improve transparency and efficiency in government processes;
2. Maria Sokhn et al. [21], proposes in their study the display of information on panels, where the panels could be offered to citizens before public meetings so that they can have full knowledge of the administrative situation. Subsequently, citizens would have the means to contribute to political implementation, taking advantage of open data, in order to propose services and improvements. In the discussion of a particular problem, users could send comments, photos and videos to enrich their considerations to provoke a debate among the participants;
3. D. Burkhardt et al. [22], reports in his study on the possibility of adding indicators to open data, to improve understanding. Open data can only indicate a possible problem through deviations from the normal level, for example, if the unemployment rate has increased significantly within a short period of time. In fact, the information alone does not explain why there was a certain deviation. It is necessary to add other information (indicators) and relate to the data to better understand the deviation detected in the data. This allows for a much higher quality in the analysis and inspection of the information; and
4. Gebetsroither-Geringer, Stollnberger and Peters-Anders [23], designed to develop a web platform that can be used as a system of dialogue between participatory citizens. The platform must be flexible and generic to provide a basis for different topics to

be addressed within a city. In this way, citizens should be able to share their ideas in the decision-making process.

Among the technological resources highlighted in the studies we have the CSV data format and the graphic library D3.JS among the most cited.

In the next section, the structure of the proposed solution is demonstrated, from access to open data, data preparation and subsequent visualization.

6 Structure of the Solution

The main objective of the proposed solution is to view information on the budgetary execution of municipal administrations, through graphics organized on information display panels in the WEB environment, in specific formats according to the group of users.

With the help of the SLR results, the following requirements were raised to achieve the intended objective:

1. Data entry must provide access to more than one open data source and allow the relationship between them;
2. It must enable the insertion of performance indicators that allow a better interpretation of the calculated values;
3. The application must allow filters by categories, by government area, and other classifications;
4. It must make it possible to select the data by period ranges and value ranges; and
5. The data output must be structured according to the user's profile. The visualization of the information must have more than one format, to attend different groups of users.

Figure 3 shows the complete structure of the solution.

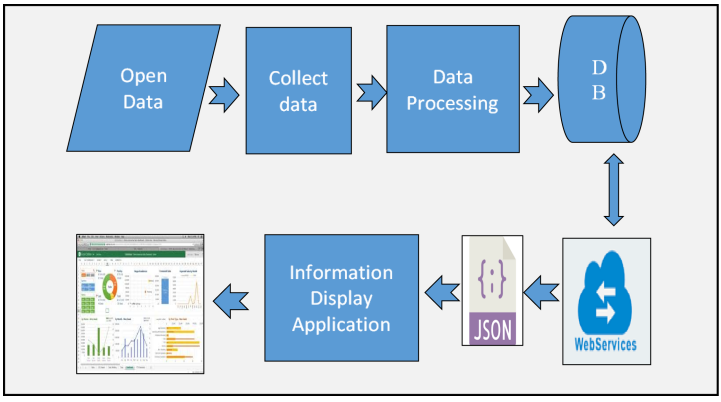

Fig. 3. Structure of the proposed solution (Source: produced by the author)

The structure presented can be divided into two parts, the first being responsible for the collection, processing of data and storage in a database and the second refers to the application of visualization of information based on data requested from this database through a WebService API.

The data collection provides access to more than one open data source, including data from the municipality itself, and the possibility of inserting performance indicators, while the visualization application must allow different formats of dashboards for different user profiles.

The main source of data for the application is available on the São Paulo Court of Auditors portal [6]. The portal provides data in the CSV format of budget revenues and expenses and debt data for municipalities in the State of São Paulo as of 2014. Access is via the address https://transparencia.tce.sp.gov.br, making it possible to choose the desired municipality and financial year.

The application performs data import in CSV format into SQL database. The relational database manager MySQL was used, created in a WEB environment.

A RESTful WebService application with PHP language and the SLIM Framework was developed especially to meet the data requests of the application.

Based on the user's choices, the application will be responsible for requesting the necessary data through a WebService application. The application will receive the data requested in JSON format in return and will be responsible for rendering the graphics that will be displayed in a panel of information visualizations.

For the development of the information visualization application, to make graphical presentations of data received in a WEB environment, the graphic library D3.JS was used, together with HTML5, CSS and JavaScript, non-proprietary resources.

The graphic library D3.js is among the main technological resources for creating information visualizations, according to the SLR. D3.js was the main resource used in the development of this application.

D3.js is a free-to-use JavaScript library, with incredible resources for creating dynamic and interactive visualizations, and has a lot of content on the internet and several galleries with templates ready for reuse. One of the galleries is available at https://www.d3-graph-gallery.com/, and has several models of graphic visualizations organized in categories such as Evolution, Maps and Flow.

D3 (Data-Driven Documents) was developed and launched in 2011 by Mike Bostock, Information Visualization specialist and graphic editor at the New York Times, to create interactive and personalized information visualizations in the web browser using SVG, HTML and CSS.

Figure 4 demonstrates the main architecture of the resources involved in the generation of graphical presentations with D3. The data stored in the MySQL database, accessed by a WebService RESTful application developed in PHP delivers the data in JSON format for D3 requests. D3 together with JavaScript manipulate the DOM by inserting data-oriented elements. Interaction and style resources are added and finally presented by HTML5 and CSS.

The main characteristics and features of D3 are: it explores modern web standards: SVG, HTML5 and CSS; data oriented in CSV, JSON and XML formats; allows you to manipulate the DOM (Document Object Model) based on your data; allows you to

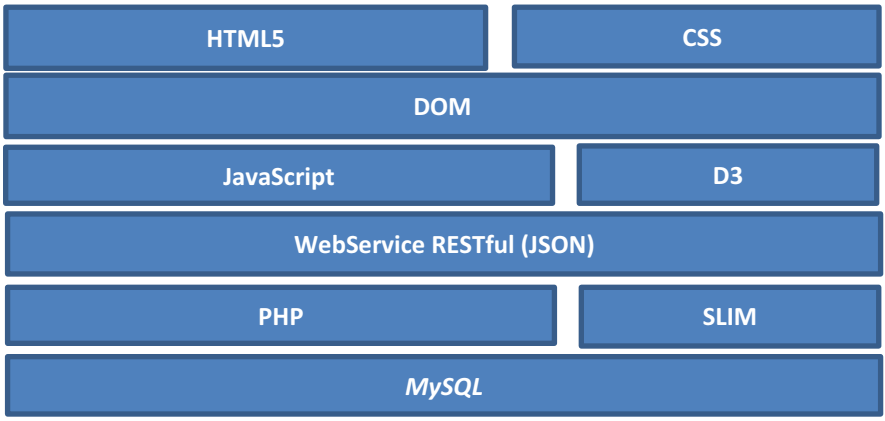

Fig. 4. Architecture of the information visualization application (Source: produced by the author)

dynamically generate new HTML elements and apply styles based on your data; dynamic properties: data can direct its styles and attributes; allows you to create graphic elements such as tables, pie charts or bars to geospatial maps; custom views with full control over your viewing capabilities; provides very powerful transition features; and great support for interaction and animation features.

The following web standards are widely used in D3: HTML5, DOM, CSS, SVG and JavaScript.

SVG (Scalable Vector Graphics) is an XML language to describe two-dimensional drawings and graphics in a vector way, either static, dynamic or animated.

One of the main characteristics of vector graphics is that they do not lose quality when they are enlarged.

The next section presents the application's features, the identification of user profiles and some visualization techniques used.

7 Application Features

In addition to the technological resources, it is necessary to define the scope of the proposed solution. Some questions need to be answered in relation to the target audience and the type of visualization desired.

For which audience do you want to make the application available? To create visualizations of functional information, it is necessary to know your target audience, to support project decisions and develop an application that meets the user's needs.

One of the main points of this work was the specification of the different user profiles, according to the participation of each element in public management. It was observed the role of each group in the municipal administration, their needs and what type of information they are interested in. This definition was fundamental for the development of the application and planning of the necessary visualizations.

7.1 User Profiles

As a result of this work, the following user profiles were identified:

- **Ordinary citizen**: it is the main focus of the information visualizations. For this group, visualizations of the information that demonstrate the application of public resources and results of public management in improving the quality of life and development of the municipality would be created;
- **Auditors and legislative agents**: responsible for authorizing the executive's projects, inspecting and approving public accounts. For this group that normally analyze a large amount of technical reports and copy of documents, the creation of graphics that demonstrate the allocation and application of approved resources in their respective categories and government projects, would be a tool that would support and facilitate the execution of your job;
- **Executive managers and public agents**: responsible for the application of public resources. For this group, it is important to create graphs that demonstrate the evolution of budgetary execution within the year, compliance with mandatory investments and respect for the limits imposed by legislation. Another need for the public manager is to assess how much has been invested in the development of the city and to monitor the evolution of the public debt;
- **Non-governmental entities, Social Organizations and the press**: entities that seek information from the government to inform and defend the interests of society. This group resembles ordinary citizens, but needs more elaborate and more detailed views; and
- **Entrepreneurs and public sector suppliers**: companies that seek information on consumption patterns and investments from the government. As it is a local database, it is important that businesspeople in the region are able to visualize the consumption needs of the local government to guide their sales strategies.

Thinking about the visualization needs for different profiles makes it possible to better plan the visualization of the information. It is important to put yourself in the shoes of a particular user and think about what would be important to them.

What about the type of information display? What you want to convey to the reader?

The intention regarding the function of the information visualizations is that it is more focused on "explaining" and "exploring". Transmit information to the reader in a specific and focused narrative, and allow discoveries free of interest and check patterns. The visualizations of the information will have a pragmatic and analytical tone, with more open and less conclusive content for an audience that seeks to interpret the information and reach their own conclusions.

It is not intended to impose an emotional and abstract tone, aiming to impress and impact readers.

These are important factors for an information visualization project. When establishing the purpose of the information visualization project, the desired tone and its function are defined.

7.2 Using the Stacked Area Visualization Technique

Stacked area graph resembles a line graph, but it presents a colored area below the line, and each series starts at the top of the previous series, and compare historical trends showing the proportion of the total that each category represents at any given time. It indicates general trends rather than individual values when comparing multiple data sets, showing how each category contributes to the cumulative total over time.

With the application of this technique, it is possible to demonstrate the evolution of income or expenses for several years, grouped into categories.

In Fig. 5, it is possible to observe the visualization in D3 using the stacked area technique. The categories can be activated or deactivated by the corresponding buttons, by the user's interaction, and the visualization is updated automatically, without the need to search for the data again. The user also has the possibility to select the type of display, for stacked area, flow chart or 100% stacked area. A table with information for each exercise is displayed when the mouse is placed over the image at the points corresponding to the X coordinate.

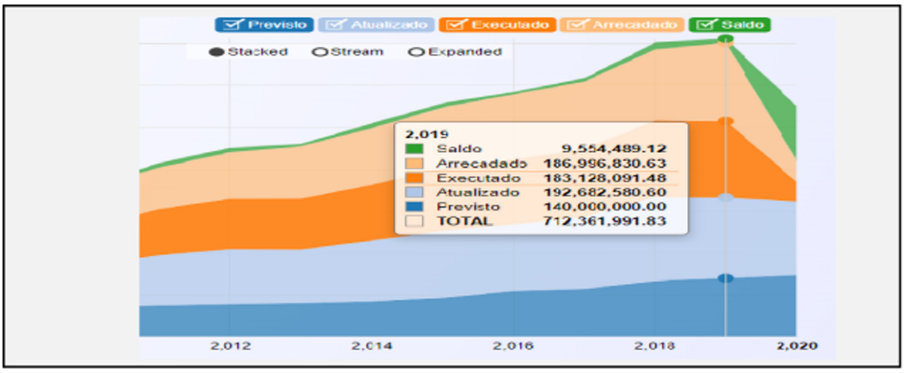

Fig. 5. D3 visualization using the stacked area technique (Source: produced by the author)

This view will be useful for multiple user profiles.

In Fig. 6, it is possible to observe the function call D3.json, containing the link "http://auditor.institucional.ws/app1/api/previsto" corresponding to the WebService application responsible for supplying the data. And in Fig. 7, it is possible to observe the return of data in JSON format related to the WebService call.

With this sequence of images, it was possible to demonstrate the operation of the application with the visualization of an interactive graphic, generated with resources from the D3.js library, the request for data through a WebService API and the return of the data in JSON format.

This information visualization using the stacked area technique was based on a model from the NVD3 library, which is a JavaScript library containing a set of components coded from the D3.js library. NVD3 allows developers to create reusable graphics, from a wide variety of models and codes that can be explored on the library's website (https://nvd3.org/).

```
d3.json("http://auditor.institucional.ws/app1/api/previsto",
        function(error, data) {
      d3.select('#chart1')
          .datum(data)
          .transition().duration(1000)
          .call(chart)
          .each('start', function() {
              setTimeout(function() {
                  d3.selectAll('#chart1 *').each(function() {
                      if(this.__transition__)
                          this.__transition__.duration = 1;
                  })
              }, 0)
          });
});
```

Fig. 6. Code snippet with the call from D3 to the WebService (Source: produced by the author)

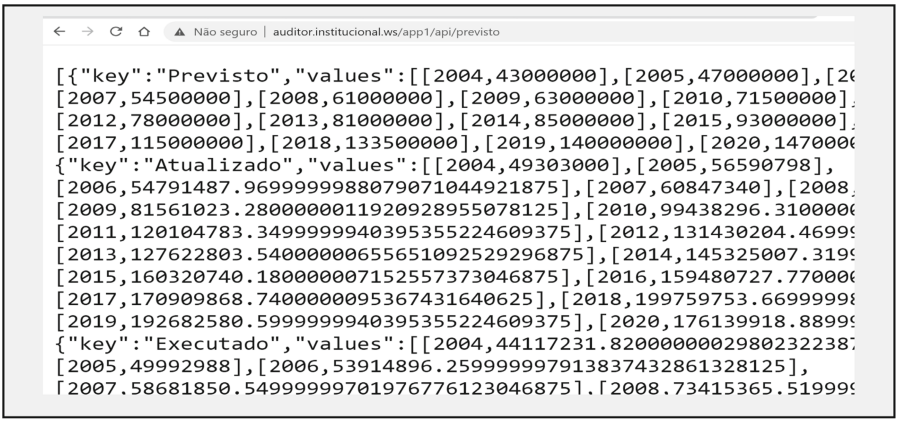

Fig. 7. Display of the JSON callback to the WebService (Source: produced by the author)

7.3 Use of Interaction and Multiple Views to Evaluate the Budget Proposal

One of the objectives planned for the application is to facilitate the evaluation by auditors and the legislature of the budget proposal for the following year, sent by the executive. This moment occurs once a year, between August and October of each year. The chief executive submits the budget proposal, which comprises a set of annexes with detailed statements of the values classified into categories and government projects distributed in the various budget units.

The evaluation of these documents first passes through the analysis of technicians and then goes on to debates and voting in plenary.

To facilitate the discussion around the proposal and support the speech and presentation of each councilor, the interesting thing is that a graphic presentation is prepared to facilitate understanding. Considering that the majority of the City Councils do not have

advisers for their Councilors, being in charge of preparing their exposition, normally their speech is not supported by a facilitating presentation.

This moment is of great importance for public administration, where the allocation of public resources for the next financial year will be defined. Which government projects will be carried out and the source of their funding.

For this moment, a visualization was planned that demonstrates the application of general resources and for each budgetary unit (Secretariats and Departments), and the amount applied in each unit can be evaluated. It will also be possible to make a comparison of the current proposal with the proposal approved in the previous year, in order to verify which Departments had an increase or reduction in the budget. Another possibility is the classification of the values into categories, which in this case divides the values proposed by the Public Agency (Legislative and Executive).

To develop this visualization, the bubble model was chosen from a gallery of models with D3 code, available on GITHUB.COM, on the page "https://github.com/d3/d3/wiki/Gallery". This model was used in 2013 by the New York Times in the newspaper's article about the American government's budget proposal for that year.

Figure 8 and Fig. 9 are the information visualization using different bubble grouping by size and colors, and show the budget proposal grouped by Management Area (Secretaries) and in General.

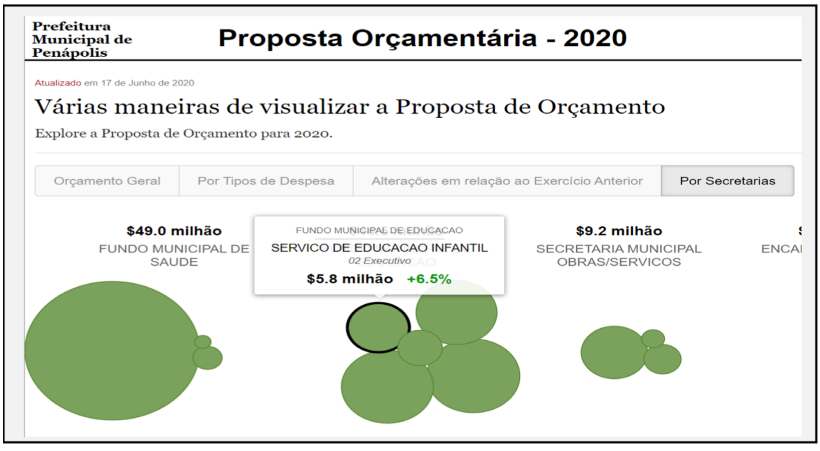

Fig. 8. Visualization in D3 with departments grouped into management area (Source: produced by the author)

Figure 8 shows visualization in D3 with the budget proposal grouped by Secretaries (Management Area). Each bubble represents a department that is part of a Secretariat. A Secretariat has at least one Department, and some Secretariats have up to five Departments. The size of the bubble indicates the value of the budget, on a scale of 1 to 90 pixels to determine the radius of the bubble, for a range of values from 0 to 50 million reais. The color of the bubble is defined by a gradient of shades of green for budgets with an increase in relation to the previous year and a gradient of shades of red for budgets with a reduction in relation to the previous year.

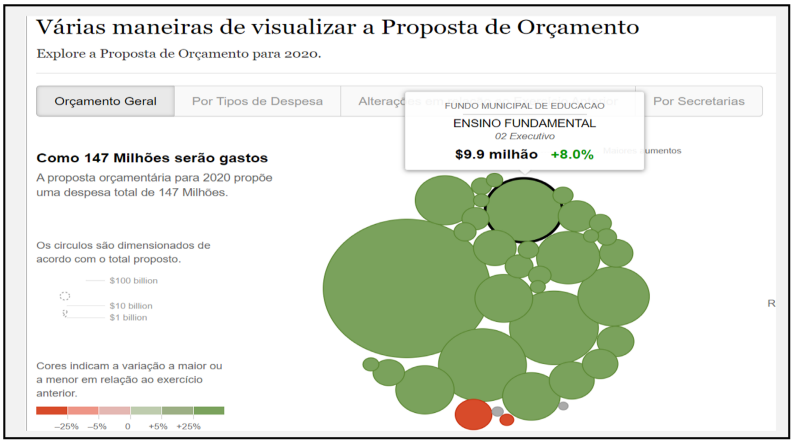

Fig. 9. D3 visualization of the general budget (Source: produced by the author)

A box with information from the Department is displayed when the mouse is placed over a certain bubble in the image.

Figure 9 shows visualization in D3 with the total of the budget proposal, and also contains the same characteristics of the grouping by Secretariats.

The same data can be evaluated in four different ways, totaled, by Secretariat, by Agency and by variation in relation to the previous year. The transition between the visualizations of the information occurs only by changing the position of the coordinates of each bubble in the image. The loading of the data with information from each Department (bubble) occurs only once, and the positioning of the bubbles is determined with D3 transition features and dynamic properties, adding interaction and animation to the visualization.

Despite starting from a ready-made model, available for use and reuse, this model is very complex with more than 10,000 lines, and had to have its structure modified according to the data structure, scale and X coordinate calculation strategy, Y, for positioning the bubbles in the image.

7.4 Using the Treemap Technique to Assess Budget Execution Within the year

In order to display the values executed at a certain point in the exercise, important information for the public manager, the visualization technique with hierarchical data and structured in a tree (treemap) was selected. This model initially allows the visualization of the total executed for each Secretariat. The size of the rectangle indicates the total value executed, ordered from largest to smallest value. The starting line of the table contains the total value executed in the exercise, and is used as the title of the path taken in the tree.

When clicking with the mouse on a particular Secretariat, the Secretariat expands into Departments. Each rectangle of the image contains its title and executed value, in addition to the total value of items subordinate to this item.

When clicking with the mouse on a specific Department, the Department expands to expense categories of this Department, such as Personnel Expenses, Services, Consumables, Permanent Material and Investment in Works/Installation.

The visualization therefore contains three levels of hierarchy, without considering the City Hall as a whole. The first level is the separation into Secretaries. The second level is separation of the Secretariat into Departments and the third level is the classification of the Department's executed values into Expense Categories.

For this visualization, a model was chosen from the D3 graphics gallery on the page "https://www.d3-graph-gallery.com". This gallery contains models with the basic structure of each type of chart, with some variations for each type.

Figure 10 shows the model with hierarchical data and structured in a tree.

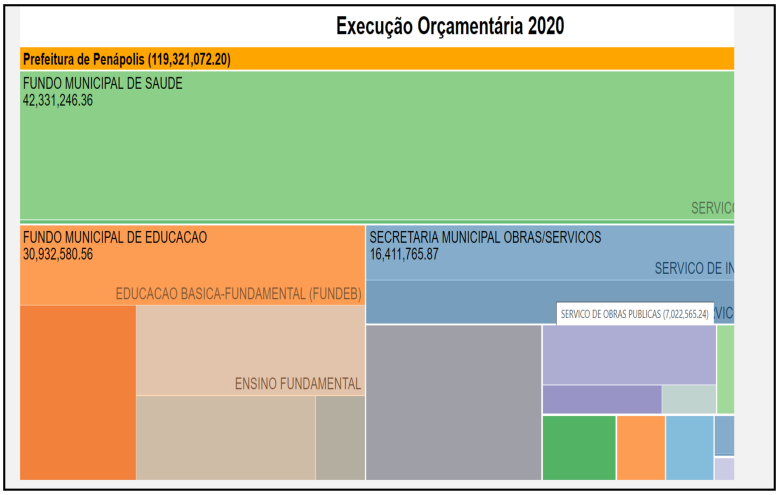

Fig. 10. View with hierarchical and tree-structured data – Treemap (Source: produced by the author)

It is important to highlight that three different approaches were used for the codification of D3. One approach was to make use of the NVD3 library, which contains several ready-made D3 models, available for reuse, exploring various visualization and interaction techniques. A second approach is related to the use of ready-made codes, previously used in a specific event or publication and made available by the authors to be modified and reused. These ready-made codes are usually available in D3 galleries on the internet. A third approach is related to the coding of D3 based on minimal structures, available in tutorials or on the library portal itself, at the address "https://d3js.org". In this case, development requires a greater coding effort, and a greater dependence on the programmer's knowledge and creativity.

The D3 library allows the free creation of new models with incredible visualization and interaction resources, being able to combine code reuse and new implementations.

The next section presents the final considerations and future work.

8 Conclusion and Future Work

The development of this work sought to present a development strategy for the application of information visualization for open government data from small municipalities.

The main data source used was the open data provided by the São Paulo State Court of Auditors [6], which is responsible for supervising the municipalities, complemented by other sources, including data from the municipality itself. The insertion of performance indicators allows the reader to better assess possible deviations from standards, such as whether the amount of investment in the educational area is related to the variation in the literacy rate.

The technological resources used and the information visualization techniques adopted were based on the SLR previously carried out, which resulted in the selection of 38 primary studies related to the theme of visualization of information from open government data. In this study it was also possible to evaluate the governmental areas that most interest the users and the purpose of the developed applications.

SLR revealed some trends:

1. That there is a clear growth in publications of open data sets, from the most varied areas, by public entities, and that there is a greater demand for applications that allow the exploration of these data in a simple and intuitive way;
2. That the proper presentation of data, motivates the exchange of information and improves communication. It is an essential tool when it comes to translating data and information into knowledge. This values the initiative to develop information visualization applications; and
3. That applications in general are based on open data from the federal government or large municipalities, whereas small municipalities do not have open data published. The data on budget execution of small municipalities, despite public transparency, are isolated without the possibility of exploitation.

In this work, it seeks to assist users with visualization of information from small municipalities, based on open data from the State and the municipality itself.

The presentation of the visualizations was also planned to be structured in user categories, with exclusive dashboards. The visualization and interaction techniques were selected according to the data set and needs for understanding the information, of interest to each user profile. The views are more attractive and interesting, for each category of users, because they are planned according to their interests.

For future work, we intend to make the application available for evaluating users of different profiles in order to improve the views of each category.

It is also intended for future work, considering that the open database contains information from all municipalities in the State, the possibility of creating views comparing data from different municipalities. It would be possible to compare income and expenses for a group of similar municipalities or even evaluate some information in the region or in the State using geolocation techniques.

Innumerable situations could be assessed using data from all municipalities in the State, such as investment in education per student, or health expenses per person. The

performance of the same supplier could be evaluated in several cities, including trying to find out if the values practiced in similar acquisitions are consistent.

Finally, this work sought to demonstrate that the development of a government data visualization application meets a current reality, the publication of an infinity of government data and information due to the legal obligation of public transparency, and the absence of a application of reading and presentation of these data, mainly for small municipalities. And implemented a new proposal for visualizations based on user profiles, seeking to serve different audiences, with different objectives in the analysis of the same database.

References

1. Brazil: Law 12.527, of November 18, 2011. Regulates access to information provided for in item XXXIII of article 5, in item II of & 3rd of art. 37 and in the 2nd of art.216 of the Federal Constitution; amends Law No. 8,112, of December 11, 1990; repeals Law No. 11,111, of May 5, 2005, and provisions of Law No. 8,159, of January 8, 1991; and make other arrangements. Federal Official Gazette, Brasilia, 19 November 2011. http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm
2. Brazil: [Constitution (1988)]. Constitution of the Federative Republic of Brazil. Presidency of the Republic, Brasília, DF (1988). http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm
3. Ramírez-Alujas, Á., Dassen, N.: Exchange rate: the advance of open government policies in Latin America and the Caribbean, Inter-American Development Bank, Technical note # idb-tn-629 (2014). https://publications.iadb.org/es/publicacion/16833/vientos-de-cambio-el-avance-de-las-politicas-de-gobierno-abierto-en-america
4. DADOS.GOV.BR. Portal: Brazilian Open Data Portal (2019). http://dados.gov.br/
5. Levine, R., et al.: The Cluetrain Manifesto. The End of Business as Usual, (The Cluetrain manifesto. The end of normal business), Reading: Perseus Book (2001)
6. TCE-SP: Court of auditors of the state of São Paulo (2020). https://www.tce.sp.gov.br/
7. EGOV-IMPROVING: Improving access to government through better use to the web. Accar, S., Alonso, J., Novak, K. (eds.) W3C Group Note (2009)
8. OKFN - Open Knowledge Foundation: Free knowledge network (2019). https://br.okfn.org/
9. Eaves, D.: The three laws of open government data (2009). https://eaves.ca/2009/09/30/three-law-of-open-government-data/
10. OpenGovData: 8 Principles of open government data (2007). https://public.resource.org/8_principles.html
11. Walsh, J., Pollock, R.: The many minds principle. Open Data and Componentization, XTech (2007), 17 May 2007. https://rufuspollock.com/many-minds/
12. Yamaguchi, J.K.: Guidelines for choosing visualization techniques applied to the knowledge extraction process. Master's Dissertation in Computer Science. State University of Maringá, Maringá.41 (2010)
13. Dagostin, D.: The use of tdics as an interactive tool in the teaching - learning process of chemical discipline in higher education. In: International Distance Education Symposium, São Carlos, Anais, São Carlos, UFSCAR (2014)
14. Card, S.K., Ackinlay, J.D., Shneiderman, B., Card, M.: Readings in information visualization: using vision to think. Morgan Kaufmann Series in Interactive Technologies, Academic Press (1999)
15. Ware, C.: Information visualization: perception for design. 3rd edition, Morgan Kaufmann Interactive Technologies Series (2012)

16. Kirk, A.: Data visualization: a sucessful design process, 1st. Ed, Packt Publishing Ltd, p. 206 (2012)
17. Cook, D.J., Mulrow, C.D., Haynes, R.B.: Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med. **126**(5), 376–380 (1997)
18. Clarke, M., Horton, R.: Bringing it all together: Lancet-Cochrane collaborate on systematic reviews. Lancet **2**(357), 1728 (2001)
19. Brereton, P., et al.: Lessons from applying the systematic literature review process within the software engineering domain. J. Syst. Softw. Elsevier Sci. Inc. **80**(4), 571–583, April 2007, ISSN 01641212. https://www.sciencedirect.com/science/article/pii/S01641212 0600197X?via%3Dihub
20. Choi, J., Tausczik, Y.: Characteristics of collaboration in the emerging practice of open data analysis. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2017), pp. 835–846. ACM, New York, NY, USA (2017). https://doi.org/10.1145/2998181.2998265
21. Sokhn, M., et al.: From data to decisions. In: Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2015 - 2016), pp. 361–362. ACM, New York, NY, USA (2016). https://doi.org/10.1145/2910019.2910032
22. Burkhardt, D., Nazemi, K., Retz, W., Kohlhammer, J.: Visual explanation of government-data for policy making through open-data inclusion. In: The 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014), pp. 83–89. London (2014). https://doi.org/10.1109/ICITST.2014.7038782
23. Gebetsroither-Geringer, E., Stollnberger, R., Peters-Anders, J.: Interactive spatial web-applications as new means of support for urban decision-making processes. ISPRS Ann. Photogrammetry, Remote Sens. Spat. Inf. Sci. **4**(4/W7), 59–66 (2018). https://doi.org/10.5194/isprs-annals-IV-4-W7-59-2018