

Platform of the Brazilian CSOs: Open Government Data and Crowdsourcing for the Promotion of Citizenship

Raul S. Ferreira
Information Technology
Department
Institute for Applied Economic
Research (IPEA)
Rio de Janeiro, Brazil
raul.ferreira@ipea.gov.br

Vagner Praia
Information Technology
Department
Institute for Applied Economic
Research (IPEA)
Rio de Janeiro, Brazil
vagner.silva@ipea.gov.br

Heraldo P. B. Filho
Information Technology
Department
Institute for Applied Economic
Research (IPEA)
Rio de Janeiro, Brazil
heraldo.filho@ipea.gov.br

Fabrício Bonecini
Department of Studies on
Democracy and State
Institute for Applied Economic
Research (IPEA)
Rio de Janeiro, Brazil
fabricio.almeida@ipea.gov.br

Andre Vieira
Department of Studies on
Democracy and State
Institute for Applied Economic
Research (IPEA)
Rio de Janeiro, Brazil
andre.vieira@ipea.gov.br

Felix Lopez
Department of Studies on
Democracy and State
Institute for Applied Economic
Research (IPEA)
Rio de Janeiro, Brazil
felix.lopez@ipea.gov.br

ABSTRACT

In Brazil and around the world, Civil Society Organizations (CSOs) provide valuable public services for society. Through CSOs, people have organized and defended their rights, communities and interests, and can fully exercise their collective potential, often acting in partnership with governments to carry out public policies and/or develop their own projects, financed by the private financing or being self-sufficient. Public transparency and availability of quality data are requirements for analyzing the strength and capacity of these organizations. Understanding the distribution of non-governmental organizations across the world and at the national scale, their areas of updating, projects in progress, and their execution capacity, is critical to promote the financing conditions of CSOs, to make it visible and to make it more effective, transparent, and strong. With these goals in mind, we developed the Civil Society Organizations Platform¹, an open, free and public on-line portal that provides a wide variety of information on the profile and performance of the population of CSOs in Brazil. Its core mission is to provide data, knowledge, and information on the role played by the almost 400,000 CSOs in activity in Brazil and their cooperation with the public administration in delivering public policies and services. We show how we developed this platform, the integration with several different databases, the challenges of working with open government data and how we integrated a lot of recent open source

¹<https://mapaosc.ipea.gov.br/>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SBSI 2017, June 5th-8th, 2017, Lavras, Minas Gerais, Brazil
Copyright SBC 2017.

technologies in all spheres of system development. The first empirical results are shown and some new features regarding public data are presented.

CCS Concepts

• **Applied computing** → **E-government**; *Sociology*;

Keywords

e-government, crowdsourcing, open data, CSOs map, big data

1. INTRODUCTION

According to [10], Civil Society is the aggregate of non-governmental organizations and institutions that manifest interests and will of citizens. The political agenda of a new regulatory framework for civil society organizations was enforced by a great number of long-established Brazilian Civil Society Organizations (CSO)² during the 2010 presidential electoral contest. These large CSOs wrote to both main presidential candidates an open letter³ making their demands. The elected government approved as a result of this public debate a new Regulatory Framework of Civil Society Organizations (MROSC)⁴, which has three priority objectives:

- Legal and regulatory innovations for the existing relationships and existing contracts between the third sector and the public administration
- Economic sustainability
- Standardization of CSO certification for tax exemptions

²<http://plataformaosc.org.br/plataforma/>

³<http://plataformaosc.org.br/em-carta-aberta-plataforma-define-prioridades-e-proximas-acoas/>

⁴<http://www.participa.br/osc#.WJ2ktVUrJxA>

As we can see in [20], United States-based NGOs believed that promoting the organizations image and fundraising were the two most important functions of new media for the NGOs, which suggests the importance of an online platform that promotes CSOs image at least at the national scale. By the year 2014, there was no public and user-friendly transparency or accountability mechanisms in Brazil for the population to access information on partnerships and agreements held by the public administration with the CSOs. Project like this would provide a valuable contribution to delivering fundamental public services: health, education, social assistance, services to the elders, aid to drug addicts, sports and culture services etc.

On the other hand, there were neither trustful data gatherings nor massive and acknowledgeable public presentation of data related to the population of Brazilian CSOs: its numbers, general characteristics, fields of activity and influence, spatial distribution on the national territory, number of people employed, amount of money and wealth circulated from the public and the private sector to the CSOs, the CSO main projects, their history and its accumulated knowledge.

[23] shows many advantages when technology is combined with CSOs, like cheap, global, many-to-many communications offered by network technologies that can provide an excellent platform for collaboration, easy publishing, widely mobilization and improved online surveillance of these organizations.

This work shows how the Civil Society Organizations Platform aggregates concise, comprehensive, and accessible information on all existing Brazilian CSOs and integrates data using several datasets providing updated information to public managers on where and what CSOs are doing, mainly in partnership with government. Besides, this platform also provides aggregated data for the decision-making of political actors, as well as public and private investors.

The article is organized as follows: This is the first section where we introduce the importance of having a digital platform for civil society organizations, public policies agencies, and population. Section 2 brings some related work about similar initiatives around the world which show systems and use cases about open data government and crowdsourcing. Section 3 explains how we develop the system and surpassed some challenges and how the architecture of this platform is organized. In the section 4 we present the first results, screenings and some features under development and we finish with section 5 where we present some conclusions, the next steps of developing this platform, and we try to point to some future directions for the field of open government data solutions.

2. OPEN GOVERNMENT DATA SYSTEMS

According to [14], electronic government (e-government) refers to government's use of technology, particularly web-based applications to enhance the access to and delivery of government information and service to citizens, business partners, employees, other agencies, and government entities.

According to [13], experiences with e-government initiatives tend to be chaotic and unmanageable and to present a number of challenges for public managers. Some of these challenges relate to infrastructures that must be built, policy issues that have to be resolved and interoperability restrictions. Figure 1 shows the intersection where open govern-

ment data solutions lies.

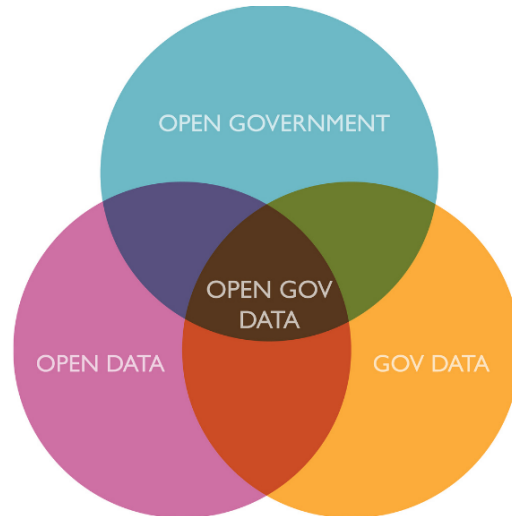


Figure 1: Open Government Data

In this type of systems, a free platform regarding CSOs that allows unrestricted access to public information is still rare in the world. Today, perhaps the most important singular initiative in the world is performed by GuideStar International[2], which is present in six countries:

- GuideStar UK⁵: consisting of more than 160 thousand organizations, it includes features for searching organizations and updating information by the organizations themselves.
- GuideStar EUA⁶: presents information related to about 1,8 million organizations and features for updating data by the organizations themselves, searching CSOs, interactive map for locating organizations and a data API available for the platform partners. Some of these features are available only through paid plans.
- GuideStar India⁷: contains information about more than 7,500 CSOs, includes functionalities for searching organizations by name or area of activity, updating data by the organizations themselves, interactive map for locating organizations, list of support programs for OSCs, and a certification program based on the estimated level of transparency and accountability held by the organization.
- GuideStar Belgium⁸: exhibits information related to about 3,000 organizations and allows inclusion and update of data by the organizations, interactive map for locating organizations, search of organizations by multiple filters, in addition to disseminating support programs for CSOs and donations to these organizations.

⁵<http://www.guidestar.org.uk/>

⁶<https://www.guidestar.org/>

⁷<http://www.guidestarindia.org/>

⁸<http://www.dono.guide/>

- GuideStar Israel⁹: contains information related to about 35,000 CSOs and has functionalities for searching CSOs by filters and updating data by the CSOs themselves.
- GuideStar South Korea¹⁰: presents information related to 9,000 organizations and features search of CSOs with multiple filters, CSOs classifications according to diverse criteria, evaluation system, statistics about CSOs and an area dedicated to disseminating and searching for donations to CSOs.

Other similar initiatives are:

- National Center for Charitable Statistics (NCCS)¹¹: with information from more than 1.5 million CSOs, has a simple search system and a classification system developed by the institution, as well as having a wide range of reports and statistics on CSOs.
- Charity Navigator^[18]: has information about 8500 organizations, with functionalities for searching CSOs using filters or relevant topics, besides a series of classifications developed by the initiative and a system for financial donations to organizations.
- Charity Commission (UK)^[17]: with a database of about 185 thousand organizations, has the functionalities of searching CSOs using various types of parameters, with updating of data carried out by CSOs, and a system for sending documents required annually to the CSOs of the United Kingdom.

Comparing the Platform of Brazilian CSOs initiative with these other initiatives, it is observed that the platform presented in this work has a range of functionalities and information about CSOs not verified in any of the other initiatives. In the other initiatives, it can be observed that some have a very focused purpose in just showing the individualized data of CSOs. On the other hand, there is a greater focus on the joint analysis of data, but they do not present a more accurate treatment for the individualized data distribution of CSOs. The Brazilian initiative contains functionalities both to present individualized information and analyses that show the global scenario of the Brazilian CSOs' performance. In addition, the features designed for the Brazilian Platform of CSOs aim to serve users with different profiles, from a user who wants to perform data analysis to a common user. The table 2.1 The ref table: comparative table displays a comparison of the platforms.

2.1 Open Data

According with Open Definition Initiative¹², open data means that data should be freely accessible by anyone to use, modify and share. In government, this is extremely important because of the open government concept, which states that citizens have the right to access the documents and proceedings of the government to allow for effective public oversight^[12]. Because of this, organizations have been proposing initiatives to discuss and promote open data and trans-

parency ^[1], ^[7], BASEL¹³, even Brazilian civil laws 131¹⁴ and 137¹⁵.

With the increasing amount of data available from governments, problems of large-scale distributed data integration, collaborative data manipulation and transparent data consumption have arisen that must be solved. In order to do this, initiatives such as ^[8] and ^[21] have been created.

^[8] is an open source infrastructure for data conversion, publishing, enhancement and access, and a community portal that educates and serves the growing international community of developers, data curators, and end users. It enables organizations to convert unstructured data from other sources to RDF. After that, it uses social semantic web and machine learning technologies to enhance the data, adding more links to it. The last step of its process is to enable queries to the data in many different ways.

^[21] presents the main transparency initiative in the United Kingdom (data.gov.uk) and lessons they've learned through building and maintaining it. The portal is a public data catalogue that points to thousands downloadable datasets, often available in CSV or spreadsheets. They present four challenges for the future of the portal, and of open data as a whole: understand how to build or reuse ontologies easily, query methods that scale across the Web, not just within well-structured data, visualization and browsing tools and populate the Linked Data Web to increase the network effects of large scale linking. Based on these mature related works presented above we can see in table 2.1 that our platform still has advantages over these initiatives:

2.2 Crowdsourcing

According to ^[4], crowdsourcing solutions can be categorized according to the types of problems they are suitable to solve. Based on that concept, the authors defined four types of crowdsourcing problems: knowledge discovery and management, distributed human intelligence tasking, broadcast search and peer-vetted creative production. For example, SeeClickFix^[15] is a platform that let users input data about problems happening in the city, like light outages. All data goes to a common location that the organization can use to improve services. Another crowdsourcing platform for citizens may be found in ^[9], where a system was developed to help citizens of city of Rio de Janeiro to supervising public transport problems in the city, where each citizen acts as a supervisor, sending complaints to public agencies and to database of the system in real time.

The platform also is composed by a crowdsourcing component present in inserting and editing pages. This pages allows CSO representatives to input data about the CSO they represent, producing thus, innovative and accurate information neither government nor citizens knew about, which will lead to new studies and insights.

3. THE CSO PLATFORM

Building an e-government solution may present many challenges to be addressed, whose solutions may follow strategies different from those adopted in general by private companies. In order to tackle these challenges starting from a

⁹<http://www.guidestar.org.il/>

¹⁰<http://www.guidestar.or.kr/>

¹¹<http://nccs.urban.org/>

¹²<http://opendefinition.org>

¹³<http://www.bis.org/>

¹⁴www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp131.htm

¹⁵http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm

Systems	Public API	CSO Amount	Interactive map	Interactive Graphics	Data updating by CSO	Media sharing
CSO Map	Yes	400,000	Yes	Yes	Yes	Yes
GuideStar UK	No	160,00	No	No	Yes	No
GuideStar EUA	Yes	1,800,000	Yes	No	Yes	Yes
GuideStar India	Yes	7,500	Yes	No	Yes	No
GuideStar Belgium	No	3,000	Yes	No	Yes	Yes
GuideStar Israel	No	35,000	No	No	Yes	No
GuideStar South Korea	No	9,000	No	Yes	No	No
NCCS	No	1,500,000	No	Yes	No	Yes
Charity Navigator	No	8,500	No	Yes	No	No
Charity Commission	No	85,000	No	No	Yes	No

Table 1: Comparative table.

solid background, we took on account some important tips from [13], who presents four stages of a growth model for e-government: (1) cataloguing, (2) transaction, (3) vertical integration, and (4) horizontal integration. By considering this model, the CSO Platform team was capable of starting from a solid background process to achieve the final result. Below we present some characteristics of the databases and architecture we adopted.

3.1 Databases

The CSO platform team is currently cataloging databases from the Brazilian Federal Government, making efforts to obtain, integrate and make the data available online. The resulting database model was developed using some important strategies:

- All SQL code are inside functions, thus we let the databases to do what they do best
- R-Tree indexing at geometry bounds: Speeds up spatial searches
- Materialized views serving the API: These views execute faster than process queries with many joins
- Triggers updating views and cache during daybreak
- Databases are isolated from each other
- We constructed many database schemas where each one has their own responsibility, for instance, the schema "portal" contains the views that serves the API services

Currently, the CSO platform database integrates 15 databases from Federal Government, gathering data related to the population of CSOs, like registration, certificates, participation in council of public policies and partnerships with governments. These databases are splitted into three blocks: CSO profile, public resources and social participation.

3.1.1 CSOs Profile

The databases from this block present information about CSOs originated from administrative records and official data about organizations certified by Federal Agencies. Besides, this information also comes from the CSOs themselves as company name, trading name, address, telephone number, e-mail, number of formally employed workers and so forth, as reported to the Annual Report of Social Information (ARSI). A description of each database are presented below and [22] shows other relevant details about these databases and their origin:

- **Annual Report on Social Information (Relação Anual de Informações Sociais or RAIS, in Portuguese):** RAIS is an annual census of the Brazilian formal labor market managed by the Labor Ministry (Ministério do Trabalho e Emprego or MTE). The sample selected for the CSO Platform contains 391,371 organizations and includes only non-profit entities formally registered as legal persons.
- **Civil Society Organization of Public Interest (OSCIP, in Portuguese):** Private law, non-profit legal entities may be granted Public Interest Civil Society Organization status provided that their organizational purposes include at least one of the following: promote social assistance, promote free education, defense and preservation of the environment, promote sustainable development and so on. This database contains about 7,000 registered organizations.
- **Social Assistance Entities Certification (CEBAS, in Portuguese) in Social Assistance, Education and Healthcare:** These databases comprise more than 4,000 organizations, presenting information about the certificate status and effective dates. CEBAS in Healthcare comprises even more data, as the organization size, type of management, administrative sphere, type of health unit and medical expertise of services.
- **National Secretariat of Solidarity Economy (SENAES, in Portuguese):** Contains 500 registers.
- **National Register of Environmental Entities (CNEA, in Portuguese):** Contains 650 registers.
- **National Register of Social Assistance (CNEAS, in Portuguese):** Contains 42,000 registers.
- **National Registry of Health Establishments (CNES, in Portuguese):** Contains 23,000 organizations. CNES is a record managed by the Ministry of Health that lists entities which have among its objectives the provision of health services. This database includes information about the administrative dependence, type of management, type of health unit and medical expertise of provided services.
- **Census of the Social Assistance System (Censo SUAS, in Portuguese):** The SUAS Census is managed by Ministry of Social and Agrarian Development and collects annually information from social assistance departments and councils throughout the country. The Census contains data on social assistance

management, physical structures, human resources and the services and benefits provided by the various social assistance outlets in Brazil.

3.1.2 Public resources

The databases in this section provide information about the organizations and resources provided through their partnerships with government, containing data related to the area of activities (culture, sports, research and development etc.), projects titles and objectives, effective dates, authorized and paid values, and many other attributes. These databases are:

- **System to Support Fiscal Incentive Laws for Culture (SALIC, in Portuguese):** The SALIC is a system of the Ministry of Culture to control the execution of cultural projects elaborated by individuals and legal entities supported by the government. This database currently contains 6,000 project records.
- **System of the Sports Incentive Law (SLIE):** Contains 500 registers. SLIE is a system of the Ministry of Sports designed to register and control proposals related to tax incentives for individuals and corporations that donate or sponsor sports projects.
- **Financier of Studies and Projects (FINEP):** FINEP is a public company linked to the Ministry of Science, Technology and Innovation that aims to develop Brazil economically and socially through public investment in Science, Technology and Innovation in projects that cover companies, universities and other public and private institutions containing 2,000 records.
- **Integrated Financial Management System (SIAFI):** SIAFI is a system of the Ministry of Finance, which is the main instrument for recording, monitoring and controlling the budgetary, financial and equity execution of the Federal Government. This database contains 100,000 records.
- **Conveners and Contracts Management System (SICONV):** Presents and process information about resources transferences from federal government to public agencies and private organizations with daily updates. Contains 110,000 conveners registered and more than 630,000 proposals.

In the field of social participation we collect **CSOs that represents institutions in National Public policies** containing 500 registers.

3.2 System Architecture

The system architecture of the Platform of Brazilian CSOs is modeled at a high level with a micro-services architecture [19], being composed of several systems with specific purposes. The figure 2 shows how the architecture is composed. As the central component of the Platform, the repository is a document-oriented NoSQL database where all data from government sources is stored. This database is fed through ETL processes and serves systems that compose the Platform of Brazilian CSOs: BI System, Open Data System and Map System.

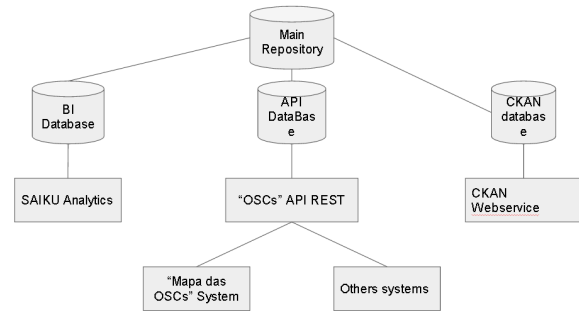


Figure 2: System architecture

- **BI System:** The purpose of this system is to make available, to the public, different ways to obtain and view information about CSOs. The system is formed by a relational database with a dimensional modeling and an analytics system, which offers user-friendly control to build queries about CSO data, as well as having several ways to visualizing data.
- **Open Data System:** The purpose of this system is to improve the publication of data about the CSOs, through the availability of data in various formats. Using a relational database, the system uses the open source system for open data [24], which offers a variety of tools to facilitate the publicity of the data.
- **Map System:** The main purpose of the Platform of Brazilian CSOs – the Map System – has the following objectives: To give more transparency to CSOs, to provide more information on the importance and diversity of projects and activities conducted by these organizations, acting as data provider, to foster research on CSOs and support public managers in making public policy decisions with the interface with CSOs.

The Map System use three databases and two systems, as seen in Figure 3. These databases and systems are divided into two components, API Map System and the Client Map System.

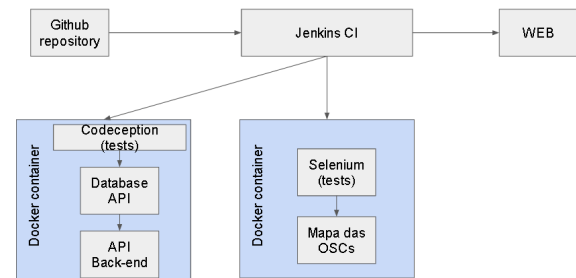


Figure 3: Automated tests and deployment

1. **API System:** The Map System API consists of a main database, another database used for caching, and a data API.
 - (a) **Main database.** Relational database, which stores the CSO data from the central repository. In addition to the structure of tables to store the data,

this database contains several functions and materialized views used as an abstraction layer for access to data through the API, described later.

- (b) **Cache system:** Developed with the NoSQL database with key/value model called REDIS[5], this database is used as cache to the main database, reducing the number of requests that this database receives.
- (c) **Data API:** Data API provides the services required to obtain and edit data from the main Map System database. This API is the external interface of the system.

2. **Map Client:** The Map System client is the web system which the end user can interact with the System. Besides functionalities of this system, there are a CSO georeferenced search system, a page with data about each CSO operating in the country, and the graphs that present CSO data in a concise way.

3.2.1 Data Integration and Knowledge Discovering

Data integration is made with Pentaho[3] and we perform knowledge discovering inside this big database with more than 2 million of registers using some data cleaning tools. Besides, we use SAIKU¹⁶ to perform BI and knowledge discovering. The platform already have some demands regarding machine learning and data mining features for next version. See future work inside Conclusion section.

3.2.2 Automated Deployment and Testing Routines

The developed test and automation flow can be viewed at Figure 3. Our code is stored on GITHUB and Jenkins¹⁷ is responsible for deliver development code to main repository to production. As mentioned before, the platform is divided in two applications: Map and API. Each one of these applications is stored inside a Docker[16] container and thus, we can provide isolation between two applications. Inside each container we put a testing tool depending on the needs. For client applications we use Selenium[11] and for back-end regarding the API we use automated tests with Codeception framework¹⁸. If test pass, then Jenkins collects code and puts in the repository and after that, updates the production server.

4. RESULTS

The system can be accessed online and the repositories regarding the systems are free accessible¹⁹, since the entire technology is made by open source tools and we promote the knowledge sharing and open data, everyone can access the code in GITHUB and contribute as want to. Figure 4 shows one system of this platform, the CSOs map interface, containing a "choropleth" representing in colors, the distribution of CSOs in the country, spatial data clustered with compressed data and main information about CSO.

Figure 5 shows the result list tab containing the same information but instead map we show some additional information and allows visualize all information about this CSO in a special CSO page. The platform also allows the CSO

representative to insert, delete or edit almost all information in this page, making a collaborative data inserting, thus, this platform can collect information that neither federal public agencies has. The developed platform allows:

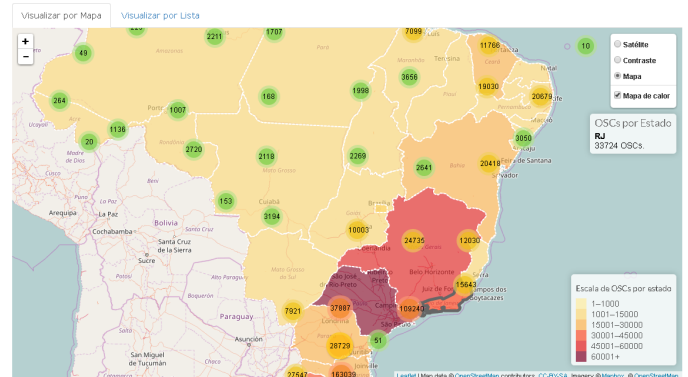


Figure 4: Main map

NOME DA OSC	CNPJ	NATUREZA JURIDICA	ENDEREÇO	DETALHAR
002/Es Grupo Escoteiro Loren Reno	7744937000110	Associação Privada	RUA PADRE JOSE CARLOS, 0, CAMPO GRANDE, Cariacica, 29146050	Detalhar
005 Guarani Esporte Clube	16764334000171	Associação Privada	RUA FERNAO DIAS, 1244, PORTO VELHO, Divinópolis, 35500432	Detalhar
0073 Associacao Missionaria Cat Sagrado	47432562000500	Associação Privada	SILVINO FERREIRA DO NASCIMENTO, 102, ITAPOA, Ivinhema, 79740000	Detalhar

Figure 5: Result List

- Streamlines and simplifies the communication and acquisition of information from and about organizations
- It provides a digital information space in which CSOs can complement data acquired from public and official sources
- It reports on all Brazilian CSOs to the world through a free public website
- It maintains an open and free technology information to enable transparency of the rendering of value-added services by intermediaries
- Source of innovation and sharing of best practices and technologies
- The Civil Society Organizations geographical search Map
- CSOs Search toolbar
- Search engine for project's funding calls
- Indicators and infographics that neither the government agencies has
- Accessibility Menu (for people with disabilities)
- Registering CSO Representatives section
- Comprehensive and Mobile-friendly CSO Profile Page

¹⁶<http://meteorite.bi/products/saiku>

¹⁷<https://jenkins.io/>

¹⁸<https://codeception.com/>

¹⁹<https://github.com/Plataformas-Cidania/>

- Daily Public Data Update
- Correction of the public resources transferred to CSOs by inflation indexes
- Tutorial for registration of representative of CSO and editing of profile page
- A complete public Data Extractor

Besides, the graphics (see Figure 6) containing information about CSOs for the first time in the country, this platform can be very beneficial for researchers and general public can collect all cleansed information used inside this platform through the CKAN[24] interface and thus, use this data through http requests without worrying to code ETL process or make data fusion processes.

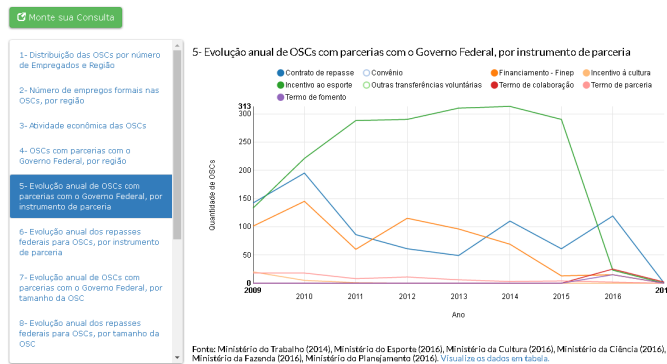


Figure 6: Graphics

5. CONCLUSIONS

We have created a broad platform on the CSOs sector in a national level, where there was no systematic information aimed for public transparency and accountability and provided to the public administration an innovative tool and new reliable public services in partnership with CSOs protecting and simultaneously empowering the CSOs helping on its path of credibility while providing tools for its long-term sustainability, through legal and institutional reform. Is also worth mentioning that this system provides trustful data for the academic sector focused on the CSOs reality, while continuously delivering reports and publications for the society. Our public service contributed to enhance the well-being of people by offering to a diverse public of beneficiaries the following goods:

- Public and private project financiers: It satisfies in the first instance the search and selection of CSOs for project financing. Allows the mapping of CSOs with experience by area of activity, helping the project funder to find those organizations that meet the requirements of working in a specific community. Allows funders to identify many CSOs for comparison purposes.
- Civil Society Organizations (CSOs): It offers to CSOs, regardless of their size, means and opportunities to make their work without costs, public and known, demonstrating their public accountability in a transparent way. Streamlines the relationship and communication

between CSOs and different project funders (sources of funds), researchers and institutions that intermediate private donations (individual and institutionalized). Allows CSOs to understand the work of other Brazilian organizations and to exchange information, experiences and good practices among themselves, facilitating the formation of networks. To a large extent, the system can potentially reduce the medium and long-term costs of resource mobilization by CSOs by providing information that will assist both CSOs and their institutional, private and/or individual funders and donors to identify projects of interest.

- Private individual and institutional donors: It supports the growth and public trust of private and institutional donation online and allows existing donation systems to strengthen. It gives donors the tools to identify, compare and track the records of the activities carried out by the CSOs they wish to support. It allows donors to have a mechanism to take responsibility for their donation, being able to verify the fundraising capacity of an organization and its correlated capacity of execution, while evaluating the impact of its work.
- Regulation and public control: Provides data and analysis tools to support public transparency. It generates ever better information, data and reports to monitor the performance of the sector in national public life. It supports the publication of objective data sent by the organizations themselves, saving the cost of control and diligence on the part of the public administration. It establishes an electronic system that can replace inefficient systems – already existing or preventively – as well as inhibits by public transparency and mutual agreement intrusive practices towards CSOs rights.
- Public Policy Managers: It allows public managers to identify territories that are over or under represented in terms of the availability and/or performance on the delivering of public services by CSOs. Supports decision-making, allowing policymakers to track trends in CSOs funding and action by area and territory.
- Professionals: It generates information that will help accountants, lawyers, and consultants advise the organizations they serve.
- Researchers and scholars in the area: It Generates key information for statistical purposes and analysis of government performance. Offer data and information to support the work of researchers and academics, qualifying the process of analysis and production of knowledge regarding the CSOs sector.

As next steps we are working with Cloudera[6] stack to perform automatic classification for CSOs, aiming to automate this hard work process that is made today by IBGE²⁰. Some analytics will be made to try find fraud problems inside CSOs universe. We expect to deliver a better processing when we move our system to a Hadoop cluster and thus, perform better results with our millions of registers with dozens of features.

²⁰ www.ibge.gov.br/

6. ADDITIONAL AUTHORS

Rafael L. Tavares (Institute for Applied Economic Research (IPEA)), email: rafael.tavares@ppgi.ufrj.br

7. REFERENCES

- [1] S. A. Aaronson. Limited partnership: Business, government, civil society, and the public in the extractive industries transparency initiative (eiti). *Public Administration and Development*, 31(1):50–63, 2011.
- [2] A. L. Bies. Evolution of nonprofit self-regulation in europe. *Nonprofit and Voluntary Sector Quarterly*, 2010.
- [3] R. Bouman and J. Van Dongen. Pentaho solutions. *Business Intelligence and Data Warehousing with Pentaho and MYSQL*, 2009.
- [4] D. C. Brabham. Using crowdsourcing in government. *IBM Center for the Business of Government*, pages 1–42, 2013.
- [5] J. L. Carlson. *Redis in Action*. Manning Publications Co., 2013.
- [6] I. Cloudera. Cdh proven, enterprise-ready hadoop distribution–100% open source, 2012.
- [7] L. A. Cunningham. The sarbanes-oxley yawn: Heavy rhetoric, light reform (and it just might work). *Conn. L. Rev.*, 35:915, 2002.
- [8] D. DiFranzo, L. Ding, J. S. Erickson, X. Li, T. Lebo, J. Michaelis, A. Graves, G. T. Williams, J. G. Zheng, J. Flores, et al. Twc logd: A portal for linking open government data. 2010.
- [9] R. S. Ferreira, R. Prata, K. P. Cotta, A. Figueiredo, V. Mororo, C. E. Barbosa, and J. M. Souza. Transport: Collaborative supervision of the public transportation. *Proceedings of the 32nd Annual ACM Symposium on Applied Computing*, 2017, in press.
- [10] P. N. Grabosky. Using non-governmental resources to foster regulatory compliance. *Governance*, 8(4):527–550, 1995.
- [11] A. Holmes and M. Kellogg. Automating functional tests using selenium. In *Agile Conference, 2006*, pages 6–pp. IEEE, 2006.
- [12] D. Lathrop and L. Ruma. *Open government: Collaboration, transparency, and participation in practice*. ” O’Reilly Media, Inc.”, 2010.
- [13] K. Layne and J. Lee. Developing fully functional e-government: A four stage model. *Government information quarterly*, 18(2):122–136, 2001.
- [14] D. L. McClure. Statement of david l. mcclure, electronic government: Federal initiatives are evolving rapidly but they face significant challenges. *Committee on Government Reform*, 2000.
- [15] I. Mergel. Distributed democracy: Seeclckfix. com for crowdsourced issue reporting. 2012.
- [16] D. Merkel. Docker: lightweight linux containers for consistent development and deployment. *Linux Journal*, 2014(239):2, 2014.
- [17] G. G. Morgan. The use of uk charity accounts data for researching the performance of voluntary organisations. *Voluntary sector review*, 2(2):213–230, 2011.
- [18] C. Navigator. Charity navigator your guide to intelligent giving. *Livestrong Foundation*, 2013.
- [19] S. Neuman. Building microservices: Designing fine-grained systems, 2015.
- [20] H. Seo, J. Y. Kim, and S.-U. Yang. Global activism and new media: A study of transnational ngos’ online public relations. *Public Relations Review*, 35(2):123–126, 2009.
- [21] N. Shadbolt, K. O’Hara, T. Berners-Lee, N. Gibbins, H. Glaser, W. Hall, et al. Linked open government data: Lessons from data. gov. uk. *IEEE Intelligent Systems*, 27(3):16–24, 2012.
- [22] B. Souto, E. Pedrozo Jr, and S. Goldbaum. Perfil das oscs e de suas parcerias com a união e mapa das oscs: bases de dados e resultados alcançados. 2015.
- [23] M. Surman and K. Reilly. Appropriating the internet for social change: towards the strategic use of networked technologies by transnational civil society organizations. *Social Science Research Council*, 2003.
- [24] J. Winn et al. Open data and the academy: An evaluation of ckan for research data management. 2013.