Exploring the impacts of web publishing budgetary information at the sub-national level in Brazil

Gisele S. Craveiro  
gisellesc@usp.br  
Jorge A. S. Machado  
machado@usp.br  
Andres M. R. Martano  
andres.martano@usp.br  
Tamiris Jesus Souza  
tamiris.jesus.souza@usp.br

See full findings at:  
colab.each.usp.br

Introduction & method

The Transparency Law in Brazil forces all public entities to web publish detailed budget data in real time. This law started in 2009, and 2013 was the final deadline for more than 5,500 Brazilian municipalities to comply with this legislation. There is a lack in the literature about how public managers and civil society organizations deal with budgetary data.

In this project, we analyze the uses of a tool Cuidando do Meu Bairro (Caring for my neighborhood) that geo-locates expenditures from open budget data. The aim of this project is to understand how access to budget information affects the relationship between civil society and public administrators in district/municipal level. Following this, we have these research questions:

- Is budget data geolocation effective for better control and supervision over the use of public resources?
- Can this project contribute to the engagement or empowerment of citizens?
- Better access to budgetary information can influence the construction of policy agendas within social organizations?

To have the answers, we carried out three workshops gathering around 250 social leaders, grassroots movements and interested citizens. In these occasions budget concepts were explained with the help of the tool. We also interviewed in depth a participant sample and public managers.

Evidence and findings

Despite the fact that budget data provided by São Paulo municipal administration has no lat./lon. coordinates, the Cuidando do Meu Bairro tool searches for textual elements that can give some detail about the location of expenditure and generates a geo-coded dataset. However, many budget activities are described in an aggregated/generic way. This prevents the Cuidando do Meu Bairro being presented as a tool to track and monitor public budget implementation in full.

Insufficient data quality could frustrate peoples initial interests and motivation to use the tool. However, very positive feedback was received from the use that could be made of the tool, including among others: invitations to give workshops and to replicate the project in other cities, a MoU with São Paulo public administration and nationwide recognition given by Brazilian media and Google Global Impact Challenge-Brazil prize. Due to these results, we conclude that budget geo-location is a good strategy to make difficult and/or abstract budget concepts, data and information more accessible to a wider audience.

The project interviews and outputs raise a series of new questions. First of all, the civil society ecosystem and the quality of its components and connections may counterbalance poor data quality supplied by governments. Results suggested the importance of better understanding the role of intermediaries and consumers, as well as their context. Besides the heterogeneity inside each category, their role can interchange depending on the situation, disrupting the static notion of passive-active relationship. In order to understand how actions and impacts are achieved is also important to understand which are the steps of data transformation, information and knowledge generation carried out through interaction of particular consumers and intermediaries. Other important findings come from the supply side of public sector information, represented through an iceberg metaphor. The visible published open data masks deeper underlying resource, process and cultural issues and constraints which potentially make open data initiatives unsustainable. On the other hand, moving towards open data can lead to changes in these underlying and hidden process, leading to shifts in the way government handles its own data.

Conclusions

The project has completed its first cycle of development with some important accomplishments and lessons learned both from the demand side (civil society organizations, academia and hacker activist) and the supply side. Using an action research methodology, the authors have worked since data collection and tool development stages: presenting and discussing budget data and public policies with civil society organizations and grassroots movements. In this scenario we could see that open data clearly disrupted the Data-Information-Knowledge value relationship, where the value of Data-in-use can be as important as the value of knowledge. Traditionally, consumers are presented with information (statistics) and knowledge, but increasingly consumers or intermediaries want to get access to the data at the core. More research work has to be done to fully address the issue of open data impact assessment methodologies, in particular, focusing on the intermediary ecosystem, rather than starting from data and tracing down.

Figure 1: Geolocated Budget Visualization with the help of the tool Cuidando-do Meu Bairro (Caring for My Neighborhood). Available at: cuidando.org.br

Figure 2: Percentage of expenses in relation to elements that allow inference of location

Figure 3: Amount of resources (in billions of Brazilian Reais) in relation to the values of activities and also the existence of information to identify the localization.

Figure 4 and 5: Models of the civil society ecosystem interaction and transformations from data to knowledge.

Evidence and findings

Despite the fact that budget data provided by São Paulo municipal administration has no lat./lon. coordinates, the Cuidando do Meu Bairro tool searches for textual elements that can give some detail about the location of expenditure and generates a geo-coded dataset. However, many budget activities are described in an aggregated/generic way. This prevents the Cuidando do Meu Bairro being presented as a tool to track and monitor public budget implementation in full.

Insufficient data quality could frustrate peoples initial interests and motivation to use the tool. However, very positive feedback was received from the use that could be made of the tool, including among others: invitations to give workshops and to replicate the project in other cities, a MoU with São Paulo public administration and nationwide recognition given by Brazilian media and Google Global Impact Challenge-Brazil prize. Due to these results, we conclude that budget geo-location is a good strategy to make difficult and/or abstract budget concepts, data and information more accessible to a wider audience.

The project interviews and outputs raise a series of new questions. First of all, the civil society ecosystem and the quality of its components and connections may counterbalance poor data quality supplied by governments. Results suggested the importance of better understanding the role of intermediaries and consumers, as well as their context. Besides the heterogeneity inside each category, their role can interchange depending on the situation, disrupting the static notion of passive-active relationship. In order to understand how actions and impacts are achieved is also important to understand which are the steps of data transformation, information and knowledge generation carried out through interaction of particular consumers and intermediaries. Other important findings come from the supply side of public sector information, represented through an iceberg metaphor. The visible published open data masks deeper underlying resource, process and cultural issues and constraints which potentially make open data initiatives unsustainable. On the other hand, moving towards open data can lead to changes in these underlying and hidden process, leading to shifts in the way government handles its own data.

Conclusions

The project has completed its first cycle of development with some important accomplishments and lessons learned both from the demand side (civil society organizations, academia and hacker activist) and the supply side. Using an action research methodology, the authors have worked since data collection and tool development stages: presenting and discussing budget data and public policies with civil society organizations and grassroots movements. In this scenario we could see that open data clearly disrupted the Data-Information-Knowledge value relationship, where the value of Data-in-use can be as important as the value of knowledge. Traditionally, consumers are presented with information (statistics) and knowledge, but increasingly consumers or intermediaries want to get access to the data at the core. More research work has to be done to fully address the issue of open data impact assessment methodologies, in particular, focusing on the intermediary ecosystem, rather than starting from data and tracing down.

Figure 1: Geolocated Budget Visualization with the help of the tool Cuidando-do Meu Bairro (Caring for My Neighborhood). Available at: cuidando.org.br

Figure 2: Percentage of expenses in relation to elements that allow inference of location

Figure 3: Amount of resources (in billions of Brazilian Reais) in relation to the values of activities and also the existence of information to identify the localization.

Figure 4 and 5: Models of the civil society ecosystem interaction and transformations from data to knowledge.

Evidence and findings

Despite the fact that budget data provided by São Paulo municipal administration has no lat./lon. coordinates, the Cuidando do Meu Bairro tool searches for textual elements that can give some detail about the location of expenditure and generates a geo-coded dataset. However, many budget activities are described in an aggregated/generic way. This prevents the Cuidando do Meu Bairro being presented as a tool to track and monitor public budget implementation in full.

Insufficient data quality could frustrate peoples initial interests and motivation to use the tool. However, very positive feedback was received from the use that could be made of the tool, including among others: invitations to give workshops and to replicate the project in other cities, a MoU with São Paulo public administration and nationwide recognition given by Brazilian media and Google Global Impact Challenge-Brazil prize. Due to these results, we conclude that budget geo-location is a good strategy to make difficult and/or abstract budget concepts, data and information more accessible to a wider audience.

The project interviews and outputs raise a series of new questions. First of all, the civil society ecosystem and the quality of its components and connections may counterbalance poor data quality supplied by governments. Results suggested the importance of better understanding the role of intermediaries and consumers, as well as their context. Besides the heterogeneity inside each category, their role can interchange depending on the situation, disrupting the static notion of passive-active relationship. In order to understand how actions and impacts are achieved is also important to understand which are the steps of data transformation, information and knowledge generation carried out through interaction of particular consumers and intermediaries. Other important findings come from the supply side of public sector information, represented through an iceberg metaphor. The visible published open data masks deeper underlying resource, process and cultural issues and constraints which potentially make open data initiatives unsustainable. On the other hand, moving towards open data can lead to changes in these underlying and hidden process, leading to shifts in the way government handles its own data.