

ABSTRACT

OPEN LEGISLATIVE DATA AND ITS IMPACT ON LOBBYING

Felipe Lélis Moreira¹

The contemporary legislative process, in the context of the information society and the rule of law, creates a new task to Parliaments: the need to communicate efficiently not only with human beings, but also with machines. Today, transparency is not limited to the mere provision of information on the internet. It is necessary to generate data in open formats, consumable by computers in order to be easily reused in digital applications.

Corporations have been using algorithms and big data to solve some of their biggest challenges. Those tools are now starting to catch on in lobbying. Several technologies collect data from many sources and compile it in a searchable database where it is possible to add an analytical layer.

Therefore, legislative documents already are being published in open standards, and once that happens more regularly and in more expected formats, the same techniques of statistical analysis, predictive modeling, and deep data-driven insights into what the law is and how it relates to business, social and economic systems can be visualized and acted upon.

The aim of the paper is to show how the legislative openness is impacting the decision-making process and which are the challenges that Parliaments will need to face to reduce inequalities related to the participation of different social actors.

First, we are going to introduce the data science and the concept of Big data². Then, we are going to show how the lawmaking process can be looked as a legislative

¹Lawyer, legal adviser and government relations specialist. Phd candidate and master in law at the Federal University of Minas Gerais (UFMG). Post graduate specialist in public policy, government management and public law. E-mail: flelismoreira@gmail.com

² Big data has been defined as the conjugation of four Vs: 1) Volume. The amount of data. While volume indicates more data, it is the granular nature of the data that is unique. Big data requires processing high volumes of low-density, unstructured Hadoop data—that is, data of unknown value, such as Twitter data feeds, click streams on a web page and a mobile app, network traffic, sensor-enabled equipment capturing data at the speed of light, and many more. It is the task of big data to convert such Hadoop data into valuable information. For some organizations, this might be tens of terabytes, for others it may be hundreds of

(big) data process, where solutions in process automation with the use of artificial intelligence and data analysis can be used for data-driven decisions. Therefore, we are going to demonstrate that the fuel of those technologies is government data, which needs to be made available in “open formats”.³

Secondly, we intend to show that some corporations have been using the data science for lobbying purposes. It is important to note that Lobbying is not a crime. Contrariwise, it is a fundamental right granted by all the democratic constitutions. Everyone lobbies when legally attempt to influence a decision of a public authority. While the face of lobbying is often a polished government relations executive trekking the halls of Parliaments armed with talking points, attending luncheons, and writing op-eds, the hidden side of the business entails hours of research and hard work. Despite the billions of dollars that corporations pour into lobbying efforts each year, the labor has remained relatively low-tech. Part of the problem is knowing how to sift through reams of information. Some firms have been using algorithms and big data to solve some of their biggest challenges and those tools are now starting to catch on in lobbying. We intent to

petabytes; 2) Velocity. The fast rate at which data is received and perhaps acted upon. The highest velocity data normally streams directly into memory versus being written to disk. Some Internet of Things (IoT) applications have health and safety ramifications that require real-time evaluation and action. Other internet-enabled smart products operate in real time or near real time. For example, consumer eCommerce applications seek to combine mobile device location and personal preferences to make time-sensitive marketing offers. Operationally, mobile application experiences have large user populations, increased network traffic, and the expectation for immediate response. 3) Variety. New unstructured data types. Unstructured and semi-structured data types, such as text, audio, and video require additional processing to both derive meaning and the supporting metadata. Once understood, unstructured data has many of the same requirements as structured data, such as summarization, lineage, auditability, and privacy. Further complexity arises when data from a known source changes without notice. Frequent or real-time schema changes are an enormous burden for both transaction and analytical environments. 4) Value. Data has intrinsic value—but it must be discovered. There are a range of quantitative and investigative techniques to derive value from data—from discovering a consumer preference or sentiment, to making a relevant offer by location, or for identifying a piece of equipment that is about to fail. The technological breakthrough is that the cost of data storage and compute has exponentially decreased, thus providing an abundance of data from which statistical analysis on the entire data set versus previously only sample. The technological breakthrough makes much more accurate and precise decisions possible. However, finding value also requires new discovery processes involving clever and insightful analysts, business users, and executives. The real big data challenge is a human one, which is learning to ask the right questions, recognizing patterns, making informed assumptions, and predicting behavior.

³ The most important characteristics of an open data are: 1) Availability and Access: the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. Also, the data must also be available in a convenient and modifiable form. 2) Re-use and Redistribution: the data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets; and 3) Universal Participation: everyone must be able to use, re-use and redistribute - there should be no discrimination against fields of endeavour or against persons or groups. For example, ‘non-commercial’ restrictions that would prevent ‘commercial’ use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

demonstrate how some of these tools works and how ordinary people, public leaders and civil society organizations could design methodologies to improve its lobbying activities.

To conclude, we will demonstrate that, few high-cost firms offers technologies for lobbying. Besides this, Data Science is an area dominated by few social actors, notably the most economically privileged. Therefore, the mere disclosure of open legislative data can contribute to the accentuation of social inequality. Thus, Parliaments needs to rethink its openness values acting as an inducer of the co-creation of technologies that can help to solve inequalities on decision-making processes.

As highlighted by Seamus Kraft⁴ “to build better government, we may just need better lobbyists”. This quote, which may be seen as extremely simplistic at first, gains prominence when lobbying is viewed as a practice of political participation on building more efficient and effective State decisions. On the other hand, talking about “better lobbyists” may be the same as talking about high-quality civic participation, broad, conscious and transparent, unrelated to the vicissitudes and crimes that sometimes ends up contaminating the legislative process.

⁴ KRAFT, Seamus. Do we need a lot more lobbyists? Available in: <<http://opengovfoundation.org/do-we-need-a-lot-more-lobbyists/>>.