TECHNICAL KNOWLEDGE NOTE

“We Feel Fine”: Big Data Observations of Citizen Sentiment About State Institutions and Social Inclusion

OVERVIEW
Motivated by the significant decline in citizen’s trust in governments over the past decades, this paper explores how policy decision makers and researchers can use social media analytics to investigate “trust”, specifically the relationship among trust in government, trust in state institutions and citizens’ collective behavior. Analysis of these complex socio-political issues using online social data requires a human in the inference loop while also benefiting from computational methods to handle large amounts of unstructured data and the inference of relevant data features. This project implemented an exploratory framework from data collection to data analysis using, among other methods, a set of tools developed by project team members that integrate visualization, interactivity and machine learning. Visual analytics provided the team with knowledge of interactive methods and investigative methodological approaches while predictive modeling provided methods to operationalize difficult concepts such as “trust” through a proxy of sentiment classification of opinion in social media data. To highlight the power of a mixed-initiative visual analytics-data science approach, this technical note describes the exploratory analysis work undertaken for analysis of collections of Tweets from Brazil, and describes further work that conceives data science methods to assist the analysis process by supporting definition of constructs of concepts of interest using social media data, and assisting the evaluation of evidence for hypotheses evaluation in an interactive-machine learning fashion. The outcomes of this project aim to support social sciences inquiry using observational social media data and World Bank operations.
BACKGROUND
According to the OECD, trust in government is an important foundation for legitimacy and sustainability of political systems, and is essential for social cohesion and wellbeing as it affects governments’ ability to govern, as it enables them to act without having to resort to coercion. Motivated by the significant decline of citizen’s trust in governments over the past decades, this paper explores how researchers and policy makers can use social media analytics to contribute to research on trust, specifically the relationship among trust in government, trust in state institutions and citizens’ collective behavior. Analysis of these complex socio-political issues inherently requires a human in the inference loop while it can also benefit from methods to handle large amounts of unstructured data (e.g., data science machine learning methods).

Data Science is defined as the study of generalizable extraction of knowledge from data with special emphasis in the power of prediction. In this paper we highlight the potential use of data science methods, particularly machine learning, to aid visual exploratory analysis of large data in the context of complex socio-political constructs. Visual analytics (VA) provides a rich toolset of interactive methods and investigative methodological approaches while predictive modelling provides methods to operationalize difficult concepts such as “trust” through a proxy of sentiment classification of online opinion in social media data.

Predictive modelling is concerned with the use of algorithms, leveraging statistical analysis, to train a model that has “learned” from the data and so, is able to make predictions or decisions about new, unknown observations, without following explicitly programmed instructions. Typically, the event to predict is in the “future”, based on learning from past data, but predictive modelling can be applied to any type of unknown event regardless of when it occurred (e.g. can we predict the sentiment valence of an opinion based on learning from other opinions?).

From the start of the project, the research team noted potential paths of development for a mixed-initiative approach where it is necessary to combine visualization, interaction and machine learning methods. An important principle of data science is that it is a process with fairly well-understood stages. Within this process, literature from both sides, Visual Analytics (VA) and Data Science acknowledges the cooperative roles of the human and the machine: Provost and Fawcett, argue that the process involves stages of application of information technology, such as the automated discovery and evaluation of patterns from data, and stages that mostly require the analyst’s business knowledge, creativity and common sense. Ribarsky et al. emphasize making “visually enabled tools”, coupling visualizations and interactions with the human visual/understanding channel for maximum throughput integrated with human understanding and judgment. Interaction is the coupling mechanism between human reasoning and analytic processes, on the one hand, and computer-based processes on the other hand. Investigative analysis such as the one conducted to understand social media data in the context of political trust theories, reflect a natural scenario for mixed-initiative systems where there is a need for human and computer to work together because the analysis involves constructing meaning from data that requires a human ability to create or extend hypotheses, yet at the same time, the data are too large, the dimensions too high, and the ramifications of a change or decision too many for a human to handle unaided.

This note describes exploratory work undertaken to analyze Tweets from Brazil using visualization of sentiment analysis, and two analytical frameworks from the VA and intelligent analysis literature: Pair Analysis and Analysis of Competing Hypotheses (ACH). The work draws upon public administration and political science literature relating to trust in government, as well as upon literature on social protest theory from social psychology and sociology. In this work, the construct of trust is operationalized as a

CONTACT
Victoria Lemieux, Senior Public Sector Specialist, Integrated Digital Solutions Group, vlemieux@worldbank.org
measure of the sentiment expressed in the content of Twitter data. The study involved developing a tool for and conducting a visual analysis of sentiment classified Twitter data to derive insights about the following three research questions: 1) how did citizens feel about their state institutions around the time of the 2014 World Cup protests, 2) how did these feelings connect to their sentiments about Brazilian Federal and State Government and politicians and 3) how did such sentiments translate into collective behaviors (e.g., social protests).

**METHODOLOGY**

The methodology for this study involved a series of choices. One of the first, and most significant, of these was the choice about how to operationalize the construct of trust. Trust has been a notoriously difficult construct to measure, with much debate and many conflicting methodological approaches. In this study, sentiment, defined as affective valence, positive or negative, was used as a proxy measure of trust. Sentiment analysis for the social web has become a useful tool for analysis of the role of emotion in online communication, public opinion and reaction to events via public texts. The use of sentiment to measure trust is also common in the field of marketing (see, for example, Boertjes on consumer trust in a cloud computing service model with textually expressed sentiment). The approach has not been used previously in studies of trust in government. Nevertheless, relying upon the use of feelings or emotions in measuring political trust is not without precedent. Feelings about politicians, institutions and groups has been used to measure political trust by Hetherington as well as an American National Election Studies survey of trust in government (ANES n.d.). Newton and Norris also link the notion of public trust to feelings. In light of these linkages, there is some justification for relying upon sentiment to operationalize the measure of trust.

Through exploratory analysis of the study aimed to track trust in institutions and general public distress, hypothesizing that an increase in negative sentiment signaled declining public trust in government. Exploratory analysis of text can be done using a toolset of options ranging from manually reading hundreds of documents to selectively querying a database. However, the reading intensity and the time cost of these options are too high. Especially when working with thousands of short texts and trying to make sense of groupings of them. Visual Analytics is a robust approach for these types of tasks. “The science of analytical reasoning facilitated by interactive visual interfaces” is an analytical approach that accounts for the need of a human-in-the-loop in the analytic sense-making process.

The methodology used in this study therefore comprised the following steps:

1. Operationalizing trust as a measure of sentiment in the content of Twitter data
2. Instrumenting trust using a sentiment classifier
3. Initial “big picture” harvest of Twitter data
4. Visual iterative exploratory analysis and text analysis (Natural Language Processing) of “big picture” twitter data to extract search terms representing the key concepts underlying our study i.e., state institutions, government, etc.
5. Use of search terms to harvest historical Twitter data for our period of interest i.e., 2014 World Cup in Brazil
6. Sentiment Classification of the harvested tweets
8. Pair analysis for exploration of data.
9. Analysis of competing hypothesis aided by our VA tool to explore the historical Twitter collections and derive findings.

To instrument the approach to measuring trust, the research team chose to use a sentiment classifier, SentiStrength, as our sentiment analysis tool considering 1) successful evaluation results reported for analysis of short texts, 2) the capability to customize a domain-specific sentiment
lexicon, and more importantly 3) the capability to customize language even without labelled training data. Research and open source initiatives have produced alternatives with good results but mainly for analysis of English language, thus, not fitting this study’s main need.

Despite identifying sentiment polarities as sufficient for many applications, texts often contain a mix of positive and negative sentiment and for some applications it is necessary to detect both simultaneously and also to detect the strength of sentiment expressed. SentiStrength accounts for this, classifying text with both a positive and a negative magnitude ranging between 1 and 5. SentiStrength is a lexicon-based classifier that uses custom dictionaries of words to identify sentiment. Besides the sentiment words, the algorithm includes booster words, negating words, question words, emoticons, and idioms if specified. The customization of the Brazilian-Portuguese dictionaries was carried out by two native speakers (inter-raters of sentiment weight for the words) having as a base dictionary, the results of the translated and adapted Affective Norms of English Words (ANEW) for Brazilian Portuguese. ANEW-BR is a set of 1,046 words with emotionality measures that were mapped to a scale of sentiment magnitude.

For the study, publicly available Twitter data was collected in two phases. In the first phase, researchers collected data in real time using STACKS, an open source, social media data collection toolkit built and maintained by project members. The first collection was bounded with a single location filter that fit roughly around the entirety of Brazil. It is important to note that only a small percentage of tweets (approximately 1 percent) contain geolocation data, so this initial collection was intended to sample a large “universe” of tweets in order to conduct an exploratory analysis of topics of discussion. Researchers implemented an iterative inductive approach to narrow down the stream of data and build up collections defined by topical themes of interest. This “big picture” approach was used in order to capture ways in which citizens communicate or discuss themes, rather than constructing topics only from dictionary terms and definitions. Ultimately, this approach was used to determine the search phrases that would more naturalistically capture topics of interest - government service institutions and citizen opinion about government to track indicators of trust – within a body of Tweets, thereby avoiding imposing the researchers’ own views on the topics of importance and the terms used to describe these topics.

Data collected and analyzed in the first phase informed the selection of search terms employed in the second phase. For the second phase, historic data was purchased from DiscoverText using a tool called Sifter, which is a reseller of data from Gnip, a subsidiary of Twitter.

For the analysis of the universe or “big picture” Tweets the researchers used IN-SPIRE text analysis tool, and n-gram based text categorization using the NLTK package to implement the analysis in Python. IN-SPIRE is a visual analysis tool for the analysis of large sets of unformatted text documents by clustering similar documents together, determining key topics or themes in each document within the collection.

IN-SPIRE’s galaxy view gave researchers a rough classification of easily identifiable popular culture themes such as TV shows, sports games, and mainstream actors and singers. By iteratively filtering out documents associated to these themes, which were not relevant to the focus of the study, political opinion commentary related to President Dilma Rousseff and Petrobras began to appear. In order to identify other documents relevant to the study researchers searched (and collected a list of) general terms of services and service institutions such as: educação, serviços, saúde, hospital, polícia, faculdade. The researchers extended the list of words people use for political commentary and to refer to institutions by filtering, searching and reading individual documents. They found commentary around the drought crisis in some states as a source of service delivery opinion.
IN-SPIRE helped the researchers to read the documents in a guided exploratory way, and to construct a list of terms to investigate further. However, due to constraints about the length of Tweets, language, and character encoding, the researchers were not able to use IN-SPIRE to define the search phrases. Instead, they used the list of seed words to perform word co-occurrence analysis. This way, they were able to identify other words used to talk about the water crisis, the current Petrobras events, references to service institutions, or general political opinion. The research team also performed bi-grams and tri-grams analysis that corroborated most of the phrases found by co-occurrence with a defined list of words.

In this project, the exploration of social media data to understand trust, consisted of a set of exploratory and hypothesis testing tasks that required a structured way to assess the text-data. The team developed a custom tool (see a snapshot in Figure 1) to facilitate the exploration with special emphasis on visualization of sentiment analysis patterns. This tool allows an analyst to:

1. Explore aggregated overviews of temporal patterns of Tweets with the ability to overlay external events flags.
2. Explore the Sentiment trends using horizon charts
3. Obtain an overview of the text content with a text cloud and line chart of individual words frequency
4. Read details of individual Tweets to understand digital opinion patterns that may not be obvious for an algorithm or from an aggregated overview of the data.
5. Filter Tweets list by search delimits (words or dates)
6. Sort data by sentiment, date and count of ReTweets if exploring the de-duplicated list of Tweets.

3.2 Pair Analysis and Analysis of Competing Hypotheses

Having developed the analytic tool and gathered relevant data, the next step involved conducting a visual analysis of the sentiment classified data. The analysis began with a simple pair analysis session aimed at making basic observations from the data about the study’s three main research questions. Subsequent analysis was informed by Heuer’s Analysis of Competing Hypotheses (ACH) methodology. ACH “is an eight-step procedure grounded in basic insights from cognitive psychology, decision analysis, and the scientific method... [it is] a proven process that helps analysts avoid common analytic pitfalls. Because of its thoroughness, it is particularly appropriate for controversial issues when analysts want to leave an audit trail to show what they considered and how they arrived at their judgment.”

Thus, the analysis proceeded with:
1. Identifying the possible hypotheses to be considered.
2. Making a list (matrix) of the hypotheses, and identifying evidence for and against each one.
3. Analyzing the strength of each hypothesis.
4. Iterative refinement of the analysis, deleting evidence and arguments that have no diagnostic value.
5. Drawing tentative conclusions about the relative likelihood of each hypothesis by trying to disprove the hypotheses rather than prove them.

For the analysis of competing hypotheses (ACH), the researchers drew upon the background literature to identify 68 hypotheses for consideration. The hypotheses can be divided into two groups: the first group derives from the literature on trust in government and the second group derives from the literature on social protest. Each hypothesis was considered in terms of the evidence in support of it, evidence against it, and the strength of the evidence in terms of diagnosticity and within sample representativeness of the evidence. The evaluation of the strength of the evidence for each hypothesis in the ACH analysis was based on a qualitative assessment.
The completed analysis offers the strongest evidence in favor of the relative deprivation theory of social protest coupled with notions of public trust that involve citizens’ assessment of government’s ability to follow through on promises, deliver on policy objectives and act with integrity. This was Hypothesis 48: “Feelings of relative deprivation result from comparison of one’s situation with a standard – be it one’s past, someone else’s situation, or a cognitive standard such as equity or justice.”

In social theory, relative deprivation is the notion that an individual or group lack something that another group has and to which they feel some entitlement. The important emphasis here is that they feel this deprivation in relation to some standard, whether static or emergent as in the case of this study (i.e., the 2014 World Cup) or other group. Evidence from this collection of Tweets indicates that it is not performance or outcomes per se, but performance and outcomes for one group (“povo Brasileiros”) relative to another (spending on campaign financing/World Cup). Many of the negative Tweets in the collection (and a good proportion of the misclassified positive Tweets) express strong sentiment (e.g., anger, disgust, discontent and disgruntlement) at the spending on the World Cup in relation to public services. Two public services in particular are the focus of Tweets in the political opinion collection: education and health, with water being mentioned most in the service delivery collection. In the public service delivery collection, the water shortage is mentioned more prominently, and there is less focus on relative deprivation in relation to spending on the World Cup and more on the behavior of state level politicians (e.g., excessive spending on campaign financing) and SABESP, the water authority. It was this sense of relative deprivation that seems most evident during periods exhibiting spikes of heightened sentiment. This theory also explains the wide variety of issues that surfaced in the Tweets, from water, to education, health, police, transport, threat of increased taxes, and inflation – all contributed to the sense of relative deprivation. In short, growing distrust (negative sentiment) was driven by people’s sense of a growing distance between the “haves” and the “have-nots.” This was also coupled with the perception that corrupt and dishonest politicians were personally benefitting, or at least not suffering like everyone else, and not delivering on promises to the people about spending, which appears to have had an amplifying emotional effect.

CONCLUSION
Exploration and hypotheses testing using Twitter data with sentiment opinion trends provide an example of a visual exploratory analysis that benefits from integrating statistical learning modeling. The sentiment analysis classifier instruments a very relevant feature within the data, establishing a framework of reference to monitor trends and
investigate the status of a specific problem. This type of problem sits in the middle of traditional “exploratory” versus “predictive” analytics space because despite the fact that the tasks at hand are exploratory, handling the volume of data or computing additional features relies on computational modeling. In this project, exploration is not only the initial stage on a linear analysis pipeline, but rather a set of cognitive tasks that structure the access to, manipulation and transformation of the data in order to make sense of the “current state of events” and to systematically test hypotheses and search for evidence. Traditional methods for analysis of social media data such as topic modelling, sentiment analysis, and social network analysis set a robust framework for investigation.

ABOUT THE GOVERNANCE GLOBAL PRACTICE
The Governance Global Practice of the World Bank supports countries in their efforts to build trustworthy institutions that can contribute to making development possible by delivering results for people. We understand the challenges facing our country clients and in particular, government officials and public sector professionals. We aim to provide the foundational support and innovative solutions through a range of financial, analytical, advisory, fiduciary and capacity building products that can help governments deliver better on development goals in every field, from social protection to infrastructure, business development and the environment.

5 Op Cit.