Citizen Trust in the United States Government: Twitter Analytics Measuring Trust in Government Sentiments

Glenn Papp  
*Niagara University*

Omar F. El-Gayar  
*Dakota State University*

Petter Lovaas  
*Niagara University*

Follow this and additional works at: [https://scholar.dsu.edu/bispapers](https://scholar.dsu.edu/bispapers)

**Recommended Citation**

[https://aisel.aisnet.org/amcis2020/social_computing/social_computing/13](https://aisel.aisnet.org/amcis2020/social_computing/social_computing/13)
Citizen Trust in the United States Government: Twitter Analytics Measuring Trust in Government Sentiments

Glenn Papp  
*Niagara University*, grp@niagara.edu

Omar El-Gayar  
*Dakota State University*, omar.el-gayar@dsu.edu

Petter Lovaas  
*Niagara University*, plovaas@niagara.edu

Follow this and additional works at: [https://aisel.aisnet.org/amcis2020](https://aisel.aisnet.org/amcis2020)
Citizen Trust in the United States Government: Twitter Analytics Measuring Trust in Government Sentiments

Completed Research

Glenn Papp Jr.
Niagara University
grp@niagara.edu

Omar El-Gayar
Dakota State University
omar.el-gayar@dsu.edu

Petter Lovaas
Niagara University
plovaas@niagara.edu

Abstract

Recent tensions and widening division in the United States (U.S.) political arena have sewn doubt in the U.S. government and democracy itself, at a time when rapid technology advances clearly ought to aid the spread of democracy. However, there has been little progress in the use of technology to increase democratic participation in the U.S. The purpose of this research is the exploratory study of how social media analytics can inform, predict, or alter trust in government sentiments which thereby inform democratic participation by studying 49,964,168 Twitter posts (tweets) from January 1, 2014 – June 13, 2019. Extant literature pertaining to E-Government research is reviewed and a theoretical framework is presented as basis for the methodology. The findings shed insight towards the workings of public trust sentiments and current events and/or governmental actions. Additionally, discussions of results pose questions for testing and evaluation in future research.

Keywords

Trust in government, trust sentiments, Twitter analytics.

Introduction

Existing research shows correlations between well-implemented government e-Services and improvement of internal relations of public administrations, relationships between administrations and citizens, and relationships between administrations and the private sector (Alcaide-Muñoz & Rodríguez Bolívar 2015; Ayyad 2017; Olphert & Damodaran 2007; Tassabehji et al. 2007). Nonetheless, the current political environment in the United States is reflective of the state of discourse in U.S. politics, in that norms and institutional integrity are being sacrificed for political party loyalty. The result is a toxic dialect, near-impossible difficulty of bipartisan collaboration, and the perception of a broken system. At the same time, however, U.S. citizens are seeing radical and expedient advances in information and communication technology (ICT) while government seemingly struggles to keep up. An exhaustive review of existing e-government research publications listed in ISI from the period of 2000-2014 was conducted and resulted in a call for non-qualitative methodologies in this field of research, as 61.98% of e-government research found used qualitative methods (Alcaide-Muñoz & Rodríguez Bolívar 2015). Furthermore, researchers found that 62.96% of existing e-government research found used case study methodologies. Accordingly, quantitative, design science research, and other hybrid methodologies are warranted in this area.

The purpose of this research is the exploratory examination of how social media analytics (Twitter, specifically) can be leveraged to identify, predict, and/or drive trust in government sentiments, with which extant literature shows positive correlations with democratic participation (Agawu 2017; Aladwani & Dwivedi 2018; Calderon et al. 2015; Hapsara 2016; Lappas et al. 2015; Maciel et al. 2016; Mossberger et al. 2013; Thomas 1998). Specifically, the research question seeks to determine which trust sentiment and perceptions exist from the past five years towards voting, attending political events, donating to political
causes, individual candidates, and the entire political system. The expected contribution of this research is to inform future research into leveraging social media analytics for gauging public sentiment to thereby inform democratic participation. Due to the medium of data being measured in this study (public digital communications via Twitter), the insights towards democratic participation ought also to inform E-Government need, and accordingly, design and implementation requirements. Relevant theories and models are used in empirical support for the study, and a theoretical contribution is made regarding the taxonomy of components of trust in government.

The remainder of the paper is organized as follows: the next section provides a literature review and theoretical background followed by a detailed description of the methodology, including data collection and corresponding analytics. The results section summarizes the findings from an analytical point of view with the aim to evaluate trust sentiment towards the political environment in general to inform future research. The fifth section presents the custom categories and their respective results and implications to future research. Limitations are addressed and a discussion of the implications of this research on future efforts is had in the sixth section, and the paper concludes with a summary of findings and contributions.

**Literature Review and Background**

E-Government, arguably the most general of the several other terms explaining the use of technology in government, is defined here as “...the delivery of [government] information and services online through the Internet or other digital means” (Agawu 2017, p. 3). E-Government is divided in to four categories: (1) government-to-citizen (G2C), (2) government-to-business (G2B), (3) government-to-government (G2G), and (4) government-to-employee (G2E). Following research conducted by Agawu (2017), this research will only also be concerned with the first category, G2C. Agawu (2017) additionally proposes three trends of E-Government activity: (1) increasing access to content, (2) digitization of the service loop, and (3) expanding or creating new government functions. A quick comparison of the United States Federal Government’s E-Government initiatives with these trends as a metric against other countries' national governments around the globe clearly demonstrates the need for progress (Sundberg 2019).

While Agawu (2017) researches the government-to-citizen (G2C) tenet of E-Government through the lens of security and status, the research purpose here demands another set of principles to more accurately guide the methodology. Thomas (1998) reviews the three conceptions of trust—fiduciary trust, mutual trust, and social trust—and connects them to the production of trust. Tassabehji et al. (2007) furthers this effort and uses modes of trust production in the development of a trust verification agent (TVA) artifact used to generate citizen trust in adopting E-Government platforms. These modes of trust production are a central piece of this research’s theoretical backing as well, and they are as follows: (1) characteristic-based trust, tied to personal characteristics that are difficult or impossible to change; (2) process-based trust, tied to reciprocity with exchanges of equal intrinsic or economic value; and (3) institutional-based trust, which can be achieved either by individuals or the organization entirely, and/or administration of laws, regulations, insurance, and other practices (Tassabehji et al 2007; Thomas 1998). These two frameworks represent this research’s theoretical basis for the trust in government concept.

Another component of the theoretical framework used in this research adopts components of SocioCitizenry theory by Aladwani & Dwivedi (2018), whereby approved adaptation, caused by improved knowledge and enhanced relatedness, is predicted by both trust configuration (split in to both content-based and engagement-based trust) and quality anticipation (determined by content meaningfulness and interaction engagingness). As this research is only looking at trust, we will not define the unrelated tenets in detail. Regarding the related trust production mode, according to Aladwani & Dwivedi (2018), content-based trust is defined as “...one’s perceptions regarding the congruence between the content attributes of governmental social media services and one’s trust needs,” while engagement-based trust is defined as “...one’s perceptions of the match between the engagement attributes of governmental social media services and one’s trust needs” (p. 263). These two definitions, although accurate, are similar at their core and conceptually seem to have some overlap with the previously discussed characteristic-based and process-based trust components, respectfully. Accordingly, this research will combine the two definitions and concepts both theoretically and practically.

Finally, as the United States Federal Government currently has no digital government platform allowing citizen-initiated democratic participation, it is thought appropriate to use a voting framework to most
closely represent the concept of democratic participation. Robertson (2005) provides a framework for contexts of voting which specifies the action of voting as counting least towards civic engagement, followed by decision making (including comparison tools, decision making aids, and record keeping), information gathering (including media dialog, targeted messages, and conversation), and finally culture and belief (including life experiences, party affiliations, and special interests). In this context, given the previous justification for combining characteristic-based trust with content-based trust and combining process-based trust with engagement-based trust, it appears the first combination is more closely associated with culture and belief, while the second combination is more closely associated with information gathering as a process. The third mode of trust production, institutional-based trust, also appears to bode well alongside decision making, as the institutions are often the only entities permitted to enact and enforce laws and regulations.

Figure 1. Theoretical Framework

As shown in Figure 1, it is the intent of this research to define three new trust production modes with the support of extant literature in this field. In assessing the combination of definitional concepts of characteristic-based trust, context-based trust, and culture and belief, this research will adopt the mode of behavioral trust, defined as one’s perceptions regarding the alignment between the expected behavior and actual behavior of the government service(s). Accordingly, in assessing the combination of process-based trust, engagement-based trust, and information gathering, this research will adopt the mode of operational trust, defined as one’s match between the operational attributes of government service(s). Finally, the Thomas (1998) and Tassabehji et al. (2007) definitions of institutional-based trust will be used for what is coined in this research as institutional trust, as previously justified.

Multitudes of extant literature have either used social media analytics to study a research problem, or call for such methodologies (Calderon et al. 2015; El-Gayar et al. 2019; Jaidka & Ahmed 2015; Jamal et al. 2015; Lappas et al. 2015; Lin 2018; Mossberger et al. 2013; Sobaci & İ 2017). Although some literature in this realm focuses the security of such initiatives (Agawu 2017; Sundberg 2019), most research in this area is focused on problems outside the scope of E-Government (Calderon et al. 2015; El-Gayar et al. 2019; Jamal et al. 2015) or conducted outside of the United States (Jaidka & Ahmed 2015; Sobaci & İ 2017). Accordingly, it is the aim of the publication of this research to initiate a new sect of E-Government research within the United States focused on leveraging social media analytics to help drive E-Government initiatives.

Research Design and Methodology

To study citizens’ perceptions and sentiment regarding trust in government, we leveraged Crimson Hexagon, a social media analytics platform for data collection and analysis (Brandwatch 2020). Crimson Hexagon (CH)’s Forsight platform uses a supervised and/or unsupervised text analysis algorithm called
ReadMe developed by Daniel Hopkins and Gary King (Hopkins & King 2010). Typically, research with this tool begins with a targeted, specific question. The one employed in this study is as follows: which trust sentiment and perceptions exist in the United States from the past five years towards voting, attending political events, donating to political causes, individual candidates, and the entire political system? This specific research question is then articulated into a Boolean query bound by certain filters such as date range, data sources, keywords for inclusion and exclusion, language, and geographical restrictions, etc., to accumulate the data. Next, the data collected is analyzed, either with predetermined categories for sentiment and/or opinion analysis (supervised via the buzz monitor), or custom categories manually trained by the user (unsupervised via the opinion monitor).

**Data Collection**

As shown in Figure 2, tweets were collected that contained trust or its synonym with a mention of a political figure, party, process, or institution while excluding spam, retweets, and governmental social media accounts. Where S represents any sentiment or action encompassing sentiment, T representing a target of that sentiment (in this case limited to political ideology, political figures, and governmental institutions relevant to the 2014, 2016, 2018 or 2020 elections), and I represents irrelevant or biased tweets and/or accounts (i.e., in trying to measure citizen sentiment, government accounts should not be included; retweets and external links should not be counted per CH training materials), the formula used in creating the query in Figure 2 is as follows: (S + T) – I. Data sources were limited to Twitter only, language was limited to English, and location was limited to the United States. The data accumulated was limited to the date range of January 1st, 2014 (Jan-01-2014), to June 13th, 2019 (Jun-13-2019).

![Figure 2. Search query](image-url)

Following the example of Hopkins & King (2010), for which they use ‘George W.,’ ‘Dubya,’ and ‘King George’ to capture all blog posts about President George W. Bush (p. 232), a similar approach was taken in the creation of this query where the only other qualifying terms used outside of first and last name are those that do not explicitly include the first or last name of the target, T, but are known to be used to identify T (i.e. ‘Drumpf’ and ‘make America great again’ for President Donald J. Trump, ‘HRC’ and ‘I’m With Her’ for 2016 Democratic Presidential Candidate Hillary Clinton). A similar study was conducted leveraging Twitter post analytics in measuring trust in government sentiments; however, the similar study gauged the attitude of citizens in response to government use of social media, meaning the query was limited to only government Twitter handles or direct responses to those handles (Franks 2016, p. 13). Although the approach used in the similar study is systematic and targeted, the methodology and research question in this study are novel in that the data collection and analyses exclude direct communications from official
government handles, and instead specifically target sentiment expression towards or about government figures, ideologies, and institutions directly from citizens.

For the unsupervised opinion monitor, the definition of custom categories and manually training at least ten posts as related to each category is required. According to CH guidelines, it is crucial that (1) posts that are trained do not overlap with other categories as the trained tool can break apart multiple custom categories within the same tweet, and (2) each category is trained with a minimum of 25 tweets for larger data sets (Brandwatch 2020). Table 1 shows the categories adopted for this study, the number of posts trained for each, and an example tweet representative of the topic and sentiment the text analysis model will identify for that category. Four groups were established: behavioral trust, operational trust, institutional trust, and irrelevant. Within the first three groups, two categories were created: positive and negative. The irrelevant group had one category titled off-topic.

Three of the seven categories denoted above were found to be relatively difficult to train for: operational trust positive, operational trust negative, and institutional trust positive. Although the volume of raw tweets approaches 50 million, the training module of CH does not show the user for every single tweet—instead, groups of tweets are represented by a singular, exemplar tweet which echoes the sentiment of dozens, hundreds, or thousands of similar tweets. After initially parsing through the data with targeted second-level queries that were thought to produce tweets exclusive to individual categories, it was quickly discovered that some categories were much easier to find several of these exemplar tweets within the training module, and some were more difficult. For example, in assessing the behavioral trust positive and negative categories, any person who has used or uses Twitter will not be surprised at the massive volume of tweets discovered that encompass sentiments regarding individuals in government (i.e. “I love/hate [insert political figure]”). However, it is much more difficult to think of a Twitter user who is motivated to tweet solely about positive and/or negative sentiments regarding an institution without also mentioning an individual (behavioral trust) or a process (operational trust). When presented with this issue, and considering CH guidelines, it was thought best to keep the purity of tweets’ relation to their respective categories intact rather than to inflate the number of trained posts and concurrently tainting the categories’ training data as needed.

<table>
<thead>
<tr>
<th>Category</th>
<th>Keywords</th>
<th>Example Tweet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral Trust – Positive</td>
<td>vote for, for president, [name] [year], love</td>
<td>I think I’m feelin’ the Bern. #DemDebate</td>
</tr>
<tr>
<td>Behavioral Trust – Negative</td>
<td>can’t vote, can’t trust, distrust, hate</td>
<td>No I don’t like Bernie I don’t trust Bernie He’s an empty suit</td>
</tr>
<tr>
<td>Operational Trust – Positive</td>
<td>vote counts, gotv, get out the vote, democracy</td>
<td>I was the 9th voter at my polling place this morning at 8am...hope it picks up! #vote</td>
</tr>
<tr>
<td>Operational Trust – Negative</td>
<td>unfair, flawed, rigged, sham</td>
<td>Also, I think 65% of the Legislative process is spent on recognizing people all day in the Senate and House.</td>
</tr>
<tr>
<td>Institutional Trust – Positive</td>
<td>right thing, trust, faith, confidence</td>
<td>@Jaketapper @RickSantorum I trust the FBI Not sleaze.</td>
</tr>
<tr>
<td>Institutional Trust – Negative</td>
<td>shame, corrupt, phony</td>
<td>The Senate Judiciary Committee is the definition of acting in bad faith.</td>
</tr>
<tr>
<td>Irrelevant</td>
<td></td>
<td>I cast my Heisman ballot for #TuaTagovailoa! Click the player you think deserves the Heisman House vote. (@NissanUSA)</td>
</tr>
</tbody>
</table>

Table 1. Crimson Hexagon Custom Categories
Results

Over a period of more than five years (Jan-01-2014 to Jun-13-2019), we collected 49,964,168 English-language tweets. Figure 3 shows the volume of tweets on a monthly basis. Descriptive statistics for the sample at hand include 60% male and 40% female, only 16.2% of tweets included information about age, and of those tweets, 79.9% reported to be over the age of 35. Although women living in the United States are typically slightly underrepresented on Twitter when compared to the U.S. population (52% women, 48% men in U.S. population, while 50% women, 50% men in the U.S. are Twitter users (Wojcik & Hughes 2019), only 65.4% of tweets included information about gender. Accordingly, not much is to be taken from the proportions of gender or age due to the number of users who omit this information. Nov-09-2016 had the record highest volume while Dec-25-2014 had the record lowest volume. Furthermore, in terms of predefined sentiment categories and excluding irrelevant posts, 73.5% of posts were found to contain negative sentiment while 26.5% contained positive sentiment (5% of the total posts were excluded as irrelevant). Regarding emotion, again excluding the irrelevant tweets, 34.7% of tweets were found to contain sadness, 28.9% were found to contain disgust, 19.3% were found to contain fear, 9.1% were found to contain joy, 4.2% were found to contain anger, and 3.8% were found to contain surprise.

![Figure 3. Monthly Volume of Tweets in Sample](image)

As shown in Figure 4, the three modes of trust production defined and explained in the second section are used here as custom categories and are displayed. As behavioral trust certainly takes up most of the conversation at 85% (23% positive, 62% negative), operational trust has an 10% share of voice (3% positive, 7% negative) while institutional trust takes 5% (less than 1% positive, 5% negative). We also see a significant reversal in sentiment from positive to negative in share of voice of behavioral trust during this time period, with positive sentiment in this category decreasing by 31% and negative sentiment in this category increasing by 34%.

In Figure 5, upon excluding behavioral trust from analysis, a notable tick upwards is visible in both operational trust and institutional trust sentiments with several spikes in each. It is also clear that a reduction of institutional trust sentiment follows the election cycle (Jan-2016 to Nov-2016) but increased beyond pre-2016 levels after the 2016 election. Also, of note is the significant increase in operational trust negative sentiment across the entire time range of the sample, in addition to the percentage value of those tweets that were found to be irrelevant. When including behavioral trust in the analysis, only 5% of tweets are found to be irrelevant, whereas that percentage increases to 27% when excluding behavioral trust. This discrepancy will be further discussed in the next section.

In Figure 6, upon reducing the window of time to Jan-01-2016 to Jan-01-2019 and excluding behavioral trust from analysis, a significant increase in tweets classifying as institutional trust (specifically, negative sentiment) is shown around November 2016. During the same period, operational trust fell by 14% (the same amount that institutional trust negative rose), with 10% of that decline showing in operational trust positive and the remaining 4% showing in operational trust negative. Also notable is a peak of negative sentiment in both institutional and operational trust around August 2017.
Figure 4. Share of Voice and Sentiment Analysis of Three Trust Production Modes' Volumes and Percentages

![Figure 4](image1)

Figure 5. Share of Voice and Sentiment Analysis of Institutional and Operational Trust Proportion for Entire Time Range

![Figure 5](image2)

Discussion and Limitations

The most notable spike in Figure 6 occurs at the beginning of August 2017. During this time, we see institutional and operational trust negative sentiment with a steep increase, while both institutional and operational trust positive sentiment nearly disappears from the graph. News articles show that the most significant political events that occurred during this period of time were President Trump’s remarks in Charlottesville (Shear & Haberman 2017) and the beginning of escalating tensions between the United States and North Korea (Baker & Sang-Hun 2017). With more targeted queries, reducing the sample size...
and time range, and increasing amount of posts trained, future research could attempt to use this sequence of events to hypothesize cause and effect with these and/or other events. The analytics also point to a trend in the last five years where raw volume of user tweets concerning trust in government has increased, and with that increase, behavioral trust in government sentiments have largely trended negative. Additionally, after November 2016, a shift occurred where raw volume of tweets concerning operational and institutional trust increased and never returned to pre-November 2016 levels.

Regarding limitations of this study, as noted earlier, the authors found it relatively challenging to train three out of the seven categories, due to cross-contamination of tweets that were offered by the CH training module—even with targeted second-level queries intended to reveal explicit posts that would only apply to one such category. The three categories—operational trust (positive), operational trust (negative), and institutional trust (positive)—had a large proportion of tweets that threatened cross-contamination of the training data, as many posts that would qualify for one such category also included data that qualified it for another category, i.e. ‘I trust [insert institution] will look in to [insert political figure/institution]’s corruption’. This type of cross-contamination is explicitly discouraged per CH training guidelines, and it was accordingly decided to choose purity of the training data over increasing the number of trained posts. As a ramification of this choice, the significant increase of tweets found irrelevant when excluding behavioral trust (from 5% to 27%) may be caused to this training discrepancy, and should be taken in to account when interpreting the data.

Further, the demographics from the sample are not necessarily representative, as Twitter users may omit gender and age. Although efforts were made to leverage publicly available data to sanity check the sample (for example, using Frommer (2017) to verify no Russia-linked Twitter handles were in the sample, manually looking through individual tweets pre- and post-training, etc.), these efforts are minimal when the sample exceeds more than 49 million tweets. Although we see a general increase in raw volume of tweets across the sample, it should be noted that (1) the generated query targets political figures in the more recent 2016, 2018, and 2020 elections but only the sitting President in 2014, President Barack H. Obama, and (2) Twitter usage tends to increase over time, so proportions tend to be more analytically valuable over raw numbers. Furthermore, using only Twitter as a data source not only excludes every person who does not use social media, but other social media platforms as well.

**Future Research**

This research aimed in part to determine whether, and if so, how social media analytics could be leveraged to identify, predict, and/or drive trust in government sentiments. As no other research has used social
media analytics in the realm of assessing trust in government outside of direct government communications or responses, this research was exploratory and did not aim to test a set of hypotheses. However, many questions emerged during this exploratory analysis. For example, what caused or contributed to the long-term shift in share of voice post-November 2016? It is assumed that, as more users join and use Twitter as time goes on per the technological revolution, the share of voice between the three trust production modes will change. What types of events cause positive sentiment of the three trust production modes to increase? What types of events cause negative sentiment of the three trust production modes to increase?

A key question required in the drive to leverage social media analytics to predict and/or drive trust in government sentiments is how these tenets of trust in government play together. According to the findings above, given the limitations discussed above and the scope of this research effort, behavioral trust currently takes the largest piece of share of voice on political Twitter in recent years—even amidst a sharp increase in institutional and operational trust sentiments (albeit, negative). We have used existing frameworks and models in this research to inform a model of trust in government sentiments, but how do the constructs in our model inform each other? Clearly, they are interdependent, but to what extent? What are the exceptions? Many other hypotheses, future research topics, and methodologies could be explored or tested. One example is a more comprehensive look across multiple types of social media, and not solely Twitter. Methodological research regarding using CH could also be useful, as another limitation could be the search query itself in Figure 2; although efforts were also made in making the query as inclusive and specific as possible (i.e. using online thesauruses to identify as many synonyms of keywords as possible), there is always room for improvement.

Conclusion

E-Government is certainly not straightforward in its definitions, terminologies, and commonly used vernacular. However, the United States’ efforts in E-Government have been dismal to say the least, and it is theorized and supported by both extant literature and this research’s analysis and findings that trust in government plays a fundamental role in E-Government success (Aladwani & Dwivedi 2018; Ayyad 2017; Hapsara 2016; Lappas et al. 2015; Olphert & Damodaran 2007; Tassabehji et al. 2007; Thomas 1998). One of the ways the U.S. could see a successful E-Government initiative, thereby increasing democratic representation of this individual, is by researching and learning how to identify, predict, and/or drive trust in government sentiments through social media analytics. Specifically, this research used Crimson Hexagon’s Forsight platform and discovered a trend over the last five years where Twitter tweets concerning trust in government sentiments have increased in raw volume, and that increase has seen with it an increase in negative behavioral trust sentiments, an increase in operational and institutional trust raw volume, and a stark increase in negative institutional trust sentiments. As mentioned in the last section, future research should explore the multitude of questions and potential hypotheses that were discovered throughout the analysis, findings, and discussion, and were accordingly out of the scope of this research effort. Additionally, the provided theoretical framework explaining the three modes of trust production in government could be tested, validated, improved, or otherwise challenged.

REFERENCES


