THE TALK ABOUT TRUMP
(and other highlights of the second wave of presidential candidate announcements)
August 4, 2015

A PEORIA Project Report
Associate Professors Michael Cornfield and Lara M. Brown
The Graduate School of Political Management (GSPM)
The George Washington University

In Partnership with Zignal Labs

INTRODUCTION

What is the **PEORIA** Project?

Welcome (back) to the **PEORIA** Project, the GSPM's harnessing of Zignal Labs's realtime, cross media story-tracking platform to analyze the "public echoes" arising from the 2016 presidential campaigns. While others during the "invisible primary" will investigate each candidate's poll standing, dollars raised and spent, and endorsements won, we track and measure **words**—the chatter about the candidates and the echo of their campaign messages in both mainstream and social media.

PEORIA is an acronym for **P**ublic **E**choes **O**f **R**hetoric **I**n **A**merica, chosen as an allusion to the old vaudeville and marketing phrase "will it play in Peoria?"

Our fundamental premise is that how candidates and their messages play on the trail with the media and the public both affect and reflect the voters' presidential preferences. When a candidate says and stages it right, it resonates positively with the public, creating an echo that benefits the campaign. Of course, the opposite can also occur with negative echoes. From positive to negative, people respond to crafted messages,

brands, catch-phrases, sound bites, slogans, and gaffes[1] as they surface in news and social media, affecting their choices down the road.

The **PEORIA** Project follows the candidates and their campaign messages, measuring the public echoes that surface in all types of media.

What does this second PEORIA report examine?

Our first report analyzed the public echoes during the period from March 15 to May 15, 2015, focusing on the formal presidential candidacy announcements and the initial branding attempts of these campaigns. This companion report does the same for the period May 16 to July 19, 2015. (We added a few days to include a week's worth of conversation about Scott Walker's announcement. The candidacies of John Kasich and Jim Gilmore formally launched too late for our research.)

We assess the textual data to determine the relative effectiveness of the candidates' performances and the campaigns' communication strategies. Then, we assign each candidate an "echo value" rating, which ranges on a scale from 1 to 11.

In this 63-day period, **twelve** individuals formally declared their candidacy for the 2016 presidential election:

Rick Santorum	May 27
George Pataki	May 28
Martin O'Malley	May 30
Lindsey Graham	June 1
Lincoln Chafee	June 3
Rick Perry	June 4
Jeb Bush	June 15
Donald Trump	June 16
Bobby Jindal	June 24
Chris Christie	June 30
Jim Webb	July 2
Scott Walker	July 13

[1] Of course, the public responds to images as well. We presume that any image which has a significant effect on candidate reputation and voter choice becomes a topic of discussion and acquires its own caption or summary title, e.g. "Dukakis in the tank" and "Bush looking at his watch." We, thus, pick up memorable images through the words that are commonly used to describe them.

What Can Candidacy Announcements Do for a Presidential Campaign?

A candidacy announcement is as carefully crafted a campaign message as exists in electoral politics. Its staging today occurs as much online as it does in physical space and across mass mediated channels. It is designed by campaign strategists to make a favorable first impression of the candidate on the public within the context and narrative of the race. The announcement is timed and phrased with other likely candidates in mind. It also attempts to cast the candidate as the best person to serve as the president in this day and age. Accordingly, campaigns sow a differentiating identifier into their announcement texts: a phrase that will set the candidate apart from their competitors, providing an advantageous comparison.

Notably, it is possible that rejection of the announcement's fanfare (e.g., silence, outrage) is mostly what registers as a public echo in the aftermath of the event. Again, not all public echoes are positive, nor are they always what campaigns intend. As such, a campaign's ability to detect and adjust to either non-existent or negative feedback in a short time is also part of the story of echoes.

METHODOLOGY

Here we present information on the Zignal Labs platform and our own metrics as applied to the data the platform contains.

A) Zignal Labs

What does the data universe contain?

SOCIAL MEDIA: Every single tweet, publicly available Facebook user posts, every single mention in social/online video (YouTube, Vimeo, MediaBistro), 30+ million blogs. NEWS and MAINSTREAM MEDIA: news stories from more than 100,000 online outlets including licensed content, all LexisNexis News Content (print news, magazines/journals/newspapers,etc), all television closed caption content from 900 channels in every media market in the US.

Who qualifies as a presidential candidate?

The database indexes information about major candidates whose names appear in the leading polls and political insider publications. For the period March 15 to May 15, 2015, eight candidates announcing within that time period were analyzed as a set. For the period May 16 to July 20, 2015, the twelve candidates announced within that time period were analyzed as a set.

What counts as a mention?

Any tweet, news story, blog, video, LexisNexis story or broadcast clip (closed captioning) that matches a query (a query is a combination of certain keywords or phrases). For this project and related ones, Zignal has built a custom database with real-time continuous queries of the presidential candidates' names. Multiple mentions within a content unit or "document" are not counted extra.

How is share of voice calculated?

Share of voice is calculated by summing up the mentions (across all media types) in each candidate profile and taking the ratio of that candidate's total to the entire set of candidates.

How are sentiment classifications (positive, negative, and neutral) determined?

Sentiment is determined using natural language processing technology (NLP). Zignal's NLP algorithm assigns a positive, negative, or neutral score to every document to provide an overall sentiment rating. Frequency, intensity, and sentence structure are factored into the model. For example, "love" has a higher score than "like", but an overall negative prediction will still occur if negations such as "not" or "neither/nor" are present within the sentence. Adverbs also serve as multipliers, with phrases like "very good" scoring higher than "good." The backbone of this algorithm is a Recursive Neural Tensor Network, a type of deep learning algorithm that allows us to continually modify and fine-tune our model as time goes on. Unfortunately, sentiment detection is still not an exact science, and NLP fares poorly when sarcasm is present or the overall diction is ambiguous.

*Over the course of the project, GWU has the opportunity to manually override and/or correct sentiment which helps train and improve the models' performances. In addition, the project only reports "**net sentiment**," positive less negative or vice versa as the case may be. This move assumes that erroneous classifications are randomly distributed, and that the directionality of sentiment is a fairer albeit thinner indicator than reporting percentages from all three categories.*

How are other indicators determined?

Popular Tweets: number of retweets that a tweet gets

Top Issues: a second/third level of filtering. Profile queries (the candidates' names) control what gets ingested into the platform, and issues are tags that categorize the data ingested. Top issues is thus a sorted list of the most frequent tags by candidate.

B) The PEORIA Project

What is an echo?

We use the word "echo" to denote the aggregate verbatim appearances (i.e. mentions) in the database of words and phrases issued by a campaign. An echo is not the same as "resonance," the absorption of campaign messages into the mindset of listeners and speakers.

What are Announcement Echo Dates?

The number of days between the mentions peak on announcement day and the leveling off of mentions before any subsequent rises.

What is a Web Site Share?

> *A mention that includes a url of a page on one of the campaign's web sites. A candidate's web site serves as a hub where the campaign tries to convert attention or interest into favorable attitudes and enduring support, particularly through the use of landing pages to collect visitors' e-mails and other contact information and to solicit and collect donations.*

What are the mainstream and social media?

See the data universe answer under the Zignal Labs section.

What is the GSPM Echo Rating?

A summary assessment combining the objective data described above and in the slides with our subjective sense as longtime students of presidential campaigns (slide 27).

If the candidacy announcement is the political equivalent of a fanfare, then the echo varies from:

11 - Historic. *When the books are written and standards are invoked for future announcements, this will be up there.*

7 - Memorable. *Recall association with the candidate; effective branding of the person in the race of these times.*

4 - Noticed. *The intended brand echoed in the news and social media spheres.*

1 - Crickets. *Virtually imperceptible echo in society, albeit searchable as reaction data, and therefore more than zero.*

FINDINGS

1. **(See Slide #4.) Since June 16, Donald Trump has overwhelmed the presidential campaign conversation.** He garnered as many news and social media mentions in one month as Hillary Clinton did in four months, close to ten million each.

- Trump's dominance is all the more impressive given that the number of Trump mentions during the first 30 days of this period (May 16-Jun 15) was close to zero. **Bernie Sanders is the only candidate besides Trump who increased his share of voice during the "post-Trump" period (June 16-July 19).**

- The candidates divide by total mentions into three tiers (these will be discussed more later): Conversation Starters (in red), Reserves (in blue), and on the Taxi Squad (in yellow). Trump was the only Starter in the GOP field, joined by two Democrats, Clinton and Sanders. Five GOP candidates who nationally all attracted between 3 and 8 percent of Share Of Voice, make up the Reserves. The remaining candidates garnered less than 3%.

- That said, both Chris Christie and Rick Perry appear to be more in the conversation when the talk is only about GOP candidates. Hence, it would appear that out of the 20 candidates in this report; eight of them are seriously in contention for the GOP nomination.

- Total Mentions over the 63-day period show us how much people were talking about the 2016 presidential contest. The total mentions of all candidates during this period was 21,848,035 (about 14,449 per hour); about 15.2 million were the Republicans only – AND notably, almost half of that was about Trump, in only the second half of the time period.

- To put the presidential campaign conversation in social context, we note (as we did in our first report) that talk about the Kentucky Derby on the day it was run averaged 16,670 mentions per hour (between 6AM and 9PM Pacific Time). So even with Trump, the Run for the Roses attracted more talk/interest than the Run for the White House.

2. **(See Slide #5.) Talk about Trump spiked seven times between June 16 and July 20. (See Slide #6) This muffled the public echoes of the candidates who followed him, and (See Slide #7) severely dampened the echo for Jeb Bush who announced the day before Trump.**

- The talk about Trump peaked or spiked seven times in the last half of our study period. All seven of these Trump Towers attracted more mentions than any of the other twelve newly announced candidates did at any time. In fact, even when talk about Trump was in a relative valley, the mention level still exceeded the mentions for all the other candidates on all days except for five candidate

announcement spikes, for Perry, Bush, Jindal, Christie, and Walker. Two of Trump's Towers occurred on the same day as Christie and Walker.

- Many candidate announcements barely registered in the month of Trump's dominance of the conversation.

- The only candidate who has garnered a larger spike than Trump's largest, on his announcement day, was Hillary Clinton (over 800K), as shown in the previous PEORIA report.

- Trump talk muffled the announcement echo of Bush and those who followed. The average share of voice during pre-Trump announcement echoes was 30.9%. After his announcement it was 12.3%. In the previous PEORIA report, the average share of voice for each candidate during their echo period was 35%.

- No candidate announcement suffered as much deprivation of conversational attention as Bush. Slide #7 shows the announcement share of voice for the three days before, the day of, and the three days after the event. This allows for a more standard comparison of how they performed on their share of voice. **Bush went from having over 80% share of voice on his Announcement Day to having only 15% SOV in the three days after.** That drop is large in magnitude and clearly is related to Trump announcing on June 16th – the very next day after Bush announced.

3. **(See Slide #8) The drown-out was especially pronounced in social media,** where Trump received 61% of public talk in the two-month period (Bush was next highest among incoming candidates at 12%; Walker and Christie followed with 5% each). In the news media, Bush had nearly as large a share of voice (25%) as Trump (26%).

- As we discussed in our last report, we're interested in similarities and differences between mainstream and social media voices. While it is too early to understand which ways the causal arrow goes (do the news media cause social, vice versa, or are they independent conversations) we are keeping track of the differences.

- Our prime research questions in this regard: Will one type of media identify the eventual party nominee earlier than the other? Who are the darlings in each category? What are the dynamics between types of media and how do those change with different circumstances?

- For now what is clear is that the mainstream media distributed attention across the field of (new) candidates much more than social media, where The Donald dominated, and Bush and Walker fared next best.

4. **(See Slide #9.) Trump did poorly in converting mentions into web site shares, with a low Echo Conversion rate of 0.2%.**

- The Echo Conversion rate measures the ratio of campaign web site shares to total mentions. It indicated the effectiveness of the candidate's campaign organization and network of supporters.

- While Trump has a large number of website shares relative to the other newcomers in our second study period, his Echo Conversion rate is tied for the poorest among those candidates where a rate is detectable.

- In contrast, Jindal and Santorum had good conversion rates despite their low volume of mentions.

- The mean number of mentions helps us place this in perspective. It stood at 1,018,451 for the period. This was less than for the first period we studied (1,289,326), notwithstanding the presence of talk about Trump. However, subtract Trump and the mean for the second period drops to 465,174, greatly below that for the first period.

5. **(See Slides #10 and #11.) Overall talk expanded but (See Slide #12.) became more negative after Trump entered.**

- We now pull back to look at both time periods together, spanning just more than four months. BT, or Before Trump, denotes the three months prior to his candidacy announcement; AT, After Trump, refers to the 33 days following.

- Trump certainly pumped up the volume of mentions – more than double the number of mentions per day (from 212K to 473K).

- Before Trump it was the Hillary Show. And while she still has a larger total number of mentions, she has been running for more than two months longer than Trump.

- Generally speaking, the conversations have become more negative since Trump entered the race. Only 5 candidates had positive conversations swirling around them: three Democrats – Chafee, Clinton, and Sanders – and two Republicans, Carson and Fiorina.

- Those two Republicans were away from the fracas of accusations and insults involving Trump. The five most involved in these hostile exchanges all suffered negative net sentiment, including Trump. Cruz, who expressed support for Trump, and Paul, who stayed neutral, fared better.

6. **(See Slide #13.) Trump changed the main issue topic of conversation to immigration….and away from an expected discussion of campaign finance.**

- This is somewhat amazing – Trump also changed the main topic of conversation. Given that finance reports were due at the end of June, and that the numbers would be mostly reported between 6/30 and the end of the period (7/19), it is surprising how little the conversation about money increase. These were the first finance reports for the candidates: the so-called "invisible primary's" first test in actual comparative numbers. They were also the first in this cycle of large dollars being collected apart from the campaigns, in Super PACs and non-profit organizations. Yet the number of mentions about money only increased by only about 4,000 more per day (a 35% increase).

- By contrast, the controversial and emotional but second-ring issue of immigration leaped ahead in total volume after Trump's entry into the race, more than doubling the number of mentions per day (a 54% increase). In addition (but not shown here), immigration became one of the five most-talked about issues for every candidate entering in the second two-month time period except for Scott Walker.

- In essence, Trump's entry, and of course his incendiary remarks about undocumented immigrants coming across the Mexican border, changed the conversational agenda.

7. **(See Slides #14 through #17.) Analysis of top tweets reveals a significant topical divide: for some candidates, the top subject was the campaign and the issues; for a second group, celebrities, holidays, and breaking news; and for a third group, Trump himself. Only the first group's tweets stayed on message.**

- As with the Echo Conversion Rate, the content of the most shared campaign tweet is an indicator of campaign proficiency, in this case the capacity to get out one's message.

- This table (14) shows the most shared campaign tweet from the newly announced twelve candidates during the three-month Before Trump time period. Only three campaigns, Bush, Chafee, and Santorum, were able to have their announcement Tweet become the most retweeted of the period. In all cases, the volume was low.

- In marked contrast, there was more volume and more on-message content in the most shared campaign tweets from the previously announced eight candidates during this pre-Trump period.

- The highest volume for a second group candidate belonged to Bush, at 1,378. Hillary Clinton's campaign announcement tweet during this period was retweeted over 95,000 times. Ted Cruz managed more than 13,000, and Rand Paul, Marco Rubio, Bernie Sanders, and Ben Carson all had their campaign tweets retweeted over 2,500 times. While they had up to two months longer for the tweets to float through and be retweeted, still these numbers are much larger than the second wave's numbers. The contrast of focus on garnering campaign support is dramatic.

- This table (16) displays the content of the most shared campaign tweet from the newly announced twelve candidates during the post-Trump period.

- Two candidates, Trump and Webb, had their announcement tweet shared most.

- Three candidates, Graham, Bush, and Jindal, had a tweet about Trump shared most.

- Two candidates, Perry and Christie, had a tweet about non-political content shared most.

- In the After Trump period, Hillary Clinton's campaign tweet about climate change was extraordinarily successful. It was retweeted more than 57K times, while The Donald's best tweet only made it to 11K. The Bernie Sanders tweet about Charleston also amassed an impressive number of retweets.

8. So, is Trump leading the Republican pack in conversational terms? (See Slide #18.) On volume and share of voice, very much so. (See Slides #19 and #20.) But on message retweeting and mentions, it's a mixed bag.

- Is Trump Really Leading? In other words, are there signs that voices in the conversation might be converted into Republican votes and delegates?

- Taking a look at the Republican field (without Kasich), we created a weighted index of total mentions. First, we divided the four+ months into 4 periods (March 15-April 15; April 16-May 15; May 16-June 15; June 16-July 19). Then, we weighted each period so that the more recent counted as the more important. As such, we multiplied the candidates' total mentions in the first period by 1; the candidates' total mentions in the second period by 2; the candidates' total mentions in the third period by 3; the candidates' total mentions in the fourth period by 4. Then, we added those together, and calculated the share of voice (in %). As you can see, Donald Trump, who had nearly a zero share of voice prior to the last period still wins out. Beyond this, the field (at least on SOV) appears to break into our three tiers – Starters (above 10%) Trump Cruz and Bush; Reserves (between 3-7%) Walker Rubio Paul Christie and Perry; and Taxi Squad (below 3%) Huckabee Carson Jindal Graham Santorum Fiorina and Pataki.

- A very different three tiers emerge when we follow the same calculation method (weighting each period and adding them together), but use the number of retweets for the campaign's Top Campaign Tweet over the number of Twitter mentions. Hence, the total number of retweets was weighted for each period and the total number of Twitter mentions was also weighted. Using the weighted numbers, a percent of retweets per Twitter Mentions was then calculated.

- As can be seen (19), Pataki and Carson had the highest percentages, but some of this is simply a function of volume (see next slide). Also, Pataki was tweeting about the Women's Soccer Ticker Tape Parade in NYC in the 4th month (the one with the weighting); Chris Christie, similarly, was tweeting about the World Cup Soccer Final during the 4th month.

- Rand Paul is the only candidate who makes into the top six on both of our Twitter-based metrics. He is also the only one whose tweets relayed his issue positions (on tax cuts, NSA security, etc.). The other Republican candidates' tweets were either responding to events (gay marriage/SCOTUS decision; Planned Parenthood; Iran Deal) or to Donald Trump's comments about Senator McCain and others.

- These (20) are the un-weighted Twitter retweets and Twitter mentions. As becomes clear, aside from Donald Trump, Ted Cruz not only has a higher volume of mentions, but also a higher number of retweets than all of the other GOP candidates. It would seem that he only appears in that middle tier on the previous chart because so much of his volume happened in time periods 1 and 2. Simply ranking the candidates without weighting would likely show that the top tier (by RT volume) is Trump, Cruz, Carson, and Paul; the middle tier is Scott Walker, Jeb Bush and Marco Rubio, and the third tier includes the rest. This provides a different take than the other two looks at this field.

CONCLUSION (See Slide #21.)

For our entire 123-day time period, our summary rating for the twenty candidates who launched a campaign for the presidency is as follows:

Cruz, Clinton	9
Sanders	8
Trump	7
Bush	6
Paul, Walker	5
Rubio	4
Carson, Christie, Santorum	3
Huckabee, Jindal, Perry	2
Chafee, Fiorina, Graham, O'Malley, Pataki, Webb	1

- These ratings assess how well the campaigns are leveraging the conversations about their candidates. The scale goes from 1 [Crickets] through 4 [Noticed] and 7 [Memorable] all the way to 11 [Historic].

- While Donald Trump would probably score an 11 on volume and share metrics alone, the inclusion of echo conversation and message re-tweets drops his rating to 7. It is not clear that his campaign is converting talk into support that will in turn translate into votes and delegates at the Republican National Convention.

- Because of these conversion metrics, Ted Cruz outranks Trump. So do two Democrats, Clinton and Sanders.

- In the final analysis the talk about Trump seemed to hurt more candidates than it helped.