

Pandit: An Election Analysis Site

Ray Vichot
Georgia Institute of Technology
1451 Mecaslin St NW #8205
Atlanta, GA 30332-0165
+1 646-270-7670
rayv@stop-making-sense.net

ABSTRACT

Pandit is a site that brings together various aspects of election coverage, specifically numerical data, media and campaign generated narratives, and temporal data of events that have occurred.

Keywords

Election, primary, delegates, pundit, analysis, timeline, narrative

1. INTRODUCTION

Pandit is a website that utilizes various tools of the medium in coordination in order to link various forms of information pertinent to an election in order to answer questions about the media's relationship to the election process. Specifically, Pandit answers questions regarding the role of narrative creation during a campaign cycle and whether it is the media who creates these narrative frameworks, or if it is an outgrowth on their reporting of the campaigns themselves.

The rationale behind Pandit is to allow users to analyze election data and reporting for themselves. The name Pandit is a knowing reference to the origin of the word pundit. A pundit is defined as "a scholar, a teacher, particularly one skilled in Sanskrit and Hindu law, religion, music or philosophy." The word was borrowed from British Colonialists and, over the years, the meaning has slowly changed from someone who possesses knowledge and learning to an opinionated, bombastic media personality. In naming my project Pandit, I'm attempting to return to the original meaning of the word and re-center knowledge about the electoral process back to the individuals involved.

While this instantiation of Pandit is using the 2008 US Democratic Primary as its subject matter, the basic idea of the system should be extensible to cover any election and possibly any long term events, such as the occupation of Iraq or the developing food shortages in the developing world.

2. CURRENT SOLUTIONS

Things exist that already address my problem space, but they are limited in their scope because they concentrate on separate elements of the issue in question. There are, of course, news sites, which report the events that take place throughout a campaign. Also there is commentary about these events, both as part of a

serious newscast as well as satirical or comedic commentaries as well. These sources provide the narrative aspects of a campaign, reporting on the various claims one candidate has about another's and the general pattern of the nationwide campaign. However these sites do not focus much on the actual outcomes of races and, when they do, they usually attempt to conflate them with the various narratives in place.

In contrast, there are sites such as Democratic Convention Watch and Electoral Vote which attempt to present only hard data regarding the aspects of the election. The sources are great at seeing a relatively objective picture at the trajectory of an election, but again, they are not complete insofar as explaining the complete electoral picture.

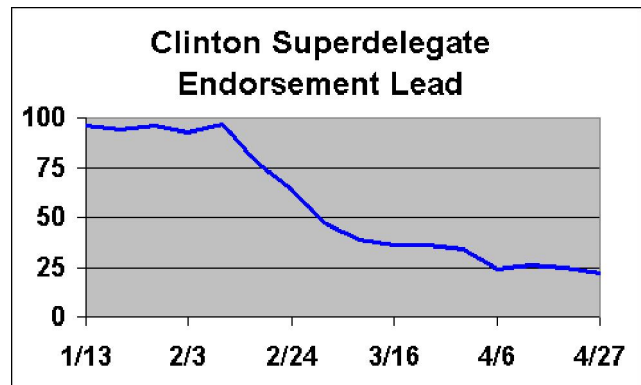


Figure 1: Chart from DemConWatch presented with no context to explain why this data is the case.

Finally both these sources types are both guilty of not placing enough emphasis on the recent history of a campaign. In news sites on the 24 news cycles, new stories and narratives, pop up, are discussed, and then fade away as the next story breaks. In their focus on the "horse race" of the election, they lose sight of the election in context. Data sites, while providing a history of the change in numbers, also do poorly at showing the election from any historical context, usually only explaining the data with one or two paragraphs and letting the numbers speak for themselves. This can lead to someone who wishes to understand what has happened during the election process very confused and muddled as to how one type of data converges with the other.

3. PANDIT

My Data Model

Pandit deals with three types of data. The first is numerical data from election results and poll data. The second are reports of events that occurred with the election process such as campaign stops, debates, and actual elections and caucuses. Finally I deal with media sources that are not strictly reporting on events such as Op-Ed pieces in newspapers, commentary from 24 hour news channels and pop cultural commentary such as found in programs such as the Daily Show and Saturday Night Live as well as online sources such as the D.C. gossip blog, Wonkette.

Organization / Architecture

These seemingly disparate sources of data are organized in two ways. One way is that they are organized by the various topics a specific article or clip is about. These tags are the aggregated and placed in a tag cloud that changes week by week, showing what are the important topics being discussed in any time period by the various news pieces. This show how these various political narratives rise and fall in importance as time goes on.

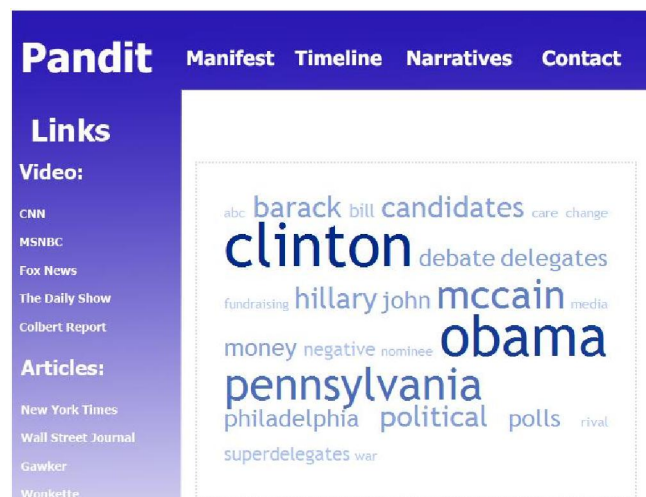


Figure 2: Example narrative set for the week of April 20, 2008

Secondly, these data sources are organized into timelines that are both accessible on their own or, more powerfully, are accessible

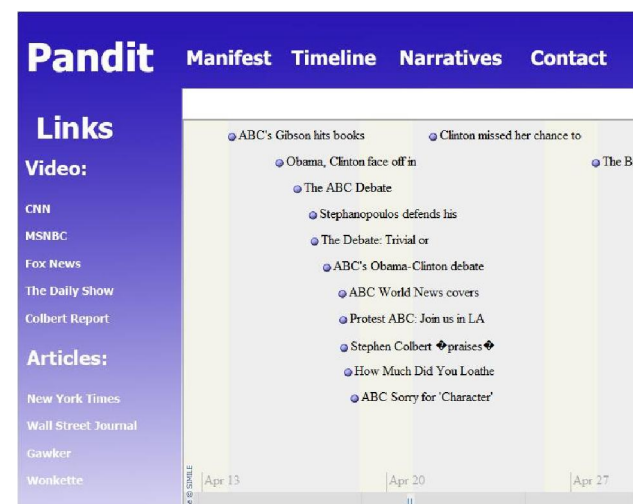


Figure 2: Example timeline related to the term "ABC".

from the narrative cloud themselves. The timelines are constructed using the Timeline tool created by the MIT SIMILE project. Timeline is a javascript based application that can parse XML data to create a temporal visualization that is easily customizable and usable. This way on can select a prominent narrative cloud, such as stories that talk about ABC, as shown in figure 2. Then one can see the timeline of the stories that deal with the network and the election, specifically focusing on the networks treatment of the Philadelphia Debate of April 16.

Interface / Interaction

The Interface consists of a straightforward website, which utilizes various conventions of the media (Text tabs as links, images that are clickable). The main page (Figure 4) consists of Fata pulled from aggregation sources DemConWatch and Electoral Vote. This allows users to quickly survey the current status of the primary process.

Users can then navigate to the Timeline page, which has a timeline of the major events of the election, in this case relaying data about when the various primaries and caucuses were held, along with data on results, pulled from CNN's election coverage.

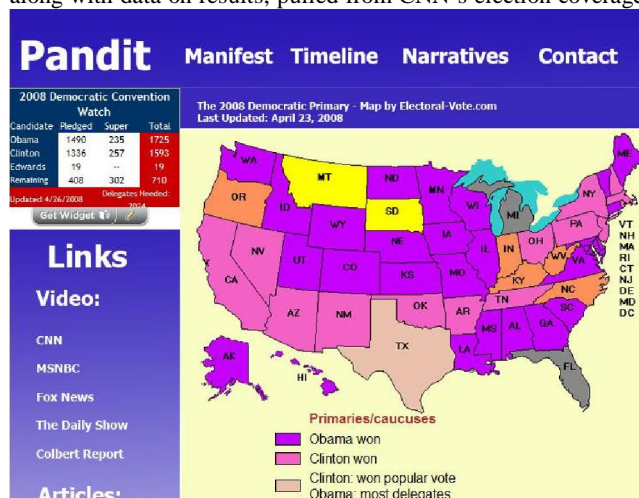


Figure 4: Pandit Main Page

Besides the timeline, the user can navigate to the Narratives section which will bring up the narrative cloud that a user can view to see what threads are most talked about for a specific week and then click through to gain a timeline specific to that thread.

A user would enter the site and see the current delegate and superdelegate count as well which states have been won by which candidate and which states still need to hold primaries. Then the user, if they were interested in seeing how Hillary won Pennsylvania, could click on the Narratives and select either Clinton or Pennsylvania, if they were interested in more the primary itself or in the coverage/portrayal of Clinton the candidate herself.

This would open up a timeline that, in the first case would show events such as Hilary being booed in Iowa, her crying in New Hampshire which led to a win, her campaign string of losses between February 5th and March 4th, to the current situation where it seems impossible for her to win. It would also notes stories such as Tina Fey's endorsement on SNL, her drinking in Pennsylvania, and her 'obliterate Iran' commercial that aired just before the primary.

In the second case, the user would get stories that dealt more with the voters of Pennsylvania, how each candidate performed at their events, and various talking heads talking about the importance/unimportance/semi-importance of this primary. There would also be coverage of the Results and what it means to the election in general.

4. FUTURE IMPLEMENTATIONS

Aggregate Timelines using Multiple Narratives

One limitation with the interface at the moment is that each narrative thread is individually linked to its timeline. What would be worthwhile would be to allow a user to aggregate several narrative threads together and generate a timeline that can incorporate multiple elements into one temporal display. This would allow users to maybe see correlation between the rise of one narrative thread and the fall of another.

Separate Data Page

Currently the objective data such as delegate count is placed on the main page. This has some advantage as it allows users who are interested in the "horse race" aspect of elections can be appraised as to the current status of the election. However, as is currently implemented is somewhat inelegant and deserves a separate page of its own, with methods of incorporating that data, specifically superdeleaget endorsements with the timeline model.

Auto-Updating Timelines

Timelines are currently written by hand, due to the fact that metadata is woefully inconsistent per site. A further implementation would include a method of extracting tags from the texts of articles pulled from feeds and then generating the tag cloud as well as placing them in the timeline without manual hardcoding.