Search engines such as Google mediate a substantial amount of human attention, acting as algorithmic gatekeepers and curators as people seek and navigate news and information online. A 2014 survey found that 51% of American respondents got news in the last week via a search engine (Media Insight Project 2014), most likely with either Google or Bing, which together represent a practical duopoly, dominating about 86% of the US market for desktop search (“Rankings—comScore, Inc” 2017). Yet, as proprietary systems, little is known about search engine information biases and how they may be shaping the salience and quality of users’ information exposure. This is of particular importance in considering how voters become informed during elections. The primary focus of this chapter is to begin to illuminate the multifaceted roles that Google’s search technologies have played in algorithmic information curation in the 2016 US elections.

The concern over the power of search engines in mediating access to information is not a new one. As early as 2000 Introna and Nissenbaum expounded on issues of accountability and transparency (Introna and Nissenbaum 2000), and these concerns have been echoed in more recent work (Laidlaw 2009; Granka 2010). Research on the potential for search engines to impact people has enumerated a number of possible repercussions, including the potential to affect attitudes (Knobloch-Westerwick, Johnson, and Westerwick 2015), to alter perceptions based on image
presentations (Kay, Matuszek, and Munson 2015), and to lead to anticompetitive tendencies by privileging preferred products or pages (Hazan 2013; Edelman 2010; Edelman and Lockwood 2011; Alam and Downey 2014).

The fundamental reason why search rankings are so powerful is that the order of information has a substantial impact on human attention and reliance: not only do people click on top results more than lower results (Agichtein et al. 2006) but also they believe something is more relevant if it is in a higher position (Pan et al. 2007). In the media environment, search engines have the power to “grant visibility and certify meanings” as Gillespie (2017) has written. Because search engines act as gatekeepers to information, the way they are designed and wielded can yield subtle or overt power. For instance, the prevailing design value baked into (and optimized for in) Google search is that of “relevance,” but this may come at the cost of considering other possible design values such as fairness or information diversity (Van Couvering 2007). In China the state exercises its will to censor and block politically inexpedient information via the predominating nationalized search engine, Baidu (Jiang 2013, 2014).

In politics the most startling evidence for the power of search engines to impact election processes comes from Epstein and Robertson (2015). In their laboratory study, they showed that by manipulating the ordering of supporting information for candidates in a mock search engine, they could shift the voting preferences of undecided voters. Other work has explored differences between search engines in terms of the types of sources surfaced for election-related queries. Muddiman (2013) found that Google ranked campaign-controlled and advocacy pages higher than Yahoo, for instance. Research has shown that users search for candidate information during key campaign events such as debates and gaffes (Trevisan et al. 2016), however query strategies used to seek information about candidates are an underresearched topic.

Legal scholars have explored the complexities of applying regulation to search engines (Grimmelmann 2014; Bracha and Pasquale 2008) in order to quell their potential for abuse. One of the key barriers to regulating search engines in the US jurisdiction is that their output is currently interpreted as free speech by courts (Laidlaw 2009). While there may be other regulatory options for ensuring search engine accountability, such as setting up nonpublic limited-access courts (Bracha and Pasquale 2008), these only partially address accountability of such systems due to their nonpublic nature. Given the challenges to the regulation of search engines, an alternative to achieving algorithmic accountability (Diakopoulos 2015) of search engines, and the approach we have taken here, is to audit their results by gathering data around specific candidate-related queries.
Yet there are a number of confounds to studying search engine results using an auditing approach. Factors such as the query terms used, language setting, geolocation, search history, and logged-in status, not to mention randomization or A/B testing that the search engine may be undertaking make it difficult to see the outputs of the system as stable (Ørmen 2015; Xing et al. 2014; Kliman-Silver et al. 2015). Moreover, the unavoidable constitution of the search engine as a sociotechnical system (Gillespie 2017) that embeds feedback loops with user inputs make it difficult to disambiguate the role of the algorithm versus the role of the human user. And this is further compounded by the knowledge that different demographics and users may exhibit different types of information-seeking behavior (Weber and Jaimes 2011). The choice of query terms can itself lead to politically biased results, as can the “universe” or input bias of the pages indexed by the search engine. Teasing apart differences in results arising from query term selection, input bias, and actual algorithmic ranking bias is a methodologically challenging proposition (Magno et al. 2016; Kulshrestha et al. 2017).

Taking these methodological challenges into account to the extent possible, this chapter details four distinct case studies/audits. Each case illustrates how Google mediates candidate information in different ways: (1) *Search Results*: the set of sites ranked on the first page of search results, including their support or opposition to the candidate; (2) *Issue Guide*: the presentation of an issue guide integrated into search results on candidates constituted from algorithmically curated quotations from each candidate gleaned from news coverage; (3) *In the News*: the presentation of news information about each candidate as framed in the “In the News” section, which occupies privileged screen real-estate at the top of search results; and (4) *Visual Framing*: candidates are visually framed in search results as a consequence of the image selections that Google makes.

The following subsections detail each of these case studies, examining why and how each facet of the search engine is important to consider with respect to how voters gather information about candidates. We then conclude the chapter with a comparison and contrast of the cases, including methodological challenges and an elaboration on where additional work is needed to better understand how search engines inform voters.

**CASE 1: SEARCH RESULTS**

Previous research has indicated that “biased search rankings can shift the voting preferences of undecided voters by 20% or more” (Epstein and Robertson 2015). The impact of that bias may be felt more if users assume
that the results are neutral and, therefore, trustworthy. To begin to articulate such biases in a real search engine a crowdsourced analysis was used to explore differences in how primary candidates in the 2016 US presidential election were presented by Google.

Methods

Search result links were collected on December 1, 2015, from nonpersonalized Google searches for each candidate’s complete name (i.e., “first-name last-name”). The top 10 ranked results for each of the 16 candidates (13 Republicans and 3 Democrats) were collected. The focus on the first page was grounded in knowledge of search users’ behavior: users end up clicking on the first 10 results 70% of the time (Loiz 2014).

The process to determine whether the linked webpages returned by the search were positive or negative for each candidate was crowdsourced using the Amazon Mechanical Turk (AMT) microtask market. Each link was presented to three separate AMT workers, who determined whether the linked webpages returned by the search were favorable or oppositional toward each candidate.

Several iterations of the instructions for workers were tested with pilot data. As a basis for that process, an initial sample of the websites was rated by the researchers and measured against the ratings of AMT respondents. Language and instructions for the task were adjusted until there was a good level of agreement between the researcher rating and the respondents’ ratings. The intraclass correlation coefficient between ratings done by AMT and the researcher’s baseline rating was 0.912.

The final version of the AMT task included simple instructions on how to read the websites. The respondents were instructed to only evaluate the linked webpage, to consider only the primary content of the webpage, and to decide whether the webpage favored or opposed a particular candidate. Workers were directed not to consider source reputation, but instead to focus on the content of the specific article.

We incorporated strategies into our task design to help ensure quality responses from crowdsourcer workers (Kittur, Chi, and Suh 2008) including asking respondents to answer a series of four questions: (1) “Please summarize the content of the website in one sentence”; (2) “To what extent does the linked webpage favor (candidate name)?”; (3) “To what extent does the linked webpage oppose (candidate name)?”; and (4) “Please explain the reasoning behind your ratings in a few sentences. Be sure to include relevant excerpts from the page that support your ratings.”
Questions 1 and 4, which were answered via free text, were used to control the quality of the information. Question 1 was designed so that respondents had to explicitly demonstrate that they, in fact, had read and understood the page. Question 4 was designed to ensure that respondents had fully considered the page; it also helped explain the respondent’s reasoning. Any response that did not include answers to Questions 1 and 4 were rejected, and another worker was then asked.

In questions 2 and 3, each responder had to rate each website on scales of “favorability” and “opposition” to a particular candidate, on a scale of 1 (“not at all”) to 5 (“a lot”). This rating was not done using a single scale because some websites were found to be more complex than merely favoring or supporting; many of them included different angles and opinions on the same candidate. This was the case, for instance, for news articles that had quotes from both Democrats and Republicans about the same candidate, or that recounted both controversies and accomplishments by the same person.

At the end of the crowdsourcing process, each page had three scores for favorability and three scores for opposition, which were averaged. The averages of the favorability and opposition scores were subtracted from each other to calculate an average net positivity score. Those average net positivity scores were then averaged by candidate and party.

Analyzing Net Positivity

The average net positivity score for all websites returned by the top 10 search results was 1.2 on a scale of -4 (very opposed) to 4 (very favorable). The Democrats had higher average scores (1.9) than the Republicans (1.0). The candidate with the highest average was Democrat Bernie Sanders, with 3.2. The lowest was Republican George Pataki, with 0.3.

Analysis included not only the average net positivity score but also the total number of positive and negative web pages for each search result. Any result with an average net positivity score between -0.5 and 0.5 was considered neutral. Below and above those thresholds, they were considered, respectively, negative and positive. Democrats had, on average, 7 positive search results in those top 10, whereas GOP candidates had an average of 5.9. Sanders was also the only candidate with no negative pages surfaced in his first 10 web results; the average candidate had 2.1. Sanders had 9 positive websites compared to the average of 6.1, and one neutral, whereas the average number of neutral results per candidate was 1.8. Republican Ted Cruz had the most negative results, with five, compared with five positive results.
The results for candidates tended to decline in positivity as you move down the rankings on the first page. On average, the net positivity score of the first result was 2.2. By the 10th ranked result, the average score dropped to 0.9 (Figure 13.1).

Looking closer at the individual results, the negative websites tended to appear in the bottom of the first page of Google results. Campaign websites—which included candidate webpages and social media profiles—were clustered at the top. The negative websites tended to be news articles that were mostly critical of the candidates.

Google presented a higher proportion of negative news articles about Republican candidates than their Democratic counterparts. Democrats tended to have more official sites, social platforms, and, to a lesser degree, more positive coverage on the first page. Six of the top 10 results for Bernie Sanders were run by the campaign. They were a Facebook profile, a Twitter profile, a YouTube account, and three campaign website pages. The other links were to a Wikipedia page and three news stories, all favorable to Sanders. Ted Cruz, on the other hand, had only two websites and a Facebook profile. The other seven links were news articles, and for Cruz, those were mostly negative.

The way Google ranks and selects websites, therefore, opens the way for digitally savvy campaigns to take hold of the top 10 results. Official sources of information appear to be privileged and come to dominate the most coveted and attention-getting positions on the results page.

![Figure 13.1](image)

*Figure 13.1*: The decay in the average positivity score on the first page of web results for presidential candidates in the US presidential election of 2016.

*Source*: Original chart produced by Daniel Trielli.
On February 1, 2016, Google deployed a new feature for users searching for US political information regarding the presidential election candidates (Schonberg 2016). Whenever a user searched the name of a candidate in the presidential primaries taking place at the time, the results page displayed a list of 16 political issues (later expanded to 17) each containing several position statements by the candidate. These statements were sourced from news articles that were linked alongside the quotes and, according to our contact with Google public relations the guide, would “only show issue topics from candidates when they have submitted a statement or if there is a sufficient amount of news articles that include relevant quotes from them that our algorithm has found” (see Figure 13.2).

Google’s ability to edit together and inject information that it deems relevant to candidates is another instance of how their dominance may affect the public’s ability to get fair and balanced information. This case study was motivated by a need to understand how the design of the new feature might bias how people perceive the candidates.

The analysis presented relies on data collected by observing the information box constituting the issue guide. The data collected includes quotes by the candidate, the name of the candidate, the broad issue or topic the quote is referring to, the rank position of the quote within its topic, the link to the article from which the quote was extracted for the infobox, the website that is the source for that article, and the date the article was published.

Data was collected three times, at three-week intervals, on April 1, April 22, and on May 13, 2016. In the first two collections, the tool presented quotes from five candidates: Democrats Hillary Clinton and Bernie Sanders, and Republicans Ted Cruz, John Kasich, and Donald Trump. By the time the third collection was performed, both Kasich and Cruz were removed from the list, since they had withdrawn from the campaign.

Quantifying Statements

The initial analysis involved counting the number of quotes by each candidate presented by the information box and then looking at the change across each data collection. The first discovery was that the number of quotes increased over time. No quotes were removed between each measurement. New quotes were added as the candidates made new statements.
that were covered by new articles. This increase over time happened with all candidates.

However, over time a disparity between the totals from candidate to candidate was detected. On average, candidates had 191 statements listed in the tool on April 1. But after separating by ideology, the Democratic candidates had an average of 258, while the Republicans had an average of 147. This continued through the second collection (average of 287 statements for Democrats versus 166 for Republicans) and the third, after two Republicans dropped out (300 and 160). In this final collection, Hillary Clinton had more statements than Donald Trump in 13 of the 17 topics; Donald Trump had more statements in three topics; and there was an equal number of statements in one topic.
The challenge, then, was to try to determine whether that disparity could be the result of an imbalanced input of news articles that preceded the aggregation by the information box. If the volume of coverage of the candidates was divergent, there would be fewer articles about some candidates, which would lead to fewer statements extracted and presented. Fewer news articles about Donald Trump would, therefore, indicate bias in the coverage indexed, not necessarily the Google information box curation algorithm. However, searches for each of the three remaining candidates conducted on Google News showed the opposite. In May of 2016, Google News had 941 indexed articles about Trump, 710 about Clinton, and 630 about Sanders. Data from LexisNexis corroborated that Trump was the focus of a larger number of news stories during the election (“U.S. Presidential Campaign Tracker” 2016).

**News Sources**

We also considered the distribution of sources for the articles from which statements were extracted, in order to determine whether a source or group of sources was dominant in the tool. In the May 13 collection, 326 different news organizations served as the sources for the statements of the candidates. Some sources had more articles represented than others. Out of a total of 735 statements from the three remaining candidates, 63% came from sources that, individually, had a small participation—1 to 5 articles listed in the information box. Another 24% came from sources that were in the middle tier—6 to 20 statements. And 14% of the statements came from four news sources, that is, 1.2% of the total of 326 sources: The Hill, Politico, The Washington Post, and The Huffington Post.

Additionally, Google itself was the source for 17 statements each by Bernie Sanders and Hillary Clinton. Unlike Donald Trump, Sanders and Clinton decided to use a specific Google-based solicitation for self-presentation on issues, and those were also included in the information box. This difference in how the candidates provided (or did not provide) information accounts for only about 12% of the disparity in totals we refer to in the previous section.

The next step was to determine whether there was an ideological imbalance in the sources from which the statements came. A database of ideological leanings of websites from prior research (Bakshy, Messing, and Adamic 2015) was used to determine whether there was a political bias in the news sources used in the information box. The database scores websites from -1 to 1, the negative values meaning they are preferred (by Facebook shares
and likes) by users who are left-wing and positive values meaning they are preferred by users who are right-wing.

Applying the bias scores from the study to the 83 overlapping sources used by the Google issues guide, provided an average bias score of -0.16, meaning that overall there is a slight liberal bias in news sources. However, those 83 sources cover only 50% of all the articles in the information box. When considering only the top 10 sources for the Issues tool, all of which were covered by the political bias database, the bias score was more liberal at −0.31.

Editorial Choices

Editorial choices for the Google information box were also analyzed. The first element in that inquiry was the order in which the topics were listed. Previous research has indicated that items listed higher on the page will get more attention from searchers (Agichtein et al. 2006). In the first data collection, the order of issues was: Immigration, Abortion, Guns, Foreign Policy, Taxes, Health Care, Economy and Jobs, Civil Liberties, Crime and Safety, Environment, Education, Budget and Spending, National Security, Medicare and Social Security, Veterans, and Energy. In the following data collections, it was mostly unchanged, except for the addition of Gay Marriage, in the sixth position between Taxes and Health Care. It is also noteworthy that when the results page was initially presented to the user, the list was trimmed and expandable; only the top four topics were displayed, lending even more significance to the prioritization of the first topics.

The order of topics is not alphabetical, indicating some other prioritization of the issues. A similarity with polls conducted by Gallup (Newport 2016) and the Pew Research Center (Doherty, Kiley, and Jameson 2015) and with Google Trends was investigated to see whether the order was determined by public interest or search trends, but no correlation was found. Additionally, the addition of Gay Marriage indicates a human editing process for two reasons: first, the statements shown in this category were previously listed under “Civil Liberties”; and second, Google Trends showed no spike in searches related to same-sex marriages that could explain its surge in the electorate’s interest.

In conclusion, a variety of strategies used in this inquiry surfaced the presence of biases in the Google information box, with respect to representation of candidates in their proportion of statements, dominance of sources, political bias of sources, and editorial choices. Still, the absence of algorithmic transparency (Diakopoulos and Koliska 2016) regarding how
the tool was designed and how it curated information limited our ability to establish why we observed these patterns.

**CASE 3: IN THE NEWS**

At the top of the result page for searches about newsworthy topics during the 2016 US elections Google featured a box with related news articles labeled “In the News” (later relabeled as “Top Stories” in November 2016 after the election). Because of its prominence, the Google “In the News” box had the potential to lead users to particular news sources, and to direct information seekers to particular narratives and frames around candidates, thus shading their impressions. With so much potential for guiding attention, questions of ideological diversity and news-value of the articles “In the News” become critical and motivate the current case study (Figure 13.3).

For this analysis article links were collected from Google using a non-personalized browser every hour between May 31 and July 8, 2016, resulting in a total of 5,604 links. Along with the URL for the articles, metadata was collected, including the name of the source for the article (usually a news organization), the text of the link (usually the title of the article), the string of text that shows how long the story has been posted or updated (“X hours ago”), and the exact time the link was scraped. The 5,604 links were nonunique, meaning that they could point to the same article in multiple instances. In total, there were 972 unique articles collected.

![Screen shot showing the Google “In the News” box for Donald Trump, collected October 17, 2016.](source: Screenshot of webpage taken and cropped by authors.)

**Figure 13.3:** Google “In the news” box for Donald Trump, collected October 17, 2016.

*Source: Screenshot of webpage taken and cropped by authors.*
To determine whether particular sources dominated the Google “In the News” box, the total number of links collected (including repetition) from each website (i.e., top level domain) was measured against the total number of links displayed. Out of the 113 news sources, 60 had articles linked nine or fewer times during the time frame of collection. Another 35 sources had between 10 and 50 links displayed in the Google “In the News” box. Sixteen sources had between 50 and 500 links showcased. The remaining two sources combined, CNN and New York Times, account for 2,476 links (44.2%) listed. Of the 5,604 links listed in the Google “In the News” box, 1,276 came from CNN and 1,200 originated from the New York Times domain. For example, between June 2 and June 6, an article titled “Hillary Clinton’s Evisceration of Donald Trump” appeared in 90 different hourly measurements, indicating a strong staying power and emphasis in Google’s results.

However, the sources that are most frequent in the Google “In the News” boxes do not rely only on single articles that are repeated multiple times. In fact, when only unique links are considered, CNN and the New York Times also have high prevalence in the ranking. Out of the 972 unique articles displayed, 152 belonged to the New York Times and 142 to CNN, meaning that, combined, they amounted to 30% of the unique articles. In third place, NBCNews.com had 69 articles, or 7% of the total unique links.

The Google “In the News” box also prioritizes the three links that it presents. There are always three links that are displayed in the tool, but the first in that list has more prominence due to additional information displayed. All of the first listed links include a short summary, or snippet of the article, and 99.4% of them also include an image. The distribution of sources in those first links listed was also measured, to investigate whether or not some sources also had more dominance in the premium spot. Again, CNN and the New York Times dominate the first listed link, with 1,211 out of 1,868, or 64.4% of them being from these two sources. This is a startling concentration of attention oriented toward only two publishers of information.

When evaluated on a weekly basis, some variation in the ranking of the sources for articles was detected. The most clear example is the Washington Post, a news organization that was in the top five sources of links for the week of June 12 to June 18, when it rose to third place.

One hypothesis is that the Washington Post had a temporarily higher profile than usual, since it became the focus of a controversy concerning the presidential election coverage. The increase of the newspaper in the Google “In the News” box rankings coincided with the then-candidate
Donald Trump criticizing the Post on Twitter and revoking their credentials for campaign events. However, more research is necessary to conclude that becoming the focus of news itself could itself propel a news organization to higher prominence in the box.

**Freshness**

Using the metadata relating to time, we found that “freshness” of articles also appears to be a relevant factor for prevalence in the “In the News” box. When all articles are considered, 30.5% of all links are marked as being less than three hours old. A smaller proportion, 4.8%, of links listed were more than one day old, and occurred mostly on weekends. Additional research is needed to determine the specifics of how Google’s algorithms considers article timestamps and update times. For example, the article “Hillary Clinton’s Evisceration of Donald Trump,” may have had staying power because it was continually updated and considered “new” by Google because of an updated timestamp.

**Non-News Organizations**

Not all linked articles came from news organizations. Of the 5,604 total links, 214 came from Twitter (the official accounts of Hillary Clinton, Donald Trump, and Chelsea Clinton were the originators of those links), 58 from YouTube, and 24 from the Federal Bureau of Investigation. Also included in the list of sources are websites that are controversial sources of news, such as Breitbart News, Infowars, and Russia Today.

It is important to note that the Google “In the News” box shifted from a strict news organization background to a more open aggregator of news-related links in 2014. Until that year, the box was called “news for . . .”; then, it changed to “in the news.” That was the year in which users started noticing non-news links appearing in the box, such as press releases from companies and content from Reddit and YouTube.

The results presented earlier can lead to different threads of inquiry. The analysis sheds light on the disparity in attention that some sources of news achieve via the box, but it still remains unclear why such disparities came about. There is a clear predilection toward fresher and newer content, privileging the daily news cycle or at least the appearance of it via the updated publication date of articles. The inclusion of non-news links is a complicating factor, raising questions about Google’s very definition of a news source.
CASE 4: VISUAL FRAMING

The first three case studies presented focus on textual results from search engines such as links to articles and websites. However, Google also presents images on the main results page. Google’s image box is located in the top right section of the main results page, and typically contains five to seven images. The box draws attention to specific images of the candidates, and to the articles or information linked via these images. In this way, Google may contribute to shaping users’ attitudes and perceptions by visually framing candidates in particular ways.

Previous work suggests that the visual portrayal of candidates in election campaigns may have an impact on electoral outcomes. For instance, the competence of a candidate is inferred from attractiveness (Todorov et al. 2005), and beauty itself is positively associated with votes (Berggren, Jordahl, and Poutvaara 2010). Furthermore, people are unlikely to change their candidate pick even after they are provided additional information regarding candidates’ competencies (Todorov et al. 2005).

In this case we examine how Hillary Clinton and Donald Trump were framed in the images surfaced on search results pages by Google including analyses of emotional and gestural content, the sources of images and their political leanings, and image rank positions. Images, their sources, and the ranking position (i.e., first, second, etc.) were collected for the queries “Hillary Clinton” and “Donald Trump” once per day from September 3 until October 28, 2016. As a baseline, we also collected images from Google Image Search for each candidate for the same time period, resulting in 353 images of Clinton and 298 of Trump after removing images that did not contain the candidate, or that contained multiple faces.

In our sampling period there were nine unique images from the image box of Clinton, and 11 of Trump. Unique images were identified using a difference hash algorithm (Buchner 2017) which converts the images to grayscale, reduces them to 72 pixels, and then computes the difference between neighboring pixels producing an alphanumeric hash unique to that image. To analyze the emotional content of the images, we used an API (Microsoft Azure 2017), which provided a confidence score for each of eight emotions, comprising anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise.

Clinton was deemed happy in seven images, neutral in one, and surprised in another, whereas Trump was happy in three images, neutral in seven, and angry in one. This distribution of emotions was somewhat different from the distribution observed in Google Image search baseline images, where Clinton exhibited a neutral expression in the majority of
In terms of gestural content, Clinton was gesturing in only one of nine image box images (11%) versus 40% of baseline images, while Trump was gesturing in four of 11 image box images (36%) versus 57% of baseline images (Figure 13.4).

Altogether, these images frame the candidates in a manner typical of representation of women and men in news media (Rodgers, Kenix, and Thorson 2007; Kwak and An 2016), with Clinton portrayed as happy and more passive (fewer gestures), and Trump depicted as serious, and more energetic (more gestures). Energetic gesturing is associated with dominance and power (Burgoon and Dunbar 2006), and effective leadership is typically associated with masculine traits including dominance (Koenig et al. 2011). Therefore, these images may reinforce the gender stereotype that men are strong leaders, while women, who display characteristics incongruent with typical leadership traits, are ineffective.

Because images are hyperlinked, Google’s choice of images also serves to direct attention to specific sources of those images. Research shows that people show more interest in unconventional news photos when presented outside the context of the article (Mendelson 2001). Such atypical images in Google’s image box would be particularly powerful in directing attention to the related articles as they are also presented outside the context of the source article. This could have an impact on voters, as outlets may portray candidates differently, and may impart partisan bias within the images that are published. Sources for Clinton’s baseline images were mostly left-leaning, but left-leaning sources of image box images were even

![Figure 13.4](image_url)  
**Figure 13.4:** Percentage of Clinton and Trump images from the image box or from Google Images (Baseline) that show listed emotions.  
*Source: Original chart produced by Jennifer Stark.*
more highly represented (Stark and Diakopoulos 2017). While sources of Trump’s baseline and image box images were also mostly left-leaning, centrist sources were reduced and right-leaning sources were augmented in the image box compared with the baseline (Figure 13.5).

In terms of preferential position of the images within the image box, Google+ was always in the first ranking position, which affords the largest image, meaning that while Google is privileging their own social media platform, knowing this, the campaigns themselves can control the visual in the most dominant position. Clinton’s Wikipedia image was in second position, suggesting it is second in importance. Trump’s Wikipedia page was second also, but after several edits to the Wikipedia page’s main picture, this image was dropped altogether. The remaining rank positions of images for both candidates were less stable. Notably for Clinton, many images changed rank between the 3rd and 24th of September. During this time Clinton was reported as suffering from pneumonia, apologized for calling Trump supporters “deplorables,” and suggested Russia was intervening in the election process (Figure 13.6). The news cycle does appear to impart some variability in the visual portrayal of the candidates after the top image or two, but does not seem to be the only driver, given that the primary rank positions were independent from news. Altogether, Google’s images are selected from sources across the ideological spectrum, and prioritize images from its own social media site Google+.

There are several limitations with this case study. Our baseline consists of images collected from Google Images search results. Although we

![Image of bar charts showing Hillary Clinton and Donald Trump image sources from left, center, and right ideologies, with percentages for image box and baseline.](chart)

**Figure 13.5:** Percentage of sources for Clinton and Trump images from the image box or Google Images baseline identified from each listed ideology. *Source: Original chart produced by Jennifer Stark.*
expected this baseline to represent a universe of Google images from which Google selects images for the image box, not all images found in the image box were also in the baseline. Alternative baselines may be more representative of all images that exist for each candidate, for example, using a longer search time period, or images collected from additional search engines like Bing and Yahoo! Moreover, the bias of photographers and editors in selecting images that feed indices like Google Images is not addressed. We also relied on computer vision algorithms for emotional content information, which in turn are trained on data with its own unspecified sampling bias.

Open questions include how the images are selected for the image box, and the role that algorithms or human editors may play in those selections as well as whether source bias would be the same for a different time period, or for politicians in different countries. Our results suggest that Google’s image results reflect visual gender differences with respect to leadership roles present in society: that women are happy and passive and men are serious and active. An open question is whether such gendered visual frames, from search engines in particular, can impact perception of presidential competency.

**DISCUSSION**

Through a series of case studies this chapter has begun to characterize the range of influences that Google’s search technologies were having on information curation during the 2016 US elections. Results clearly show the myriad editorial choices that Google’s algorithms make in shaping the information environment, including a focus on official sources in the main rankings, ordering in the issue guide that may have unduly privileged
certain issues, the dominance of a small set of news sources highlighted “In the News,” and differences in the visual framing of a female versus a male candidate. These results constitute a set of observations that raise important questions for future work, and suggest political conversations that need to be undertaken about the editorial role of search engines in political life: What should be the responsibilities of Google as an important intermediary of information for voters?

Studying search engine results is a methodologically challenging undertaking. Though we were able to observe several instances of how results may shift attention toward candidates we lacked any form of scientific control, or ability to run experiments, and found it difficult to explain why we were seeing many of the results patterns that we saw. Still, we believe that it is a valuable public interest endeavor to report on the observed patterns of information that dominant entities like Google mediate, given that they can have a substantial effect on exposure to political information and bias.

For the Issue Guide case we found it valuable to consider the dynamics of the results from Google and we believe this should be an important aspect that informs the methodologies of algorithm audits of search engines in future work. While taking individual snapshots of data allows for concrete analyses, the dynamics of how the results change over time (and how the search giant may be responding to public pressure) is important to track for accountability purposes. Writing “stories” about results may be augmented by building dashboards to track and reflect search results over time.

In two of the cases, for the Issue Guide and for Visual Framing, the concept of defining baselines emerged. Defining an appropriate baseline dictates what the expectation for an algorithm should be, thus informing what is perceived as interesting and newsworthy for an audit. But what is the “right” baseline? In our work we made logical arguments about the expected input data sets that Google algorithms would use to curate quotes or images, however other reasonable baselines could be considered. Some may consider results less compelling if the baseline does not come from a sample independent of the platform under study. We believe that the public impact of these types of audits hinges on making a strong and well-reasoned claim for the expected information that a search engine should provide, and then showing that the search engine meets that expectation or violates that expectation. Additional algorithmic transparency information could inform the baseline definition process, and more generally may be considered by regulators exploring how targeted transparency policies could balance the dominance of information intermediaries.

Several of the case studies presented focused on characterizing the sources that Google surfaces in results. We think diversity is an important
frame for considering the information that voters have about candidates. For instance, although Google predominantly operates on a logic of relevance, the idea of information diversity could help to mitigate issues of political polarization that challenge society. In these cases we have considered the ideological diversity present in sources, however other definitions of diversity are also possible, such as by looking at ownership diversity (i.e., who owns a particular source), and even topical diversity (i.e., the topical frames that are apparent in information surfaced by search engines). Future work might be usefully oriented toward building community resources that can reliably tag sources according to these dimensions of diversity so that audits can easily incorporate various diversity measures.

ACKNOWLEDGMENTS

This work has been supported by grants from the Tow Center for Digital Journalism and from the VisMedia project at the University of Bergen. We have also benefited through discussions with professor Susan Moeller at the University of Maryland, and from the feedback we received from Torie Bosch, who edited the first three case studies when they appeared as articles published in Slate.

REFERENCES


