The VP of Google News also stated that "Google does not make editorial decisions about which stories to show" and that "our primary approach is to use technology to reflect the news landscape, and leave editorial decisions to publishers." To prevent fake news from running rampant on the platform, Google says "Our algorithms are designed to elevate news from authoritative sources." Very little has been said publicly about what this really means and how it is accomplished. All I could find on Google's official website that supposedly describes how trustworthy news is elevated[34] is that the algorithms rely on signals that "can include whether other people value the source for similar queries or whether other prominent websites on the subject link to the story."

Alas, it doesn't seem that much progress has been made toward uncovering the state of fake news on Google News and the company's efforts to limit it. However, the official explanatory site for Google News also states that "Our ranking systems for news content across Google and YouTube News use the same web crawling and indexing technology as Google Search," so it seems the time is right to turn now to this chapter's lengthiest section: the role Google search plays in the dissemination of fake news.

Google Search

When we search for information on Google, the results that come up—and the order they are presented in—shape our views and beliefs. This means that for Google to limit the spread of misinformation, it must find ways of training its algorithms to lift quality sources to the top without taking subjective, biased perspectives on contentious issues and also without impinging on people's ability to scour the depths of the Web. There are many pieces of this story. In this section, I will present evidence backing up the assertion that search result rankings affect individuals' worldviews; look into what factors Google's search algorithm uses to decide how to rank links; illustrate how featuring highlights from top searches has led to problematic misinformation; show how authentic links have been removed from Google through deceptive means; introduce the deep learning language model Google recently launched to power its search and many other tools; and, finally, discuss Google's efforts to elevate quality journalism in its search rankings.

[34] https://newsinitiative.withgoogle.com/hownewsworks/mission.

Ranking Matters

In August 2015, a study[35] of the impact Google search rankings have on political outlook was published in the prestigious research journal *Proceedings of the National Academy of Sciences*. One of the main experiments in this study was the following. Participants were randomly placed in three different groups. The participants were all provided brief descriptions of two political candidates, call them A and B, and then asked how much they liked and trusted each candidate and whom they would vote for. They were then given fifteen minutes to look further into the candidates using a simulated version of Google that only had thirty search results—the same thirty for all participants—that linked to actual websites from a past election. After this fifteen-minute session, the participants were asked the same questions as before about the two candidates. The key was that one group had the search results ordered to return the results favorable to candidate A first, another group had results favorable to candidate B first, and for the third group the order was mixed.

The researchers found that on all measures, the participants' views of the candidates shifted in the direction favored by the simulated search rankings, by amounts ranging between thirty-seven and sixty-three percent. And this was just from a single fifteen-minute search session. The researchers also experimented with a real election—two thousand undecided votes in a 2014 election in India. The authors stated[36] that "Even here, with real voters who were highly familiar with the candidates and who were being bombarded with campaign rhetoric every day, we showed that search rankings could boost the proportion of people favoring any candidate by more than 20 percent—more than 60 percent in some demographic groups."

The researchers go on to boldly suggest that "Google's search algorithm, propelled by user activity, has been determining the outcomes of close elections worldwide for years, with increasing impact every year because of increasing Internet penetration." I find this assertion to be a stretch—at least, the evidence to really back it up isn't in their PNAS paper—but my focus in this book is not political bias and elections, it is fake news. And the researchers here did convincingly establish that search rankings matter and affect people's views, which means there are real consequences when Google places fake news links highly in its search rankings.

[35] Robert Epstein and Ronald Robertson, "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections," *Proceedings of the National Academy of Sciences (PNAS)*, August 18, 2015, 112 (33), E4512-E4521: https://doi.org/10.1073/pnas.1419828112.

[36] Robert Epstein, "How Google Could Rig the 2016 Election," *Politico*, August 19, 2015: https://www.politico.com/magazine/story/2015/08/how-google-could-rig-the-2016-election-121548/.

Dylann Roof, the neo-Nazi who in 2015 murdered nine Black people at a church in Charleston, South Carolina, wrote a manifesto[37] in which he claims to have been inspired by Google. He described how he typed "black on White crime" into Google, and he has "never been the same since that day." He says the first site this search produced was for an organization called the Council of Conservative Citizens (CCC). It contained numerous descriptions of "brutal black on White murders." He alleges that seeing this made him question the media's attention on the Trayvon Martin case, and it motivated him to pursue a self-education journey on Google on racial matters—a journey that, as we now know, ended in unfathomable tragedy.

While this is anecdotal evidence, it is nonetheless quite unsettling and shows one of the dangers of letting algorithms decide what information we should see, and in what order. As UCLA professor Safiya Noble pointed out in her book *Algorithms of Oppression* critiquing Google, the top result for Roof's search should have been an authoritative source such as FBI crime statistics—which shows that most violence against white Americans is committed by white Americans—rather than the CCC, an organization whose Statement of Principles says that it "opposes all efforts to mix the races of mankind."

Signals the Algorithm Uses

What factors determine how highly ranked pages are in Google searches? Once again, Google won't reveal much about its algorithmic trade secrets—in part to prevent competitors from copying them, but also to prevent people from gaming the algorithm—so we only know the broadest outlines. The official company website describing the search algorithm[38] states the following: "Search algorithms look at many factors, including the words of your query, relevance and usability of pages, expertise of sources, and your location and settings. The weight applied to each factor varies depending on the nature of your query—for example, the freshness of the content plays a bigger role in answering queries about current news topics than it does about dictionary definitions." These factors are largely about finding links that appear to be good matches to the search query. When it comes to ranking the results, Google says the algorithm attempts to "prioritize the most reliable sources available" by considering factors that "help determine which pages demonstrate expertise, authoritativeness, and trustworthiness on a given topic." This sounds good, but it's quite vague. The two examples Google gives are that a

[37]Daniel Strauss, "Racist manifesto linked to Dylann Roof emerges online," *Politico*, June 20, 2015: `https://www.politico.com/story/2015/06/dylan-roofs-racist-manifesto-emerges-online-119254`.
[38]`https://www.google.com/search/howsearchworks/algorithms/`.

site is bumped up in the rankings if other prominent sites link to it (this is the essence of the original PageRank algorithm Google first launched with in 1998) or if many users visit the site after doing closely related searches.

Earlier in the history of the algorithm, the PageRank method played a more prominent role, and less attention was given to assessing the quality of information by other means. Nefarious actors figured out how to use this narrow focus on link counting to manipulate the rankings. In December 2016, it was reported[39] that fake news and right-wing extremist sites "created a vast network of links to each other and mainstream sites that has enabled them to game Google's algorithm." This led to harmful bigotry and disinformation—for instance, eight of the top ten search results for "was Hitler bad?" were to Holocaust denial sites.

Several months later, in April 2017, Google admitted[40] that fake news had become a serious problem: "Today, in a world where tens of thousands of pages are coming online every minute of every day, there are new ways that people try to game the system. The most high profile of these issues is the phenomenon of 'fake news,' where content on the web has contributed to the spread of blatantly misleading, low quality, offensive or downright false information." One change Google implemented at the time was to provide more detailed guidance on "misleading information, unexpected offensive results, hoaxes and unsupported conspiracy theories" for the team of human evaluators the company uses to provide feedback on the search algorithms: "These guidelines will begin to help our algorithms in demoting such low-quality content and help us to make additional improvements over time." I'll return to these human moderators and the role they play in shaping Google's search algorithm at the end of this section. Google also said that the signals used in the algorithm were adjusted in order to "help surface more authoritative pages and demote low-quality content," but no details were provided.

One year later, as part of a multipronged effort to "elevate quality journalism" (including a three-hundred-million-dollar outreach initiative[41]), Google tweaked the algorithm again. This time, it revealed which signals were prioritized in this

[39]Carole Cadwalladr, "Google 'must review its search rankings because of rightwing manipulation'," *Guardian*, December 5, 2016: `https://www.theguardian.com/technology/2016/dec/05/google-must-review-its-search-rankings-because-of-rightwing-manipulation`.
[40]Ben Gomes, "Our latest quality improvements for Search," *Google blog*, April 25, 2017: `https://blog.google/products/search/our-latest-quality-improvements-search/`.
[41]Kevin Roose, "Google Pledges $300 Million to Clean Up False News," *New York Times*, March 20, 2018: `https://www.nytimes.com/2018/03/20/business/media/google-false-news.html`.

adjustment and why, at least in very broad strokes:[42] "To reduce the visibility of [harmful misinformation] during crisis or breaking news events, we've improved our systems to put more emphasis on authoritative results over factors like freshness or relevancy."

Featured Snippets

In 2014, Google introduced a tool called *featured snippets*: a box of text is placed at the top of the results page for certain searches containing a highlighted passage from a top link that the algorithm believes is relevant to the search. These are not included for all searches, just ones that Google thinks are asking for specific information that it can try to find on the Web. Featured snippets help people extract information quickly from the internet since you get the answers to your questions directly from a Google search without having to choose a link, click it, then find the relevant passage buried somewhere on that site; they are also useful for voice search on mobile devices and Google's Home Assistant because the user can ask a question verbally and then the device responds verbally by reading aloud the featured snippet resulting from the corresponding Google search.

But if a Google search turns up fake news, then the answer provided by Google in the featured snippet might be wrong. And the featured snippet format strips the answer of any context and presents it in an authoritative-sounding manner, leaving the reader/listener even less aware than in a typical Google search of how unreliable the information source might be. Sure enough, featured snippets eventually made headlines for providing disastrously misinformed answers to a variety of questions.

Indeed, in March 2017, it was found[43] that asking Google which US presidents were in the KKK resulted in a snippet falsely claiming that several were; asking "Is Obama planning a coup?" yielded a snippet that said "According to details exposed in Western Center for Journalism's exclusive video, not only could Obama be in bed with the Communist Chinese, but Obama may in fact be planning a Communist coup d'état at the end of his term in 2016!"; searching for a gun control measure called "Proposition 63" yielded a snippet falsely describing it as "a deceptive ballot initiative that will criminalize millions of law abiding Californians." In the case of the Obama coup snippet, the top search result was an article debunking this fake news story about an upcoming coup attempt, but when using Google's Home Assistant, there are no search results listed—all one gets is the featured snippet read aloud.

[42] Richard Gingras, "Elevating quality journalism on the open web," *Google blog*, March 20, 2018: https://blog.google/outreach-initiatives/google-news-initiative/elevating-quality-journalism/.
[43] Rory Cellan-Jones, "Google's fake news Snippets," *BBC News*, March 6, 2017: https://www.bbc.com/news/technology-39180855.

In a blog post[44] a year later describing how the snippets tool had improved over time, the company said "Last year, we took deserved criticism for featured snippets that said things like [...] Obama was planning a coup. We failed in these cases because we didn't weigh the authoritativeness of results strongly enough for such rare and fringe queries." While this claim is superficially true, it conveniently sweeps under the rug that a prerequisite to *weighing* authoritativeness is *measuring* authoritativeness, which is a challenging, fraught issue—one I'll return to at the end of this section. Reducing misinformation by elevating quality sources is not nearly as simple as adjusting a lever labeled "authoritativeness" the way Google suggests here in this remark.

Blocking Search Results

Some fake news publishers scrape articles from the Web and repost them as their own in an effort to give the stories a wider platform and to collect ad revenue in the process. While most fake news on the Web doesn't violate any Google policies that would prevent it from being eligible to show up on search results, if an article is found to be in violation of copyright, then Google will expunge it from search listings—so these spammy fake news sites indeed get delisted.

But an extensive *Wall Street Journal* investigation[45] found an unsettling twist here: people have been gaming Google's copyright infringement request system in order to delist content that is unflattering or financially impactful to certain parties. One of the techniques used is *backdating*: someone copies a published article and posts it on their blog but with a misleading time stamp to make it appear that it predates the published article; then they tell Google that the published article is violating their blog's copyright, and the published article is removed from Google search results. When this happens, Google is delisting an actual news article on the basis of a false copyright infringement notification. Daphne Keller, a former Google lawyer and currently a program director at Stanford University's Cyber Policy Center, said that "If people can manipulate the gatekeepers to make important and lawful information disappear, that's a big deal." The *Wall Street Journal* found that not only can people indeed do this, but they have been doing so in surprisingly large numbers.

[44] Danny Sullivan, "A reintroduction to Google's featured snippets," *Google blog*, January 30, 2018: https://blog.google/products/search/reintroduction-googles-featured-snippets/.
[45] Andrea Fuller, Kirsten Grind, and Joe Palazzolo, "Google Hides News, Tricked by Fake Claims," *Wall Street Journal*, May 15, 2020: https://www.wsj.com/articles/google-dmca-copyright-claims-takedown-online-reputation-11589557001.

Google received copyright removal requests for fewer than one hundred thousand links between 2002 and 2012, whereas it now routinely handles more than a million requests in a single day. In order to scale up to this magnitude, a company spokesperson said that Google has automated much of the process so that human review is mostly not needed. In 2019, eighty percent of the nearly one-quarter billion links flagged for copyright infringement were removed from Google's search listings. However, after the *Wall Street Journal* uncovered numerous cases of fraudulent violation notifications, Google restored more than fifty thousand links that had been removed.

One of the clusters of fraudulent requests the *Wall Street Journal* found concerned Russian-language news articles critical of politicians and business leaders in Ukraine. These articles were taken off Google after various organizations including a Russian edition of *Newsweek* filed a copyright violation request—but it turned out these organizations were all fake, the supposed Russian *Newsweek* had nothing to do with *Newsweek*, it was just using *Newsweek's* logo to deceive Google into thinking the copyright violation notification was legitimate.

There is a secondary harm to these deceptive methods for tricking Google into delisting real news articles: in Google's recent efforts to elevate quality journalism, one factor the ranking algorithm considers is the number of copyright violations sites have received. A Google spokesperson said that "If a website receives a large number of valid takedown notices, the site might appear lower overall in search results." But the *Wall Street Journal* investigation established that many of the takedowns that Google thinks are valid are actually invalid and the result of deliberate disinformation aimed at Google's automated system. This opens the door for more gaming of Google's rankings by dishonest actors.

BERT

A humorous cultural trend emerging in the AI community over the past few years has been to make as many *Sesame Street* allusions as possible when naming deep learning text processing algorithms. You saw *Grover* in Chapter 2, the system developed by the Allen Institute for AI to generate text like GPT-2 with the ultimate goal of being able to detect such deep learning generated text. Currently, the two most impressive and powerful deep learning systems for text are GPT-3, which was discussed extensively in Chapter 2, and a system developed by Google called *BERT* (which stands for *Bidirectional Encoder Representations from Transformers*, but don't worry about that just yet). BERT builds on an earlier system developed by the Allen Institute for AI named *ELMo* (standing for *Embeddings from Language Models*). Alas, there is not yet an Ernie or a Snuffleupagus.

While text generation is immensely useful, it turns out that for many applications one needs something that is essentially a by-product of the inner workings of a deep neural net that occurs automatically while training for tasks like text generation: a *vector representation* of words (sometimes called a *word embedding*). This means a way of representing each word as a vector—that is, a list of numerical coordinates—in such a way that the geometry of the distribution of word vectors reflects important semantic and syntactic information. Roughly speaking, we want words that frequently appear in close proximity to each other in written text to have vector representations that are geometrically in close proximity to each other. Vector embeddings translate data in messy formats like text into the standard numerical formats that machine learning algorithms know and love.

Earlier word embeddings (one of the most popular, called *Word2vec*, was developed by Google in 2013) produced a fixed, static vector for each word. This opened the door to many breakthroughs: for example, analyzing the sentiment of words and sentences (how positive or negative they are) turns into a more familiar geometric analysis of vectors in Euclidean space, where, for instance, one looks for a plane that separates positive word vectors from negative word vectors. One of the key drawbacks in these early static approaches was that a single word can have multiple meanings (such as "stick" a landing, "stick" from a tree, and "stick" to a plan), and all the different meanings got conflated when the word was represented as a vector. In contrast, ELMo and BERT are *contextual* word embeddings, which means the vector representations are not fixed and static—they depend on the surrounding text. If you feed these systems the sentence "I hope the gymnast sticks the landing" and the sentence "the toddler sticks out her tongue," the word "sticks" will have different vector representations in each case. This allows for much more flexibility in language modeling and understanding.

BERT learns its contextual word embeddings through a massive self-supervised pre-training process somewhat similar to that of GPT-3. As you may recall from Chapter 2, GPT-3 was fed huge volumes of text, and as it read through this, it used the preceding words to try to predict the next word. BERT's self-supervised training process also involved reading massive volumes of text, but in this case a percentage of the words were randomly masked (hidden from the algorithm). BERT learned how to "predict" these missing words by guessing what they were and then unmasking them and using the difference between the guess and the actual unmasked word as an error to propagate through the neural net and adjust all the parameters so that over time the guesses become more accurate. In this way, BERT was trained to predict missing words. BERT's ultimate goal is not to predict words, but being able to predict words well is considered a good proxy for understanding them; or, from a more technical perspective, the contextual vector embeddings BERT develops internally while training on hidden word prediction turn out to be very useful for a wide range of linguistic tasks.

Actually, this masked word task helps BERT learn about words in the context of each sentence, but to get a more global perspective, it is simultaneously trained on a self-supervised sentence prediction task: it is shown pairs of sentences from the training text and learns to estimate the probability that one sentence immediately precedes the other. This task helps the word embeddings encode larger-scale meaning that extends beyond individual sentences. In case you are curious, the *Bidirectional* in BERT's name refers to the fact that it reads training text both left-to-right and right-to-left in order to get both past and future context for each word. This is a reasonable thing to do precisely because, unlike GPT-3, BERT is not aiming to predict future words—it is aiming to produce word embeddings that draw in as much context as possible. The *Encoder Representations* part of the name just indicates that words are encoded with vector representations, and the *Transformer* in the name refers to a specific deep learning architecture that is used.

When you type a search phrase into Google, you are providing more than just a list of keywords to match—often you are providing a grammatical snippet of text that Google's search algorithm needs to understand. BERT is the intermediary service that translates your search phrase into a collection of vectors that the search algorithm can then process quantitatively. In an October 2019 company blog post announcing the absorption of BERT into Google's search algorithm, Google's vice president of Search said[46] that "Search is about understanding language," and BERT has indeed been one of the most successful steps forward in allowing computers to better understand human language.

This Google blog post went on to say that "when it comes to ranking results, BERT will help Search better understand one in 10 searches in the U.S. in English" and that "Particularly for longer, more conversational queries, or searches where prepositions like 'for' and 'to' matter a lot to the meaning, Search will be able to understand the context of the words in your query." To illustrate the types of improvements users should expect, the blog post included an example of a user searching "Can you get medicine for someone pharmacy." Previously, the top search result was an article that included each of these individual words, but it didn't answer the question the user was attempting to ask; with the new BERT-powered search, the top result was an article specifically addressing when and how people can pick up medications for others at the pharmacy.

Google evidently underestimated itself with the "one in 10" figure, because almost exactly one year later in another company blog post[47] Google declared

[46]Pandu Nayak, "Understanding searches better than ever before," *Google blog*, October 25, 2019: `https://blog.google/products/search/search-language-understanding-bert/`.
[47]Prabhakar Raghavan, "How AI is powering a more helpful Google," *Google blog*, October 15, 2020: `https://www.blog.google/products/search/search-on/`.

that "BERT is now used in almost every query in English." And in September 2020, Google announced[48] that BERT was also being used "to improve the matching between news stories and available fact checks." In Chapter 9, I'll cover fact-checking tools in depth; for now, what's relevant here is that BERT is used to automatically scan through lists of human fact-check reports and figure out which ones pertain to a given news article.

But BERT also featured prominently in a recent debacle that landed Google in the news in an unflattering light. Stanford-trained computer scientist Timnit Gebru is one of the world's leading experts on ethics in AI and algorithmic bias; she is a cofounder of the organization *Black in AI*; and, until recently, she was one of the leaders of Google's Ethical Artificial Intelligence Team. But she was abruptly fired from Google in December 2020.[49] She was working on a research paper with several other Google employees when a request from the higher-ups came in asking her to either withdraw the paper or remove the names of all Google employees from it. She refused and demanded to know who was responsible for this bizarre request and their reasoning behind it, but Google leadership rebuffed her demand and instead fired her.

This move was shocking, not just to Gebru but to the entire AI community. Gebru is widely respected and recognized for important pioneering work; Google's iron-fisted handling of this incident did not sit well with most people. The optics of Google censoring and then firing a prominent and beloved Black woman AI researcher from a leadership role on an ethics team were very poor, to say the least. And what was the topic of Gebru's research paper that stirred up all this controversy in the first place? The paper was titled "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?", and it was on the dangers—ranging from environmental costs to inscrutability to harmful biases—involved in large deep learning language models like BERT.

We now come to the final topic of this lengthy section, and of the entire chapter, which is how Google has been trying to adjust its algorithm so that accurate information rises to the top of search rankings and fake news is relegated to later pages of search results.

[48]Pandu Nayak, "Our latest investments in information quality in Search and News," *Google blog*, September 2020: `https://blog.google/products/search/our-latest-investments-information-quality-search-and-news`.
[49]Bobby Allyn, "Google AI Team Demands Ousted Black Researcher Be Rehired And Promoted," *NPR*, December 17, 2020: `https://www.npr.org/2020/12/17/947413170/google-ai-team-demands-ousted-black-researcher-be-rehired-and-promoted`.

Elevating Quality Journalism

In February 2019, at an international security conference in Munich, Google released a white paper[50] on the company's efforts "to tackle the intentional spread of misinformation—across Google Search, Google News, YouTube and our advertising systems." While mostly repeating the philosophies and general approaches that were already sprinkled across various Google blog posts and corporate documents (and already mentioned in this chapter), this white paper does include a few remarks and insights that helpfully shine some additional light on certain details.

In attempting to elevate quality journalism, Google's search ranking algorithm needs to assess the trustworthiness of news sites. The white paper clarifies that these assessments are not just overall measures, they depend specifically on the scope of the search phrase: "For instance, a national news outlet's articles might be deemed authoritative in response to searches relating to current events, but less reliable for searches related to gardening." It also clarifies that the "ranking system does not identify the intent or factual accuracy of any given piece of content." In other words, everything—true and false—is allowed to show up on Google, and Google's search algorithm does not try to determine which particular links contain valid information versus misinformation; instead, it just tries to rank more highly the sites it deems more generally trustworthy in the context of the present search.

In June 2020, the Parliament of the United Kingdom published a policy report[51] with numerous recommendations aimed at helping the government fight against the "pandemic of misinformation" powered by internet technology. The report is rather forceful on the conclusions it reaches:

> The Government must make sure that online platforms bear ultimate responsibility for the content that their algorithms promote. [...] Transparency of online platforms is essential if democracy is to flourish. Platforms like Facebook and Google seek to hide behind 'black box' algorithms which choose what content users are shown. They take the position that their decisions are not responsible for harms that may result from online activity. This is plain wrong. The decisions platforms make in designing and training these algorithmic systems shape the conversations that happen online.

[50]Kristie Canegallo, "Fighting disinformation across our products," *Google blog*, February 16, 2019: https://www.blog.google/around-the-globe/google-europe/fighting-disinformation-across-our-products/.

[51]"Digital Technology and the Resurrection of Trust," *House of Lords*, Select Committee on Democracy and Digital Technologies, Report of Session 2019–21: https://committees.parliament.uk/publications/1634/documents/17731/default/.

While preparing this report, Parliament collected oral evidence from a variety of key figures. One of these was Vint Cerf, Vice President and Chief Internet Evangelist at Google. He was asked: "Can you give us any evidence that the high-quality information, as you describe it, that you promote is more likely to be true or in the category, 'The earth is not flat', rather than the category, 'The earth is flat'?" His intriguing response provided a sliver of daylight in the tightly sealed backrooms of Google:

> The amount of information on the world wide web is extraordinarily large. There are billions of pages. We have no ability to manually evaluate all that content, but we have about 10,000 people, as part of our Google family, who evaluate websites. We have perhaps as many as nine opinions of selected pages. In the case of search, we have a 168-page document given over to how you determine the quality of a website. [...] Once we have samples of webpages that have been evaluated by those evaluators, we can take what they have done and the webpages their evaluations apply to, and make a machine-learning neural network that reflects the quality they have been able to assert for the webpages. Those webpages become the training set for a machine-learning system. The machine-learning system is then applied to all the webpages we index in the world wide web. Once that application has been done, we use that information and other indicators to rank-order the responses that come back from a web search.

He summarized this as follows: "There is a two-step process. There is a manual process to establish criteria and a good-quality training set, and then a machine-learning system to scale up to the size of the world wide web, which we index." Many of Google's blog posts and official statements concerning the company's efforts to elevate quality journalism come back to this team of ten thousand human evaluators, so to dig deeper into Cerf's dense statement here, it would be helpful to better understand what these people do and how their work impacts the algorithm. Fortunately, an inside look at the job of the Google evaluator was provided in a *Wall Street Journal* investigation[52] from November 2019.

[52]Kirsten Grind, Sam Schechner, Robert McMillan, and John West, "How Google Interferes With Its Search Algorithms and Changes Your Results," *Wall Street Journal*, November 15, 2019: https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753.

While Google employees are very well compensated financially, these ten thousand evaluators are hourly contract workers who work from home and earn around $13.50 per hour. One such worker profiled in the *Wall Street Journal* article said he was required to sign a nondisclosure agreement, that he had zero contact with anyone at Google, and that he was never told what his work would be used for (and remember these are the people Cerf referred to as "part of our Google family"). He said he was "given hundreds of real search results and told to use his judgment to rate them according to quality, reputation and usefulness, among other factors." The main task these workers perform, it seems, is rating individual sites as well as evaluating the rankings for various searches returned by Google. These tasks are closely guided by the hundred-sixty-eight-page document these workers are provided. Sometimes, the workers also received notes, through their contract work agencies, from Google telling them the "correct" results for certain searches. For instance, at one point, the search phrase "best way to kill myself" was turning up how-to manuals, and the contract workers were sent a note saying that all searches related to suicide should return the National Suicide Prevention Lifeline as the top result.

This window into the work of the evaluators, brief though it is, helps us unpack Cerf's testimony. Google employees—presumably high-level ones—make far-reaching decisions about how the search algorithm should perform on various topics and in various situations. But rather than trying to directly implement these in the computer code for the search algorithm, they codify these decisions in the instruction manual that is sent to the evaluators. The evaluators then manually rate sites and search rankings according to this manual, but even with this army of ten thousand evaluators, there are far too many sites and searches to go through by hand—so as Cerf explained, these manual evaluations provide the training data for supervised learning algorithms whose job is essentially to extrapolate these evaluations so that hopefully all searches, not just the ones that have been manually evaluated, behave as the Google leadership intends.

While some of the notable updates to the Google search algorithm have been publicly announced[53] by the company (several were mentioned in this chapter), Google actually tweaks its algorithm extremely often. In fact, the same *Wall Street Journal* investigation just mentioned also found that Google modified its algorithm over thirty-two hundred times in 2018. And the number of algorithm adjustments has been increasing rapidly: in 2017, there were around twenty-four hundred, and back in 2010 there were only around five hundred. Google

[53] These announcements typically appear in Google blog posts, but a convenient list and description of the substantial ones has been collected by a third-party organization called the *Search Engine Journal*: `https://www.searchenginejournal.com/google-algorithm-history/`.

has developed an extensive process for approving all these algorithm adjustments that includes having evaluators experiment and report on the impact to search rankings. This gives Google a sense of how the adjustments will work in practice before turning them loose on Google's massive user base. For instance, if certain adjustments are intended to demote the rankings of fake news sites, the evaluators can see if that actually happens in the searches they try.

Let me return now to Vint Cerf. Shortly after the question that led to his description of Google's "two-step" process that I quoted above, the chair of the committee asked Cerf another important, and rather pointed, question: "Your algorithm took inaccurate information, that Muslims do not pay council tax, which went straight to the top of your search results and was echoed by your voice assistant. That is catastrophic; a thing like that can set off a riot. Obviously, 99 percent of what you do is not likely to do that. How sensitised are your algorithms to that type of error?" Once again, Cerf's frank answer was quite intriguing. He said that neural networks (which, as you recall, are the framework for deep learning) are "brittle," meaning sometimes tiny changes in input can lead to surprisingly bad outputs. Cerf elaborated further:

> Your reaction to this is, "WTF? How could that possibly happen?" The answer is that these systems do not recognise things in the same way we do. We abstract from images. We recognise cats as having little triangular ears, fur and a tail, and we are pretty sure that fire engines do not. But the mechanical system of recognition in machine-learning systems does not work in the same way our brains do. We know they can be brittle, and you just cited a very good example of that kind of brittleness. We are working to remove those problems or identify where they could occur, but it is still an area of significant research. To your primary question, are we conscious of the sensitivity and the potential failure modes? Yes. Do we know how to prevent all those failure modes? No, not yet.

In short, we trust Google's algorithms to provide society with the answers to all its questions—even though it sometimes fans the flames of hate and fake news and we don't entirely know how to stop it from doing so.

Recall the quote I included earlier from Google's white paper: "[Google's] ranking system does not identify the intent or factual accuracy of any given piece of content." The *Wall Street Journal* investigation discussed in this section noted that Facebook has taken a more aggressive approach to removing misinformation and said Google publicly attributes this difference in approach to the fact that Facebook actually hosts content whereas Google merely indexes it. But in private a Google search executive told the *Wall Street Journal* that the problem of defining misinformation is incredibly hard and Google "didn't want to go down the path of figuring it out."

Summary

Fake news and harmful misinformation appearing at or near the top of Google search results became a widely discussed topic after Trump unexpectedly won the 2016 election. Many people started to blame Google and the other tech giants for their role in the election and in eroding the very notion of truth. Google responded by making a series of adjustments to its ranking algorithm—often with the assistance of an army of low-paid contract workers—over the ensuing years aimed at bringing trustworthy links to the top of searches and pushing less reliable ones lower down. In this chapter, I presented a variety of examples where this played out, gathered what technical details I could about the closely guarded search algorithm, and looked into the public statements and general strategies Google has employed in this effort to elevate quality journalism. I also discussed instances of misinformation, deception, hateful stereotyping, and blatant racism that surfaced on other corners of Google such as maps, image search, and autocomplete.

In the next chapter, I tackle another aspect of Google: how its advertising platform provides the revenue stream for a huge fraction of the fake news industry. Facebook is also brought into the fray, though its advertising platform fans the flames of fake news in a rather different way, as you will soon see.

CHAPTER 7

Avarice of Advertising

How Algorithmic Ad Distribution Funds Fake News and Reinforces Racism

One of the incentives for a good portion of fake news is money.

—Fil Menczer, Professor of Informatics, Indiana University

When we think of Google supporting the fake news industry, the first thing that comes to mind is how, as described in the previous chapter, it serves up an audience with its various search products. However, there is an entirely separate way—less obvious but extremely influential—that Google supports the fake news industry: financially through ad revenue. The first half of this chapter focuses on the mechanics and scale of Google's algorithmic ad distribution system, the extent to which it funds fake news organizations, and the reluctant steps Google has taken over the years to curtail this dangerous flow of funds.

N. Giansiracusa, *How Algorithms Create and Prevent Fake News*,
https://doi.org/10.1007/978-1-4842-7155-1_7