OMICS A Journal of Integrative Biology
Volume 23, Number 0, 2019

DOI: 10.1089/omi.2019.0078

**Research Article**

# Sentiment Analysis of the News Media on Artificial Intelligence Does Not Support Claims of Negative Bias Against Artificial Intelligence

Colin Garvey and Chandler Maskal



## Abstract

Artificial intelligence (AI) is a hot topic in digital health, as automated systems are being adopted throughout the health care system. Because they are still flexible, emerging technologies can be shaped significantly by media representations as well as public engagement with science. In this context, we examine the belief that negative news media coverage of AI—and specifically, the alleged use of imagery from the movie *Terminator*—is to blame for public concerns about AI. This belief is identified as a potential barrier to meaningful engagement of AI scientists and technology developers with journalists and the broader public. We name this climate of risk perception the ''Terminator Syndrome''—not because of its origins in the movie of the same name *per se*, but because such unchecked beliefs can terminate broad public engagement on AI before they even begin. Using both quantitative and qualitative approaches, this study examined the hypothesis that the news media coverage of AI is negative. We conducted a sentiment analysis of news data spanning over six decades, from 1956 to 2018, using the Google Cloud Natural Language API Sentiment Analysis tool. Contrary to the alleged negative sentiment in news media coverage of AI, we found that the available evidence does not support this claim. We conclude with an innovation policy-relevant discussion on the current state of AI risk perceptions, and what critical social sciences offer for responsible AI innovation in digital health, life sciences, and society.

**Keywords:** artificial intelligence (AI), digital health, risk governance, technology policy, AI and risk, public engagement

## Introduction

Artificial intelligence (AI) has become a hot topic in a wide range of science and technology fields. AI is expected to broadly impact the digital transformation of health care, not to mention automation in allied fields such as biomedical research, data science, and integrative biology. Impacts on digital health powered by AI and automation are already manifesting at multiple levels in medical research ranging from study design, data collection, and real-time data analysis to societal applications of Big Data (Bohannon, 2015a; Char et al., 2018; Forsting, 2017; Garvey, 2018d; Just et al., 2017; Kaiser, 2018; Meyer et al., 2018).

The current enthusiasms for AI in life sciences and health care can be better understood in the following brief historical context, however. AI has evolved under three broad paradigms or ''peaks'' spanning over six decades: (1) good-old-fashioned AI (GOFAI) (1950–60s), (2) ''expert systems'' (late 1970–80s), and (3) ''machine learning (ML)'' (2010–present) (Garvey, 2018c; Matsuo, 2015; Ziad and Emanuel, 2016). In between these peaks, there were periods of ''troughs'' characterized by low optimism and lower funding, also known as ''AI Winters'' (Fast and Horvitz, 2017; Hendler, 2008).

The GOFAI paradigm was driven by symbolic logic to make ''machines who think'' (Haugeland, 1985; McCorduck, 1979, 2004). Later, expert systems narrowed the focus from general intelligence to human expertise in specific domains, such as chemistry and medicine (Buchanan and Shortliffe, 1985; Feigenbaum et al., 1988). The current ML paradigm extrapolates patterns directly from Big Data, usually through a training period involving millions of trial-and-error loops (Jordan and Mitchell, 2015).

This requires considerable computational power and memory, which is why the recent successes of ML algorithms in image (e.g., in radiology) recognition and classification (Esteva et al., 2017), speech recognition and language translation (Hirschberg and Manning, 2015), and games like *Go* and

Science and Technology Studies Department, Rensselaer Polytechnic Institute, Troy, New York.

poker (Brown and Sandholm, 2018; Moravčík et al., 2017; Silver et al., 2016, 2017, 2018) owe as much to recent hardware innovations as clever programming (Stone et al., 2016).

Human values and politics (i.e., the constitution and contestation of power) manifest themselves in various forms and contexts in science and technology. Scholars on the politics of knowledge production have for decades described the ways in which human values and power decisively shape the entire trajectory of science and emerging technologies such as AI—from new idea conception to implementation science (Agalianos, 2006; Bijker and Law, 1992; Bradshaw et al., 2013; Campbell, 2005; Eubanks, 2017; Forsythe, 2001; Harding, 2006; Hoffman, 2017; Sabanović, 2014; Winner, 1980).

Yet public engagement in science and technology, if designed with guidance from critical social science, can help surface the human values and power acting on emerging technologies, thereby contributing to transparency and accountability in science, as well as technological democracy (Barber, 1998; Carroll, 1971; Dotson, 2017; Eglash and Garvey, 2014; Feenberg, 2017; Sclove, 1995; Stilgoe et al., 2014; Woodhouse and Patton, 2004; Wynne, 2006).

While systems science applications forge ahead in medicine (Dzobo et al., 2018; Konstorum et al., 2018), the suite of AI technologies continues to be championed and contested in parallel. In this context, several questions are necessary to understand the AI future(s).

- Will AI scientists build ''AI for Social Good,'' as many industry leaders (Dean and Fuller, 2018; Horvitz, 2017; Nadella, 2016) and academics claim (Floridi et al., 2018; Taddeo and Floridi, 2018)?
- A wide range of companies, think tanks, interest groups, and government agencies have taken up AI for Social Good as an organizing theme. Why then is the public allegedly concerned about and distrustful of AI? For example, one study found that the majority of Americans is ''worried'' about AI's effect on the job market (72%), the impact of hiring algorithms (67%), and the development of driverless cars (54%) (Pew Research Center, 2017). The tech industry now recognizes that a lack of public trust may pose one of the greatest barriers to AI adoption in areas from driverless cars to health care (AAA, 2018; Davenport, 2018; Dujmovic, 2017; Wakefield, 2018).
- Finally, how and under which epistemologies will the concerns of the public about AI be addressed? Will the AI community's attempts to ameliorate public concerns and rectify the ''trust crisis'' (Agency Staff, 2018) allow for critical approaches to technology assessment, or remain constrained to narrowly framed market efficiency arguments?

For answers, broad public engagement with AI is essential. Such engagement exercises should be mindful of politics of knowledge production, and epistemological (Özdemir and Springer, 2018) and other forms of diversity (Callahan et al., 2016) in the engagement practices themselves, not only because the field of AI itself is in a gender (Simonite, 2018) and racial ''diversity crisis'' (Snow, 2018) but also because of the cognitive and democratic benefits such diversity brings to collective decision making (Hong and Page, 2012; Landemore, 2013; Lindblom and Woodhouse, 1993). The politics of and the power embedded in public engagement (who is represented, included, or excluded in engagement, and why?) should also be borne in mind to prevent such exercises from being transformed into hollow pageantry to improve public relationships or a ''tick the box'' exercise devoid of meaningful exercises of technological democracy (Stilgoe et al., 2014; Wynne, 1992).

Allowing for dissent and disagreement is valuable to achieve critical and reflexive perspectives at the science and society interface (Garvey and Chard, 2018; Lindblom, 1990). Moreover, forced consensus on emerging technologies does not bring about sustainable or responsible innovation (Sarewitz, 2011, 2015).

Moving forward on AI and society research, there are barriers to public engagement (Garvey, 2018a). Many in the AI technology community hold an antagonistic stance toward the media. This is not unique to AI; the contentious relationship between science and the media has been studied for decades (Nelkin, 1987, 1989, 1998). As Nelkin points out, scientists often ''interpret critical reports about science or technology as evidence of an anti-science or antiestablishment bias'' and respond by attempting to control journalistic access (Nelkin, 1987: 155). For example, the Conference on Neural Information Processing Systems (NeurIPS, formerly NIPS), currently the largest AI conference with over 8000 attendees (Shoham et al., 2018), banned journalists from the tutorials and workshops in 2017, and from the workshops in 2018 (Shead, 2018).

The belief that negative media coverage of AI—and specifically, news media's alleged use of imagery from the movie *Terminator*—is to blame for public concerns about AI has been widespread in the AI community for years (Shead, 2018). This belief poses a nontrivial ''psychocultural barrier'' (Dotson, 2015) to broader public engagement on AI (Garvey, 2018a), thereby reducing democratic societies' capacities for responsibly governing the manifold risks posed by AI technologies (Garvey, 2018b).

We name this climate of risk perception the ''Terminator Syndrome''—not because of its origins in the movie of the same name *per se*, but because such beliefs can terminate broad public engagement with AI before it even begins.

Using both quantitative and qualitative approaches, this study examined the hypothesis that news media coverage of AI is negative. We conducted a sentiment analysis of news data spanning over six decades, from 1956 to 2018, using the Google Cloud Natural Language API Sentiment Analysis tool. We found that the available evidence does not support the claim of a bias toward negative sentiment in the news media coverage of AI. We conclude with an innovation policy-relevant discussion on the current state of AI risk perceptions, and what critical policy studies offer for ''AI and society'' scholarship and responsible innovation for emerging applications of AI in health care, life sciences, and society.

## Materials and Methods

The term ''artificial intelligence'' dates to 1956, when it was first used in the title of a conference at Dartmouth College (McCarthy et al., 2006). Because this date is generally accepted as the origin point for the AI field (Crevier, 1993), we decided to analyze the period from 1956 to 2018.

Traditional approaches to sentiment analysis of the news media typically involve manually classifying articles from a selected corpus, delimited either by temporal range or article availability (Martin, 1993). However, our intention to analyze news article sentiments over AI's entire history, combined with the recent explosion of news articles about AI, made manual approaches to sentiment analysis low throughput and thus inappropriate for this study. We therefore decided to automate the task by employing a high-throughput methodology utilizing natural language processing (NLP) software to perform sentiment analysis on a larger (what is essentially Big Data) corpus of AI news articles than could be feasibly coded by hand. To the extent that heuristic search and NLP both emerged from AI (Stone et al., 2016), this study used basic AI tools to study the news media coverage of AI itself. This study was approved by the authors' institutional review board (IRB).

We began by selecting news sources to establish a corpus of articles. We accessed online news sources through their application programming interfaces (APIs), a common method for aggregating news from multiple sources into streaming feeds at a single location (e.g., Google News). Our review of available news API services ultimately informed the subsequent study design. To compensate for potential bias in coverage, we had planned to populate our corpus of articles from sources spanning the full ideological spectrum, expanding out from allegedly neutral news organizations to include media outfits on both ends (e.g., conservative and progressive) of the political spectrum. However, we found that most news outlets had either commercialized or otherwise restricted access to their APIs.

We therefore adopted a research design utilizing depth and breadth searches on separate sources of AI news media. *The New York Times* (NYT) API was chosen for *depth* search, as it offers news dating from 1891 to the present, as well as a limited selection of articles from Associated Press, The International Herald Tribune, Reuters, CNBC, International NYT, and Internet Video Archive (NYT, 2019).

The news aggregator News API was chosen for *breadth* search, as it offers articles from over 30,000 different sources, but only includes articles published within the last 3 months of the search (for unpaid users) (News API, 2019).

With historical depth and contemporary breadth datasets, we believe the source material presented in this study to be an adequate representation of a broad mediasphere. The NYT has long been regarded as the flagship news outlet in the USA, a publication of record known for relatively minimal partisan bias and broad coverage of events. In contrast, News API aggregates a diversity of news sources, from mainstream news outlets to technology blogs, thus ensuring that this study captures media reflective of broader post-print news and Internet cultures.

Using the search queries ''A.I.'' and ''artificial intelligence,'' we retrieved 8470 articles from NYT over the period from 1956 to March 27th, 2018, and 3906 articles from News API over the period from December 14th, 2017, to March 14th, 2018, for a total of 12,376 articles. Figure 1 shows an overview of the data collection process.

The news search function examined the body text, headline, and byline of each article for the query text. When a match is found, the API only returns the article's headline, a ''snippet'' of the body text (usually the first line of the article), and metadata (such as date, section, and source). Although the articles' body text was unavailable, social scientific research on news coverage of science has emphasized how headlines not only concisely convey the emotional and informational content of entire articles but also play important roles in forming public opinion (Nelkin, 1987). Therefore, this study utilized a simplifying assumption that

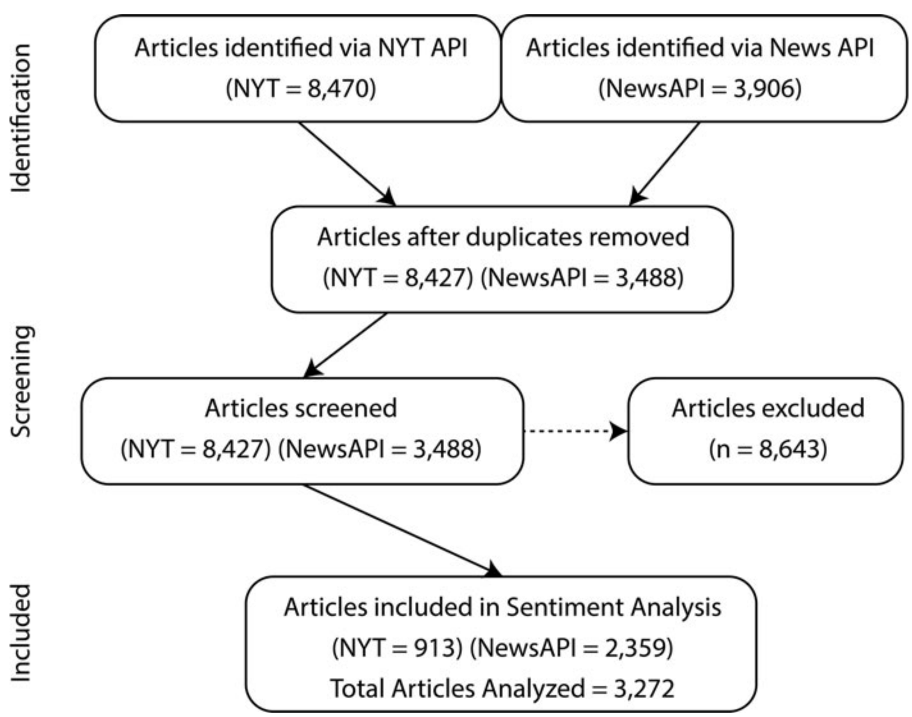


**FIG. 1.** Flowchart of the data collection and preparation process.

article sentiment can be approximated using only headline and snippet data.

After removal of duplicate articles, we had 11,915 articles in two datasets: NYT ($n = 8427$) and NewsAPI ($n = 3488$). We made the assumption that if an article's headline and snippet did not include the query terms ''artificial intelligence'' or ''A.I.,'' and it did not appear relevant to AI as suggested by its tacit sociotechnical context (see examples below), then it should be considered a false positive and removed from the dataset.

However, ''AI'' is itself a broad concept and a fluid field (Stone et al., 2016). This made rules for determining the relevance of each article challenging to define. Ultimately, because this study was interested in media representations and sentiments of AI, we decided that articles containing cognate terms such as ''driverless car,'' ''robot,'' ''thinking machine,'' and so on should be considered relevant.

For example, a headline such as ''Ava of 'Ex Machina' Is Just Sci-Fi (for Now)'' (NYT, 5/21/2015) does not explicitly mention ''artificial intelligence'' or ''A.I.,'' but it does refer to a robot from a popular sci-fi movie *about* AI. Moreover, the article's snippet, ''A question most techies don't seem to want to answer: who is making sure that all of this innovation does not go drastically wrong?'' is directly relevant to the themes of fear, sensationalism, and danger surrounding AI that this study intended to examine. Therefore, it was decided that this article and others like it should be included in the final dataset. Table 1 shows the final selection of the relevant terms.

The two datasets were then cleaned of false positives, as noted above, and reviewed by the authors. We started the analyses with the oldest article and moved forward toward the newest, read the headline first, then the snippet, and looked for the query terms ''A.I.,'' ''AI,'' or ''artificial intelligence.'' If these terms were not found in either the headline or snippet, we carried out a qualitative analysis to verify and decide if the article was related to AI and attendant sentiments, using the terms in Table 1 as guidelines.

The search queries used in this study returned some unanticipated results that attest to the importance of utilizing human intelligence (HI) (Özdemir, 2019) with reference to tacit social context (Dreyfus, 1992; Taube, 1961) in media analysis. For example, the NYT dataset contained many articles from the 1980s on ''artificial insemination,'' which described people looking for sperm donors with high ''intelligence.'' These articles were not included because they are obviously not related to AI. In addition, the search queries returned some articles headlined by ''Artificial Intelligence,'' which actually described false or ''artificial'' military intelligence; these were not included.

Table 1. Terms Used to Select Relevant Articles for This Study

| *Terms included* | *Terms not included* |
|---|---|
| Alternative terms for AI: thinking machines, smart machines, intelligent machines, etc. | Drone |
| AI techniques: machine learning, deep learning, neural networks, expert systems, etc. | Supercomputer |
| Specific AI systems: IBM Watson, DeepMind AlphaGo, Apple's Siri, Amazon Alexa, Microsoft Cortana, etc. | Artificial insemination |
| Applications of AI: driverless cars, self-driving cars, autonomous cars, etc. | ''Artificial'' (i.e., false) military intelligence |
| Cognate terms: robot (also ''-bot,'' ''chatbot,'' ''sexbot,'' etc.) | |

AI, artificial intelligence.

The data cleaning process resulted in two final datasets (NYT = 913, NewsAPI = 2359). The sentiment of all 3272 articles (headline plus snippet) was then analyzed using Google Cloud Natural Language Sentiment Analysis tool.

Accordingly, for each article, Google's Sentiment Analysis tool returns two numerical values that can be used ''to determine the overall attitude (positive or negative) expressed within the text'' (Google, 2019). The first score dimension is a ''sentiment score'' between $-1.0$ (negative) and $+1.0$ (positive), which ''corresponds to the overall emotional leaning of the text.'' We note, however, that Google offers no explanation of how these associations are formed. Sentiment scores in the negative range ($-$) indicate the presence of text associated with overall negative emotions, and higher scores in the positive range (+) with positive emotions. A positive score in the first dimension does not necessarily mean, however, that the article contained no negative emotions in certain sections, but rather that the overall sum of the emotions was positive.

The second score dimension is a ''magnitude score'' between 0.0 and infinity that ''indicates the overall strength of the emotion noted (both positive and negative) with the given text'' (Google, 2019). The magnitude score is calculated as a sum of all sentiment, both positive and negative; in other words, it is the absolute value of sentiment in the text. Therefore, if not 0.0, the magnitude score is always positive. In addition, it increases with each occurrence of sentiment in a text, so longer texts will likely have higher magnitudes, with no upper limit.

One limitation of Google's Sentiment Analysis tool is that scores of zero in the sentiment dimension can indicate either neutral sentiment throughout the text or the presence of both positive and negative sentiments that cancel each other out, resulting in an overall sentiment of zero. Here, the second magnitude score dimension can be used to distinguish between the two possibilities. Articles with both sentiment and magnitude scores near zero suggest truly neutral text, whereas near-zero sentiment with high magnitude scores suggest strong contrasting sentiments within the text.

For example, the article ''Happy Birthday, HAL; What Went Wrong?'' (NYT, 1/12/92) returned a modest sentiment score of +0.1 in the positive sentiment range and a high magnitude of 2.8. This scoring presumably reflects the emotional contrast between ''Happy'' and ''Wrong,'' whose positive and negative sentiments canceled out in the first dimension, while the strength of emotion in both words (happy and wrong) contributed to the total value of the second magnitude score dimension.

The sentiment and magnitude score data were then visualized using Tableau (Tableau Software, 2019), a data visualization software package available for free under educational license.

### Results and Discussion

News media is a pillar of technological democracy and an important conduit for public engagement in science. Because news media influences both publics' and practitioners' perception, understanding, and beliefs about emerging technologies such as AI and their medical applications, the emotional sentiment of news articles can shape the development and trajectory of digital health innovations.

We found that across both depth and breadth datasets, AI news coverage demonstrates robustly positive sentiment. Figure 2 shows the count of articles with positive sentiment (>0), negative sentiment (<0), and neutral sentiment (0). Articles with positive sentiment outnumber negative and neutral articles for both NYT and NewsAPI datasets, with 502 positive articles out of 913 total NYT articles (54.98%), and 1423 positive articles out of 2359 total NewsAPI articles (60.32%).

By contrast, only 223 NYT articles (24.4%) and 498 NewsAPI articles (21.1%) were negative. That is, articles with negative sentiment on AI accounted for less than one in four articles in our broad sampling of the mediasphere, whereas a sizable majority display positive sentiment.

Articles with neutral sentiment accounted for the smallest percentage of the datasets, with only 188 NYT articles (20.6%) and 438 NewsAPI articles (18.6%). Of these, a majority received positive magnitude scores, indicating the presence of contrasting positive and negative sentiments within the text. Only 69 NYT articles (7.6%) and 129 NewsAPI articles (5.5%) were truly neutral, meaning they had sentiment and magnitude scores of zero.

Interestingly, the average sentiment score of positive articles was nearly identical across NYT (0.3876) and NewsAPI (0.3857) datasets. Moreover, the average negative score was also quite close in both NYT (–0.3108) and NewsAPI (–0.2926) datasets.

That the vast majority of articles display some sentiment, whether positive or negative, suggests an appreciable degree of polarization in the sentiment space of AI news coverage. Moreover, the similarity of average positive and negative sentiment scores across the depth and breadth news corpuses further suggests that this polarization may be temporally robust across a broad mediasphere. One implication for public engagement is that emerging technology ecosystems with polarized sentiments can become entrenched (Collingridge, 1980), with the consequence that actors within them may become resistant to new perspectives on the emerging technology.

Figure 3 shows a histogram of the NYT article count by year, and notably, how news coverage of AI has exploded in the last decade, with considerable increases starting in 2015. The low count for 2018 is an artifact reflecting that this study only included articles up to March 27, 2018; presumably, many more were printed over the rest of the year. Interestingly, the three historical paradigms and periods of AI discussed in the introduction can be discerned from our data, with a handful of articles during the 1960s (GOFAI), a peak/trough pattern from the late 1970s to early 1990s (expert systems), and a clear boom of coverage in recent years (ML).

Figure 4 provides a historical time course of the sentiment of the NYT depth corpus, showing that yearly averages are

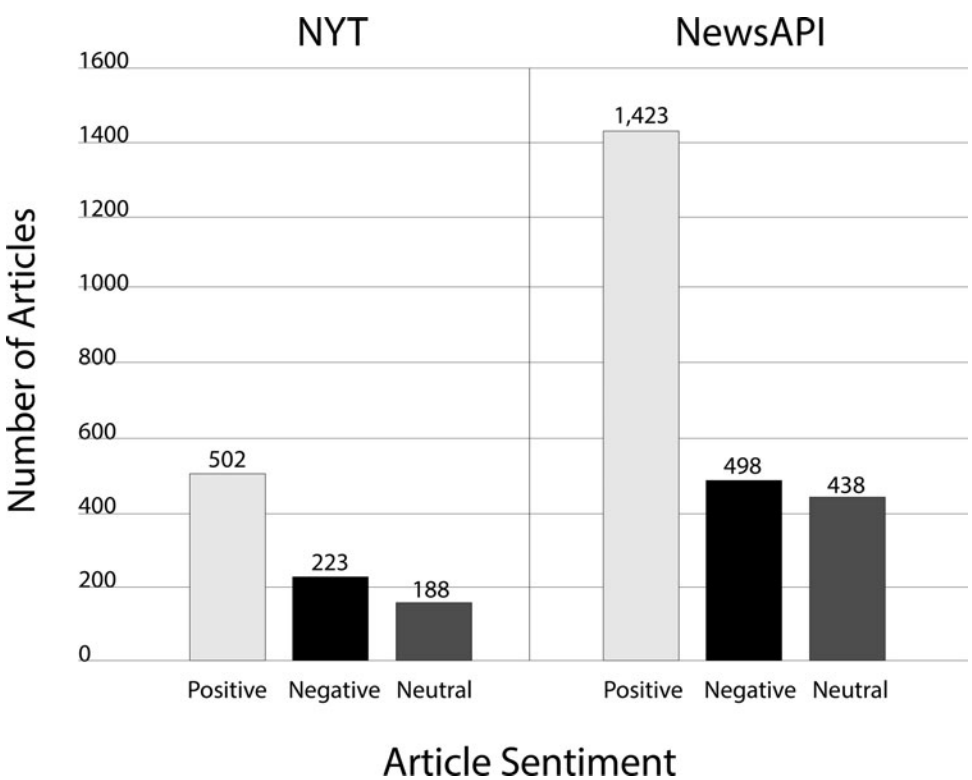


**FIG. 2.** Count of articles with positive sentiment (>0), negative sentiment (<0), and neutral sentiment (0).

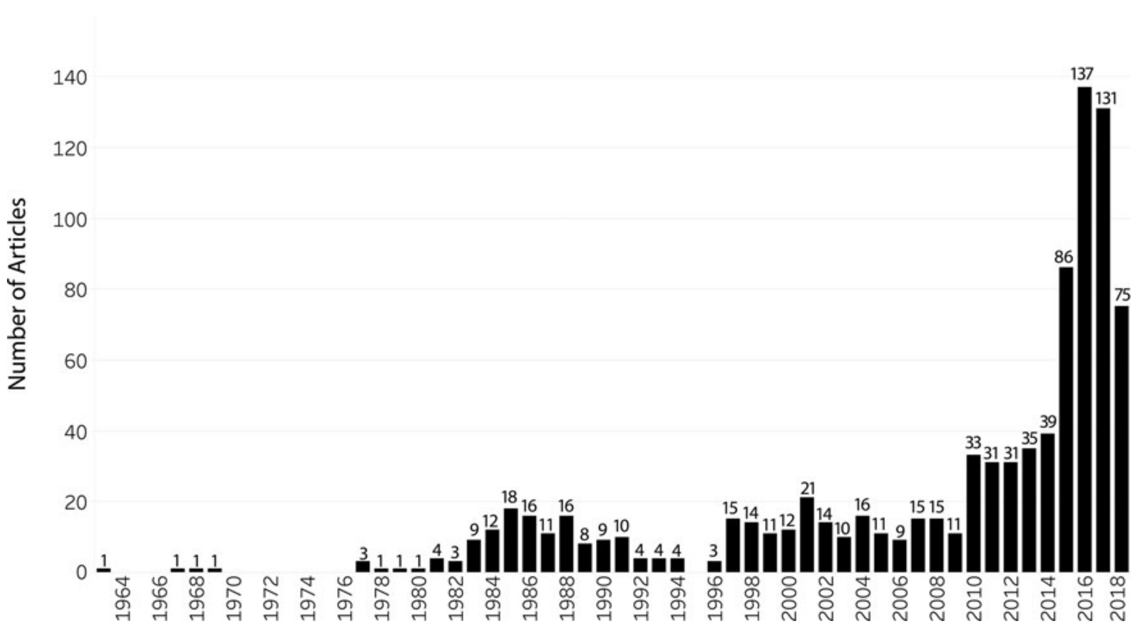


**FIG. 3.** Total count of NYT articles by year. NYT, The New York Times.

robustly positive across more than five decades, as are average sentiment magnitude scores.

Qualitative analysis using HI through discussions among the authors could not discern why the earliest article in the dataset (''Engineers Hailed for Space Work,'' NYT, 8/23/1963), which mentioned the ''Artificial Intelligentsia,'' received a negative sentiment score (–0.4). Anomalies such as this highlight one shortcoming of using black boxed AI technologies like Google's sentiment analysis tool (see our discussion of this problem below).

The extreme fluctuation of scores from 1965 to 1980 reflects, in part, the paucity of articles from that period. The next lowest average sentiment (–0.031) occurs in 1988, following the onset of ''AI Winter'' in 1987 (Fast and Horvitz, 2017), with multiple lead stories reporting on the field's failure to live up to its promises (Garvey, 2018c). For example:

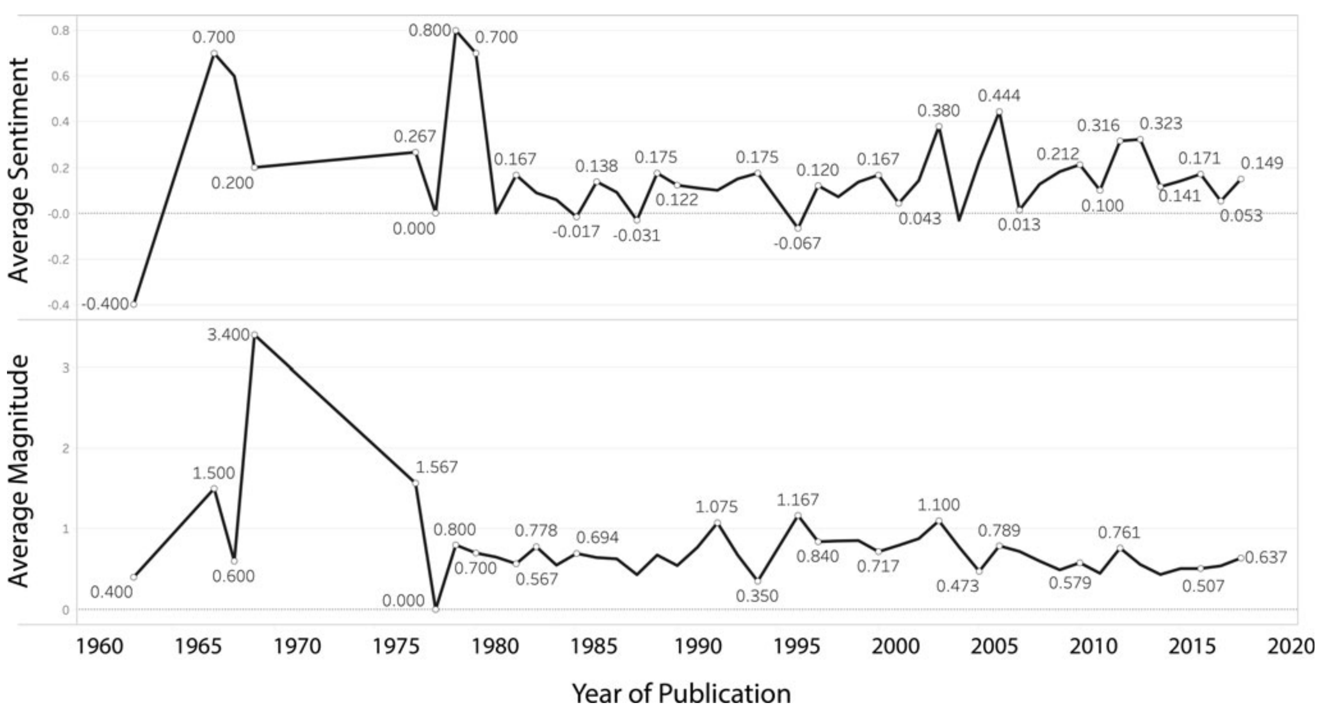


**FIG. 4.** Average sentiment and magnitude of NYT articles per year. Average sentiment is shown above and average magnitude below.

> Setbacks for Artificial Intelligence: A major retrenchment is occurring in the artificial intelligence industry, dashing the hopes of many companies that thought they would prosper by providing the technology to make computers ''think.'' (NYT, 3/4/1988)

The lowest average annual sentiment score in the entire depth corpus (–0.067) occurs in 1996, in the middle of the AI Winter of the 1990s, before picking back up in with IBM Deep Blue's 1997 win over Gary Kasparov (Campbell et al., 2002; Kasparov and Greengard, 2017). However, coverage in 1996 was scarce, with only three articles, and the low average score appears to be impacted by an idiosyncrasy of Google's tool, which ranks words like ''cockroaches'' as extremely negative (see discussion below).

Taken together, Figures 3 and 4 offer an important takeaway from this study: despite the vast increase in AI news coverage over the last decade, average sentiment and magnitude remain positive. Moreover, the most negative articles from this period reflect real-world events rather than sensational speculation or fearmongering. For example, comparatively low average sentiment score for 2017 (0.05) is due, in part, to the article, ''Google Self-Driving Car Unit Accuses Uber of Using Stolen Technology'' (NYT, 2/24/2017), which received a sentiment score of –0.9 and a magnitude of 0.9. The highly publicized legal battle that followed from Google's accusations revealed a disturbing level of recklessness in the driverless industry, which presumably contributed to public perceptions of AI (Duhigg, 2018; Lee, 2018; Somerville and Levine, 2017).

Figure 5 displays the highest (peak), average, and lowest (trough) sentiment scores, as well as average magnitude scores, per day, for the NewsAPI breadth corpus. There were more dramatic fluctuations in the trough sentiment scores than the peak scores in a given day; indeed, on 4 days (1/1, 1/13, 2/10, and 2/17), for example, there were no articles with negative sentiment, and on 2 days, the trough sentiment score was in the positive value range (1/13, 2/10). This suggests that media coverage of AI is more consistently positive with greater stability in the positive scores, and only periodically negative.

The average magnitude scores were also high on most days and well over 0.500, suggesting most articles contain sentimental valence (polarity), whether positive or negative. However, the top five magnitude peaks (12/25/17, 1/6, 1/14,



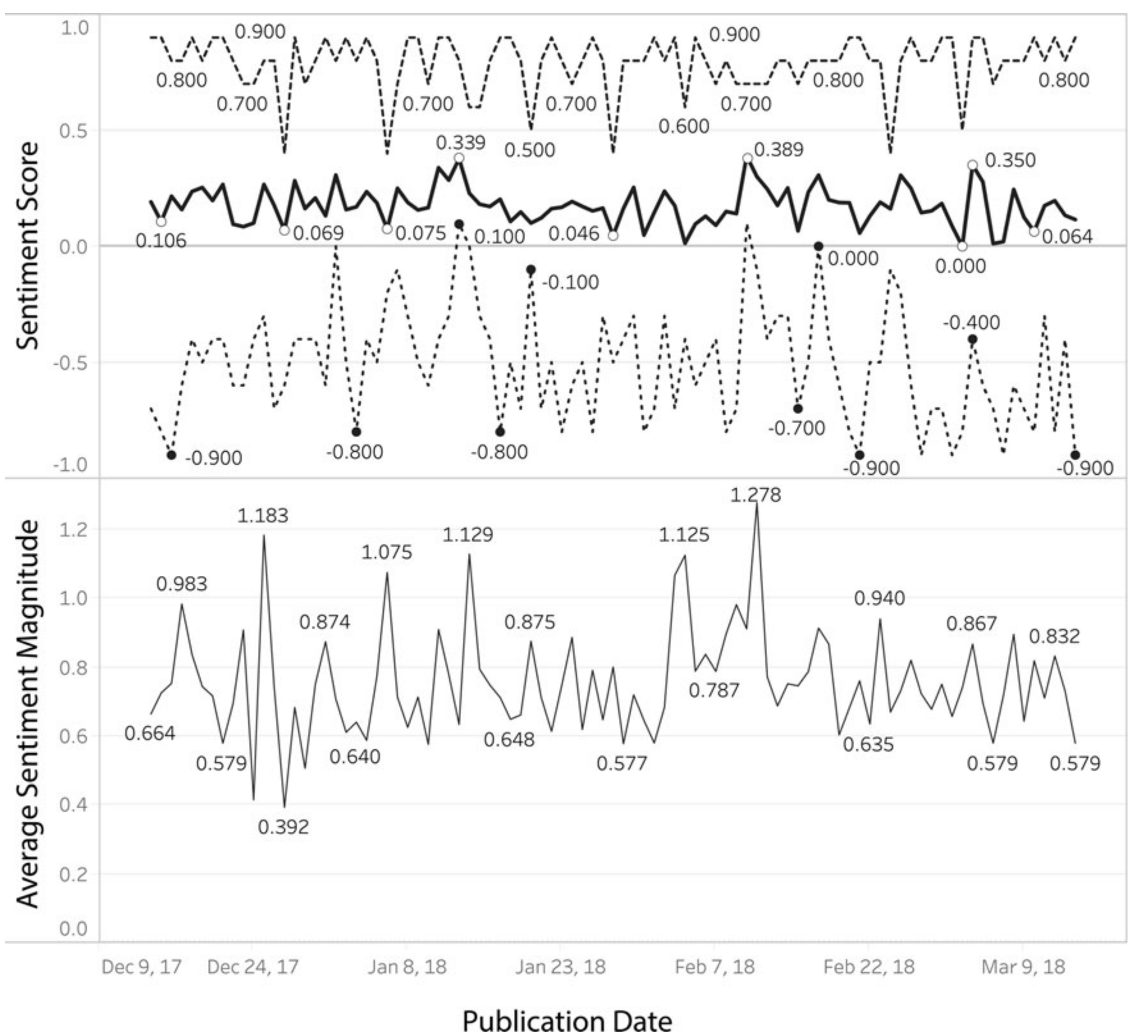


**FIG. 5.** Daily (high, average, and low) sentiment scores and average magnitude scores for NewsAPI articles. API, application programming interface.

2/4, and 2/11) occurred on days with a positive average sentiment, which implies that spikes in emotional magnitude might correlate with positive sentiment more strongly than with negative, at least within the period analyzed. Indeed, the average sentiment score over the 3-month time span never dropped below zero.

For the NewsAPI data, the lowest average daily sentiment (0.0) occurred on March 3rd, 2018. Of the 10 articles published that day, 5 were positive, 1 was neutral, and 4 were negative. The two most negative articles, from Yahoo.com (–0.8) and *Fortune* (–0.3), each carried the headline, ''Elon Musk Blasts Harvard's Steven Pinker Over Comments Dismissing the Threat of Artificial Intelligence'' (3/3/2018). Musk has been one of the foremost figures in raising the alarm about AI (Anderson, 2014; Breland, 2017; Devaney, 2015). However, our sampling of the news that day shows that even when a dystopian view of AI is reported on, its sentiment is counter-balanced by multiple positive articles in the mediasphere.

Figure 6 displays a treemap of the 659 different data sources that provided the 2359 articles for the NewsAPI breadth corpus. The size of each cell corresponds to the number of articles from each source, and the shading to the averaged sentiment scores of the articles from that source. For example, the largest cell in the upper left represents a source (Youbrandinc.com) that published 152 articles whose average sentiment is positive (0.140), while the smallest cell in the bottom right represents a source (Zohosites.com) that only published 1 neutral article (0.0). Overall, Figure 6 qualitatively depicts a positive relationship between the number of articles published and the positive sentiment. That is, the sources that provided more AI coverage offered, on average, more positive coverage.

For the most part, sources with greater coverage were also more recognized and reputable sites (e.g., Slashdot.org, CNBC, Forbes.com, Google News, *NYT*), whereas sources with fewer articles were mostly news blogs. The exceptions to this general qualitative trend are worth noting, however, as they suggest there appears to be no clear connection between political orientation and sentiment regarding AI.

The largest source with a negative average sentiment score is the contested extremist media outlet Breitbart.com (12 articles, average sentiment score = –0.017). Yet an equally contested extremist site, Infowars.com, was one of the larger news sources in our dataset with 38 articles, mostly comprised blog posts about recent market research that scored a positive average sentiment of 0.376. Identification with political extremism does not seem to indicate a clear negative or positive bias toward AI.

By contrast, Sciencemag.org, the official website of the journal *Science* and arguably a highly reputable source of

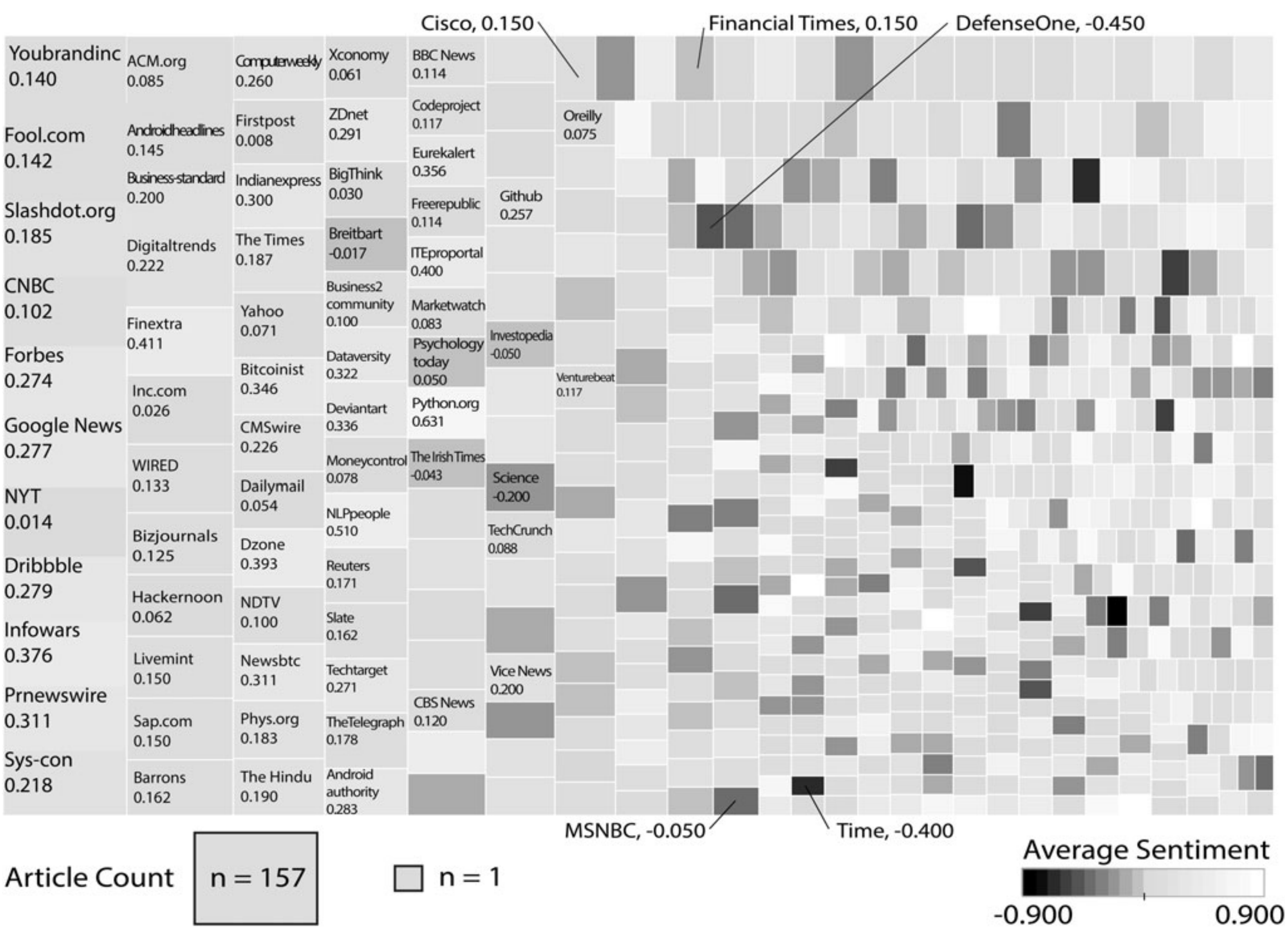


**FIG. 6.** Treemap of the 659 different data sources for the NewsAPI breadth corpus. Cell size corresponds to article count, and shading to average sentiment score.

scientific information, had five articles with a negative average sentiment of –0.2. Our qualitative analysis of the articles found no evidence of unfair negative bias. The sentiment score in the latter site was brought low by articles about real problems facing AI at the moment, with two on the ''reproducibility crisis'' emerging in AI (''Missing data hinder replication of artificial intelligence studies,'' 2/15/2018, –0.3 and ''Artificial intelligence faces reproducibility crisis,'' 2/15/2018, –0.2) and one on the authoritarian political applications of AI in China (''China's massive investment in artificial intelligence has an insidious downside,'' 2/8/2018, –0.8).

### *Contextualizing the findings*

Our most salient finding that AI news coverage, by and large, is robustly positive casts doubt on the belief (Guizzo and Ackerman, 2016; Scharre, 2017; Shead, 2018; Williams, 2015) that negative public perceptions of AI can be explained by negative media coverage of AI. There are, however, caveats.

First, we acknowledge that quantitative scoring of emotional ''sentiment'' is a reductive construction that fails to capture contextual nuance. One could argue the score says more about the programmers who created the tool than the text it processes, making this entire exercise meaningless. However, sentiment analysis tools, including Google's, are used widely online and in industry for market forecasting and other forms of analysis. For example, the AI Index 2017 and 2018 included a sentiment analysis of AI media coverage (Shoham et al., 2017, 2018). Shoham et al. (2017) found positive AI articles outnumbered negative during 2013–17, although most articles were neutral (2017: 25).

Shoham et al. (2018) extended those results, adding that ''AI articles have become less neutral and more positive, particularly since early 2016, when articles went from 12% positive in January 2016 to 30% positive in July 2016,'' where it has remained since. However, the AI Index results do not bear a strong burden of proof because the provenance of the data used is unclear. It was obtained from a business analytics firm (Trendkite.com), but beyond that, no information is provided regarding the number of articles analyzed, their sources, or the methods used. This study thus improved on the AI Index study by conducting a sentiment analysis of the AI news media coverage over a longer period, while providing transparency on data, sources, and methodology.

Second, the accuracy of Google's sentiment analysis tool can certainly be questioned. We acknowledge that other sentiment analysis tools might provide more nuanced scores. The Google tool itself is a black box that does not explain how sentiments are weighted, although our examination of the most negative articles revealed some of its idiosyncrasies. For example, consider the following article that received one of the lowest sentiment scores (–0.9):

> In Kingdom of Cockroaches, Leaders Are Made, Not Born—An international team of scientists has created artificial roaches to study ''collective intelligence.'' (NYT, 12/7/2004).

In this case, the low score is almost certainly an artifact of the Google's NLP tool, which most likely assigns very low values to the word ''cockroaches.'' It seems unlikely that this article contributed significantly to negative public perceptions of AI.

However, many of the lowest scoring articles were indeed negative, covering a range of controversial issues highly relevant to AI today: billion-dollar theft of corporate IP (NYT, 2/24/17); Facebook's attempt to predict which users are at risk of suicide (CNBC, 2/21/18); failed safety systems in driverless cars resulting in a pedestrian's death (NYT, 3/21/18); the poor performance of AI-powered hedge funds (Investopedia.com, 3/14/18); issues of consent raised by sex robots (NYT, 7/17/17); sexual misconduct by AI executives (The Hill, 12/16/17); the malicious use of AI (Reason.com, 2/21/18); the threat of authoritarian AI technologies in China (Metafilter.com, 12/15/17); and so on.

Although the Google NLP sentiment analysis tool is arguably a simple one, it is not immediately clear that more complex tools are necessarily appropriate to answering the questions posed by our study about the sentiment of media coverage of AI. For example, Fast and Horvitz (2017) conducted a far more costly and comprehensive analysis of AI coverage in the *NYT* over the period 1986–2016. Using crowdsourcing and ML classifiers, they attempted to measure ''optimism versus pessimism'' and ''engagement,'' as well as a number of hopes for and concerns about AI. However, the relevant results of this sophisticated study confirm our own: ''In general, AI has had consistently more optimistic than pessimistic coverage over time, roughly two to three times more over the 30-year period'' (Fast and Horvitz, 2017).

Moreover, in their sophisticated technical analysis, Fast and Horvitz collapse the social dimension of their data—paragraphs of text from the *NYT*—to stand in for public belief itself. For example, they interpret an increase in one classifier, ''ethical concerns about AI,'' to ''suggest an increase in public belief that we may soon be capable of building dangerous AI systems'' (Fast and Horvitz, 2017: 967). This study, by contrast, acknowledges the possibility of difference between media representations and alleged public beliefs. Indeed, that appears to be the case with AI: despite robustly positive media coverage, concerns about the technology are widespread among the public.

Third, it is certainly possible that studies on larger datasets could potentially contest our findings. Future studies should analyze sentiment from many more news sources, perhaps controlling for ideology by surveying across the political spectrum. Ethnographic research is needed to better explore the process by which ordinary people make sense of promissory technologies like AI, and the role that news media plays in that understanding. However we leave that to future studies on the subject matter.

Having addressed those points, we now ask the following: If most AI news is positive, why then is a majority of the public allegedly worried about and distrustful of AI? We consider three hypotheses.

We agree it is possible that despite the mostly positive news coverage, a small number of highly sensational, ''fear-mongering'' articles could have a disproportionate effect on public perceptions of AI. Cognitive or other biases might cause extremely negative stories (e.g., Strange, 2015) to outweigh the preponderance of less remarkable positive stories in the publics' judgment. However, plausible, the burden of proof lies with those who would advance such a hypothesis.

Moreover, the principle of symmetry (Bloor, 1991) urges caution in invoking bias to explain public perceptions without considering its role in experts' own perceptions of AI news. Biases are an inescapable fact of human cognition (Gigerenzer and Goldstein, 1996; Kahneman, 2011); hence, similar explanations can be applied to both experts and lay people. That is, by the same logic, one could argue experts' biases inure them to the steady stream of positive AI news, while making them exceptionally sensitive to the occasional negative article, which would become overrepresented in their mental model of ''AI news'' and create the perception of a biased media.

A more parsimonious explanation for negative public perceptions of AI is that the technology actually does pose considerable risks to many people, and that the public is rightly concerned (Didier et al., 2015). Consider how after early denials (Williams, 2015), some major tech leaders are now conceding that AI does pose risks, such as Google CEO Sundar Pichai, who recently acknowledged that fears about AI are ''very legitimate'' (Romm et al., 2018).

Leading AI experts are warning about the military risks (Bohannon, 2015b; FLI, 2015). Senior politicians and government officials argue that AI threatens political democracy (Graham, 2017; Nemitz, 2018). Reckless corporate behavior has already led to the first pedestrian death by driverless car (Lee, 2018; Wakabayashi, 2018). ML has been shown to systematically reproduce biases in its training data (Caliskan et al., 2017), and such algorithmic biases have entrenched discriminatory practices in range of social settings, from criminal sentencing (Angwin et al., 2016; Matacic, 2018) to facial recognition (Buolamwini and Gebru, 2018). Investigative journalists have shown how AI allows advertisers to discriminate by race, gender, and other categories (Angwin and Parris Jr., 2016; Angwin et al., 2017).

Finally, business people, technical experts, financial institutions, think tanks, governments, and much of the public agree that AI puts jobs at risk (Acemoglu and Restrepo, 2017; Arntz et al., 2016; Brynjolfsson and Mitchell, 2017; Dutton et al., 2018; Frey and Osborne, 2013; Ma et al., 2015; Pew Research Center, 2014; World Economic Forum, 2018).

What about machine takeovers and killer robots? It does not appear to be the public's primary concern. One survey of 2000 American adults found that their ''top fear or concern about possible A.I. threats or risks'' was ''A.I. taking jobs'' (30%), rather than ''A.I. turning against us'' (8.8%) or ''AI taking control'' (6.8%) (SYZYGY, 2017). Indeed, even James Cameron, director of the *Terminator* movies, suggests shifting attention to more quotidian instantiations of AI technology and their social impacts:

> People ask me: ''Will the machines ever win against humanity?'' I say: ''Look around in any airport or restaurant and see how many people are on their phones. The machines have already won.'' (Belloni and Kit, 2017)

As we look around, it is worth asking what ''digital health'' means—and ought to mean—in this burgeoning Age of AI, where a certain class of machines has ''already won.'' What role should AI play in childhood development as screens become fixtures in peoples' lives at ever earlier ages, especially as it becomes clear that screen time correlates with adverse effects on young people (Hunt et al., 2018; Turel, 2019; Twenge et al., 2018, 2019)? How ought guidelines for digital entertainment be set, when even a site as seemingly innocuous as YouTube is now known to use AI to maximize screen time by tantalizing viewers with increasingly polarizing and extreme videos (Cook, 2018; Lewis, 2018; Nicas, 2018; Tufekci, 2018)? Should private initiatives to put screens in public schools (Kardaras, 2016; Toyama, 2015) not be reconsidered in light of the fact that Silicon Valley elites are increasingly restricting their own children's use of handheld digital electronics (Bowles, 2018)?

At a minimum, we suggest that ''digital health'' should be construed broadly enough to consider these and related questions about the role of AI technologies in human life. Indeed, it seems to us that digital health cannot be understood outside the context of AI technologies, with their potential benefits and manifold risks. This context suggests an orientation for the digital health research agenda. If AI poses significant risks to a nontrivial fraction of humanity (Garvey, 2018b), then we suggest that future research in digital health could make the greatest impact by adopting a perspective of ''thoughtful partisanship'' (Woodhouse et al., 2002) and explicitly focusing on the needs of those groups at greatest risk.

Because those at-risk populations are often also those least able to influence political decision making and technological R&D (Lindblom and Woodhouse, 1993), research oriented around their concerns will help to counteract the well-documented tendency for new technologies, especially in medicine, to disproportionately benefit the wealthy and amplify social inequality (Sarewitz and Woodhouse, 2003; Woodhouse and Sarewitz, 2007). We recommend science and technology policies for promoting ''digital health'' to be reoriented accordingly.

For digital health to be more than the next bump in the hype cycle, a systems perspective that includes critical social science and public engagement as inputs will be crucial. AI is increasingly a part of ordinary life, and the public is increasingly familiar with it. Explaining away public concerns about AI as mere fear of the *Terminator*, stoked by an unscrupulous or sensationalist media, poses a barrier to greater public engagement with AI by reproducing the ''deficit model'' of the public's understanding of science (Bauer et al., 2007; Fujigaki, 2009; Ziman, 1991), which too often assumes ''lay publics are only capable of taking sentimental, emotional, and intellectually vacuous positions'' (Wynne, 2001).

Humane futures in digital health and the responsible advancement of AI technologies require that the concerns of an informed public be heard, addressed, and incorporated into research, design, deployment, and operation. For that to happen, open, transparent, and deliberative interaction between AI innovators, news media, government officials, civic leaders, and diverse publics is essential.

## Conclusions

Using the sentiment scores calculated by the Google Cloud NLP tool, we found that the majority of AI news coverage from the sources we examined is positive. If this tool is reliable and accurate, and if our news sources are representative, then our finding that a majority of AI news coverage is positive refutes the hypothesis that most media coverage of AI is negative. This casts doubt on the validity of the belief, which we have called the ''Terminator Syndrome,'' that negative media coverage is largely to blame for negative public perceptions of AI. Although our results alone are unlikely to terminate the Terminator Syndrome, we hope they contribute to facilitating greater public engagement with AI.

## Acknowledgments

No funding was received in support of this research article. Views expressed are authors' personal opinions only and do not necessarily reflect those of the affiliated institutions.

## Author Disclosure Statement

The authors declare that no competing financial interests exist.

## References

AAA. (2018). AAA: American trust in autonomous vehicles slips. https://newsroom.aaa.com/2018/05/aaa-american-trust-autonomous-vehicles-slips/.

Acemoglu D, and Restrepo P. (2017). Robots and jobs: Evidence from US labor markets. Working paper no. 23285. DOI: 10.3386/w23285.

Agalianos A. (2006). The social shaping of logo. Soc Stud Sci 36, 241–267.

Agency Staff. (2018). IBM joins Apple in blasting Silicon Valley for "trust crisis." TechCentral. https://techcentral.co.za/ibm-joins-apple-in-blasting-silicon-valley-for-trust-crisis/85495/.

Anderson L. (2014). Elon Musk's artificial intelligence fear: Machines could wipe us out. Vanity Fair. www.vanityfair.com/news/tech/2014/10/elon-musk-artificial-intelligence-fear

Angwin J, Larson J, Mattu S, et al. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Angwin J, and Parris T. (2016). Facebook lets advertisers exclude users by race. ProPublica. https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

Angwin J, Tobin A, and Varner M. (2017). Facebook (still) letting housing advertisers exclude users by race. ProPublica. https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin

Arntz M, Gregory T, and Zierahn U. (2016). The risk of automation for jobs in OECD countries: A comparative analysis. OECD social, employment and migration working papers no. 189. DOI: 10.1787/5jlz9h56dvq7-en.

Barber BR. (1998). Three scenarios for the future of technology and strong democracy. Polit Sci Quart 113, 573–589.

Bauer MW, Allum N, and Miller S. (2007). What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. Public Underst Sci 16, 79–95.

Belloni M, and Kit B. (2017). James Cameron sounds the alarm on artificial intelligence and unveils a "Terminator" for the 21st century. The Hollywood Reporter. www.hollywoodreporter.com/features/james-cameron-sounds-alarm-artificial-intelligence-unveils-a-terminator-21st-century-1043027

Bijker WE, and Law J. (1992). *Shaping Technology/Building Society: Studies in Sociotechnical Change*. Cambridge, MA: MIT Press.

Bloor D. (1991). *Knowledge and Social Imagery*. Chicago: University of Chicago Press.

Bohannon J. (2015a). The synthetic therapist. Science 349, 250–251.

Bohannon J. (2015b). Fears of an AI pioneer. Science 349, 252.

Bowles N. (2018). Silicon valley nannies are phone police for kids. The New York Times. https://www.nytimes.com/2018/10/26/style/silicon-valley-nannies.html

Bradshaw JM, Hoffman RR, Woods DD, et al. (2013). The seven deadly myths of 'autonomous systems.' IEEE Intell Syst 28, 54–61.

Breland A. (2017). Elon Musk: We need to regulate AI before "it's too late." TheHill. https://thehill.com/policy/technology/342345-elon-musk-we-need-to-regulate-ai-before-its-too-late.

Brown N, and Sandholm T. (2018). Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. Science 359, 418–424.

Brynjolfsson E, and Mitchell T. (2017). What can machine learning do? Workforce implications. Science 358, 1530–1534.

Buchanan BG, and Shortliffe EH. (1985). *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison-Wesley.

Buolamwini J, and Gebru T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proc Mach Learn Res 81, 1–15.

Caliskan A, Bryson JJ, and Narayanan A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science 356, 183–186.

Callahan BR, Hathaway C, and Krishnamoorthy M. (2016). Quantitative metrics for generative justice: Graphing the value of diversity. Teknokultura 13, 567–586.

Campbell ND. (2005). Suspect technologies: Scrutinizing the intersection of science, technology, and policy. Sci Technol Human Values 30, 374–402.

Campbell M, Hoane AJ, and Hsu F. (2002). Deep blue. Artif Intell 134, 57–83.

Carroll JD. (1971). Participatory technology. Science 171, 647–653.

Char DS, Shah NH, and Magnus D. (2018). Implementing machine learning in health care—Addressing ethical challenges. N Engl J Med 378, 981–983.

Collingridge D. (1980). *The Social Control of Technology*. New York: St. Martin's Press.

Cook J. (2018). The YouTube Kids app has been suggesting a load of conspiracy videos to children. Business Insider. www.businessinsider.com/youtube-suggested-conspiracy-videos-to-children-using-its-kids-app-2018-3

Crevier D. (1993). *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York, NY: Basic Books.

Davenport TH. (2018). Can we solve AI's 'trust problem'? MIT Sloan Management Review. https://sloanreview.mit.edu/article/can-we-solve-ais-trust-problem/.

Dean J, and Fuller J. (2018). AI for social good. Google. https://www.blog.google/outreach-initiatives/google-org/ai-social-good/

Devaney J. (2015). Hawking, Musk sign letter warning of artificial intelligence "arms race." Newsmax. www.newsmax.com/Newsfront/Stephen-Hawking-Elon-Musk-letter-artificial-intelligence/2015/07/27/id/659161/

Didier C, Duan W, Dupuy JP, et al. (2015). Acknowledging AI's dark side. Science 349, 1064.

Dotson T. (2015). Technological determinism and permissionless innovation as technocratic governing mentalities: Psychocultural barriers to the democratization of technology. Engag Sci Technol Soc 1, 98–120.

Dotson T. (2017). *Technically Together: Reconstructing Community in a Networked World*. Cambridge: MIT Press.

Dreyfus HL. (1992). *What Computers Still Can't Do*. New York: MIT Press.

Duhigg C. (2018). Did Uber steal Google's intellectual property? The New Yorker. https://www.newyorker.com/magazine/2018/10/22/did-uber-steal-googles-intellectual-property

Dujmovic J. (2017). What's holding back artificial intelligence? Americans don't trust it. MarketWatch. www.marketwatch.com/story/whats-holding-back-artificial-intelligence-americans-dont-trust-it-2017-03-30

Dutton T, Barron B, and Boskovic G. (2018). Report. *Building and AI world: Report on national and regional AI strategies*. CIFAR. Toronto, Ontario, Canada.

Dzobo K, Rowe A, Senthebane DA, et al. (2018). Three-dimensional organoids in cancer research: The search for the holy grail of preclinical cancer modeling. OMICS 22, 733–748.

Eglash R, and Garvey C. (2014). Basins of attraction for generative justice. In: *Chaos Theory in Politics*. Banerjee S, Erçetin ŞŞ, and Tekin A, eds. Dordrecht: Springer Netherlands, 75–88.

Esteva A, Kuprel B, Novoa RA, et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118.

Eubanks V. (2017). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Fast E, and Horvitz E. (2017). Long-term trends in the public perception of artificial intelligence. In: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*. Palo Alto, CA: AAAI Press, 963–969.

Feenberg A. (2017). *Technosystem: The Social Life of Reason*. Cambridge, MA: Harvard University Press.

Feigenbaum EA, McCorduck P, and Nii P. (1988). *The Rise of the Expert Company: How Visionary Companies Are Using Artificial Intelligence to Achieve Higher Productivity and Profits*. New York: Times Books.

FLI. (2015). Open letter on autonomous weapons. Future of Life Institute. http://futureoflife.org/open-letter-autonomous-weapons/.

Floridi L, Cowls J, Beltrametti M, et al. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds Mach 28, 689–707.

Forsting M. (2017). Machine learning will change medicine. J Nucl Med 58, 357–358.

Forsythe DE. (2001). *Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence*. Stanford, California: Stanford University Press.

Frey CB, and Osborne MA. (2013). Report. *The future of employment: How susceptible are jobs to computerisation*. Oxford Martin School of Business. Oxford, UK.

Fujigaki Y. (2009). STS in Japan and East Asia: Governance of science and technology and public engagement. East Asian Sci Technol Soc Int J 3, 511–518.

Garvey C. (2018a). A framework for evaluating barriers to the democratization of artificial intelligence. In: *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*. Palo Alto, CA: AAAI Press, 8079–8080.

Garvey C. (2018b). AI risk mitigation through democratic governance: Introducing the 7-dimensional AI risk horizon. In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. New York: ACM, 366–367.

Garvey C. (2018c). Broken promises and empty threats: The evolution of AI in the USA, 1956–1996. Technology's Stories. www.technologystories.org/ai-evolution/

Garvey C. (2018d). Interview with Colin Garvey, Rensselaer Polytechnic Institute. Artificial intelligence and systems medicine convergence. OMICS 22, 130–132.

Garvey C, and Chard DS. (2018). Disrupting the AAA$. In: *Science for the People: Documents from America's Movement of Radical Scientists*. Schmalzer S, Chard DS, and Botelho A, eds. Amherst: University of Massachusetts Press, 1–236.

Gigerenzer G, and Goldstein DG. (1996). Reasoning the fast and frugal way: Models of bounded rationality. Psychol Rev 103, 650–669.

Google. (2019). Natural language API basics | cloud natural language API. Google Cloud. https://cloud.google.com/natural-language/docs/basics.

Graham L. (2017). Opening statement at hearing on Russian disinformation, extremist content online. United States Senator Lindsey Graham. https://www.lgraham.senate.gov/public/index.cfm/2017/10/graham-opening-statement-at-hearing-on-russian-disinformation-extremist-content-online

Guizzo E, and Ackerman E. (2016). When robots decide to kill. IEEE Spectr 53, 38–43.

Harding SG. (2006). *Science and Social Inequality: Feminist and Postcolonial Issues*. Urbana: University of Illinois Press.

Haugeland J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge, MA: MIT Press.

Hendler J. (2008). Avoiding another AI winter. IEEE Intell Syst 23, 2–4.

Hirschberg J, and Manning CD. (2015). Advances in natural language processing. Science 349, 261–266.

Hoffman, SG. (2017) Managing ambiguities at the edge of knowledge: Research strategy and artificial intelligence labs in an era of academic capitalism. Sci Technol Human Values 42, 703–740.

Hong L, and Page SE. (2012). Some microfoundations of collective wisdom. In: *Collective Wisdom: Principles and Mechanisms*. Landemore H, and Elster J, eds. New York: Cambridge University Press, 56–71.

Horvitz E. (2017). AI, people, and society. Science 357, 7.

Hunt MG, Marx R, Lipson C, et al. (2018). No more Fomo: Limiting social media decreases loneliness and depression. J Soc Clin Psychol 37, 751–768.

Jordan MI, and Mitchell TM. (2015). Machine learning: Trends, perspectives, and prospects. Science 349, 255–260.

Just MA, Pan L, Cherkassky VL, et al. (2017). Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. Nat Human Behav 1, 911.

Kahneman D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kaiser J. (2018). Cancer institute head touts big data and basic research. Science 359, 15.

Kardaras DN. (2016). Screens in schools are a $60 billion hoax. Time. http://time.com/4474496/screens-schools-hoax/.

Kasparov GK, and Greengard M. (2017). *Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins*. New York: PublicAffairs.

Konstorum A, Lynch ML, Torti SV, et al. (2018). A systems biology approach to understanding the pathophysiology of high-grade serous ovarian cancer: Focus on iron and fatty acid metabolism. OMICS 22, 502–513.

Landemore H. (2013). *Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many*. Princeton: Princeton University Press.

Lee TB. (2018). The hype around driverless cars came crashing down in 2018. Ars Technica. https://arstechnica.com/cars/2018/12/uber-tesla-and-waymo-all-struggled-with-self-driving-in-2018/.

Lewis P. (2018). "Fiction is outperforming reality": How YouTube's algorithm distorts truth. The Guardian. https://www.theguardian.com/technology/2018/feb/02/how-youtubes-algorithm-distorts-truth

Lindblom C. (1990). *Inquiry and Change: The Troubled Attempt to Understand and Shape Society*. New Haven: Yale University Press.

Lindblom C, and Woodhouse EJ. (1993). *The Policy Making Process*. Englewood Cliffs, NJ: Prentice Hall.

Ma B, Nahal S, and Tran F. (2015). *Robot Revolution: Global Robot and AI Primer*. Bank of America/Merrill Lynch. London, UK.

Martin CD. (1993). The myth of the awesome thinking machine. Commun ACM 36, 120–133.

Matacic C. (2018). Are algorithms good judges? Science 359, 263.

Matsuo Y. (2015). *Jinkō Chinō wa Ningen o Koeru ka: Dīpu Rāningu no Saki ni Aru Mono*. Tōkyō-to Chiyoda-ku: Kabushiki Kadokawa.

McCarthy J, Minsky ML, Rochester N, et al. (2006). A proposal for the dartmouth summer research project on artificial intelligence: August 31, 1955. AI Mag 27, 12–14.

McCorduck P. (1979). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*. San Francisco: W. H. Freeman.

McCorduck P. (2004). *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* (25th Anniversary Update). Natick, MA: A.K. Peters.

Meyer A, Zverinski D, Pfahringer B, et al. (2018). Machine learning for real-time prediction of complications in critical care: A retrospective study. Lancet Respir Med 6, 905–914.

Moravčík M, Schmid M, Burch N, et al. (2017). DeepStack: Expert-level artificial intelligence in heads-up no-limit poker. Science 356, 508–513.

Nadella S. (2016). The partnership of the future. Slate. www.slate.com/articles/technology/future_tense/2016/06/microsoft_ceo_satya_nadella_humans_and_a_i_can_work_together_to_solve_society.html

Nelkin D. (1987). *Selling Science: How the Press Covers Science and Technology*. New York: W.H. Freeman.

Nelkin D. (1989). Communicating technological risk: The social construction of risk perception. Ann Rev Public Health 10, 95–113.

Nelkin D. (1998). Scientific journals and public disputes. Lancet 352, S25–S28.

Nemitz P. (2018). Constitutional democracy and technology in the age of artificial intelligence. Philos Trans Royal Soc A Math Phys Eng Sci 376: 20180085.

News API. (2019). News API: A JSON API for live news and blog articles. https://newsapi.org.

Nicas J. (2018). How YouTube drives people to the internet's darkest corners. Wall Street Journal. https://www.wsj.com/articles/how-youtube-drives-viewers-to-the-internets-darkest-corners-1518020478/

NYT. (2019). The New York times developer network: All the APIs fit to POST. https://developer.nytimes.com/.

Özdemir V. (2019). Not all intelligence is artificial: Data science, automation, and AI meet HI. OMICS 23, 67–69.

Özdemir V, and Springer S. (2018). What does ''diversity'' mean for public engagement in science? A new metric for innovation ecosystem diversity. OMICS 22, 184–189.

Pew Research Center. (2014). Report. *AI, robotics, and the future of jobs*. Pew Research Center. Washington, DC.

Pew Research Center. (2017). Report. *Automation in everyday life*. Pew Research Center. Washington, DC.

Romm T, Harwell D, and Timberg C. (2018). Google CEO Sundar Pichai: Fears about artificial intelligence are 'very legitimate,' he says in Post interview. Washington Post. https://www.washingtonpost.com/technology/2018/12/12/google-ceo-sundar-pichai-fears-about-artificial-intelligence-are-very-legitimate-he-says-post-interview/

Šabanović S. (2014). Inventing Japan's 'robotics culture': The repeated assembly of science, technology, and culture in social robotics. Soc Stud Sci 44, 342–367.

Sarewitz D. (2011). The voice of science: Let's agree to disagree. Nature 478, 7.

Sarewitz D. (2015). CRISPR: Science can't solve it. Nature 522, 413.

Sarewitz D, and Woodhouse EJ. (2003). Small is powerful. In: *Living with the Genie: Essays on Technology and the Quest for Human Mastery*. Lightman AP, Sarewitz D, and Desser C, eds. Washington: Island Press, 63–83.

Scharre P. (2017). Why you shouldn't fear ''slaughterbots.'' IEEE Spectrum. https://spectrum.ieee.org/automaton/robotics/military-robots/why-you-shouldnt-fear-slaughterbots.amp.html.

Sclove R. (1995). *Democracy and Technology*. New York: Guilford Press.

Shead S. (2018). Tech journalists locked out at top AI conference. Forbes. https://www.forbes.com/sites/samshead/2018/12/08/tech-journalists-locked-out-at-top-ai-conference/

Shoham Y, Perrault R, Brynjolfsson E, et al. (2017). Report. *AI index 2017*. Artificial Intelligence Index. Stanford, CA.

Shoham Y, Perrault R, Brynjolfsson E, et al. (2018). Report. *The AI index 2018 annual report*. AI Index Steering Committee, Human-Centered AI Initiative, Stanford, CA: Stanford University. Stanford, CA.

Silver D, Huang A, Maddison CJ, et al. (2016). Mastering the game of Go with deep neural networks and tree search. Nature 529, 484–489.

Silver D, Hubert T, Schrittwieser J, et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Science 362, 1140–1144.

Silver D, Schrittwieser J, Simonyan K, et al. (2017). Mastering the game of Go without human knowledge. Nature 550, 354–359.

Simonite T. (2018). AI is the future—But where are the women? Wired. https://www.wired.com/story/artificial-intelligence-researchers-gender-imbalance/

Snow J. (2018). ''We're in a diversity crisis'': Cofounder of Black in AI on what's poisoning algorithms in our lives. MIT Technology Review. https://www.technologyreview.com/s/610192/were-in-a-diversity-crisis-black-in-ais-founder-on-whats-poisoning-the-algorithms-in-our/

Somerville H, and Levine D. (2017). Uber-Waymo trial delayed as U.S. judge raises prospect of ''cover-up.'' Reuters. https://www.reuters.com/article/us-alphabet-uber-ruling-idUSKBN1DS26X

Stilgoe J, Lock SJ, and Wilsdon J. (2014). Why should we promote public engagement with science? Public Underst Sci 23, 4–15.

Stone P, Brooks RA, Brynjolfsson E, et al. (2016). Report. *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence*. Stanford, CA: Stanford University.

Strange A. (2015). Elon Musk, Stephen Hawking warn of artificial intelligence dangers. Mashable. http://mashable.com/2015/01/13/elon-musk-stephen-hawking-artificial-intelligence/.

SYZYGY. (2017). Report. *Sex, lies and A.I.: How Americans feel about artificial intelligence; What marketers need to know*. SYZYGY. New York, NY.

Tableau Software. (2019). Tableau: Business intelligence and analytics software. https://www.tableau.com/

Taddeo M, and Floridi L. (2018). How AI can be a force for good. Science 361, 751–752.

Taube M. (1961). *Computers and Common Sense: The Myth of Thinking Machines*. New York: Columbia University Press.

Toyama K. (2015). *Geek Heresy: Rescuing Social Change from the Cult of Technology*. New York: Public Affairs.

Tufekci Z. (2018). YouTube, the Great Radicalizer. The New York Times. https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html

Turel O. (2019). Potential 'dark sides' of Leisure technology use in youth. Commun ACM 62, 24–27.

Twenge JM, Cooper AB, Joiner TE, et al. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. J Abnorm Psychol 128, 185–199.

Twenge JM, Joiner TE, Rogers ML, et al. (2018). Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. Clin Psychol Sci 6, 3–17.

Wakabayashi D. (2018). Self-driving Uber car kills pedestrian in Arizona, where robots roam. The New York Times. https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html

Wakefield J. (2018). Google accused of ''trust demolition'' over health app. BBC News. https://www.bbc.com/news/technology-46206677

Williams C. (2015). AI guru Ng: Fearing a rise of killer robots is like worrying about overpopulation on Mars. The Register. www.theregister.co.uk/2015/03/19/andrew_ng_baidu_ai/

Winner L. (1980). Do artifacts have politics? Daedalus 109, 121–136.

Woodhouse E, Hess D, Breyman S, et al. (2002). Science studies and activism: Possibilities and problems for reconstructivist agendas. Soc Stud Sci 32, 297–319.

Woodhouse E, and Patton JW. (2004). Design by society: Science and technology studies and the social shaping of design. Des Issues 20, 1–12.

Woodhouse E, and Sarewitz D. (2007). Science policies for reducing societal inequities. Sci Public Policy 34, 139–150.

World Economic Forum. (2018). Report. *Towards a reskilling revolution: A future of jobs for all*. World Economic Forum. Geneva, Switzerland.

Wynne B. (1992). Misunderstood misunderstanding: Social identities and public uptake of science. Public Underst Sci 1, 281–304.

Wynne B. (2001). Creating public alienation: Expert cultures of risk and ethics on GMOs. Sci Cult 10, 445–481.

Wynne B. (2006). Public engagement as a means of restoring public trust in science—Hitting the notes, but missing the music? Community Genet 9, 211–220.

Ziad O, and Emanuel EJ. (2016). Predicting the future—Big data, machine learning, and clinical medicine. N Engl J Med 375, 1216–1219.

Ziman J. (1991). Public understanding of science. Sci Technol Human Values 16, 99–105.

Address correspondence to:
*Colin Garvey, PhD Candidate*
*Science and Technology Studies Department*
*Rensselaer Polytechnic Institute*
*110 Eighth Street, Sage Labs 5710*
*Troy, NY 12180*

*E-mail:* colin.k.garvey@gmail.com

**Abbreviations Used**

AI = artificial intelligence
API = application programming interface
GOFAI = good-old-fashioned artificial intelligence
HI = human intelligence
IRB = institutional review board
ML = machine learning
NLP = natural language processing
NYT = The New York Times