Rapidly Testing Policymakers' Assumptions Using Social Media: A Systematic Approach Applied to Public Health

by Dian Hu

B.S. in Systems Engineering, The George Washington University

A Dissertation submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

May 15, 2022

Dissertation directed by

David. A. Broniatowski
Associate Professor of Engineering Management and Systems Engineering

The School of Engineering and Applied Science of The George Washington University certifies that Dian Hu has passed the Final Examination for the degree of Doctor of Philosophy as of April 1, 2022. This is the final and approved form of the dissertation.

Rapidly Testing Policymakers' Assumptions Using Social Media: A Systematic Approach Applied to Public Health

Dian Hu

Dissertation Research Committee:

David. A. Broniatowski, Associate Professor of Engineering Management and Systems Engineering, Dissertation Director

Thomas A. Mazzuchi, Professor of Engineering Management and Systems Engineering, Committee Member

Erica L. Gralla, Associate Professor of Engineering Management and Systems Engineering, Committee Member

Cindy M. Liu, Associate Professor of Environmental and Occupational Health, and of Microbiology, Immunology, and Tropical Medicine, Committee Member

John Helveston, Assistant Professor of Engineering Management and Systems Engineering, Committee Member

© Copyright 2022 by Dian Hu
All rights reserved

Dedication

I dedicate my dissertation work to my family and many friends. A special feeling of gratitude to my loving parents, Jinyou Hu and Meizhen Zhong, whose unconditional love has encouraged me all these years.

I want to thank my grandfather, who wished to remain anonymous. He has acted as a role model of engineer and scholar since my childhood.

Acknowledgments

Throughout my doctoral years, I have encountered many people who had a positive effect on my research and career, and I am grateful for everything and everyone at SEAS and GWU. Especially all of my instructors for every class.

First, I would like to thank my advisor, Dr. David. A. Broniatowski, for his countless hours of reading, mentoring, encouraging, and patience throughout the entire seven years, without whom I would probably still be a front-end data engineer. I wish to thank Dr. Thomas. A. Mazzuchi, my undergraduate advisor and department chair, for his mentoring, leadership, and consistent encouragement for my entire academic career in the U.S. I wish to thank Dr. Erica L. Gralla for her instructions in several critical courses in research methodology and engineering management. I wish to thank Dr. John Helveston for his insightful suggestions for career transition. And a special thanks to Dr. Cindy. M. Liu for her mentorship and support for my recent publication.

I would like to acknowledge and thank The School of Engineering and Applied Science for allowing me to conduct my research and providing any assistance requested. Finally, I would also like to thank Dr. Issac Chang and Dr. Mitra Ahadpour for their mentorship while I was working for FDA as ORISE Fellow.

Abstract of Dissertation

Rapidly Testing Policymakers' Assumptions Using Social Media: A Systematic Approach Applied to Public Health

Policymakers in public health have been facing growing trouble recently: on the one hand, the public's trust in policymakers has been declining over the years. On the other hand, the turmoil of international politics has created an unpredictable environment for policy consistency. Much of these debates have originated, developed, escalated, and perpetuated on social media. Naturally, to understand the public's opinion of specific health issues, many researchers in public health have turned to seek answers using social media data directly.

However, studying social media is not a trivial task. Scholars in qualitative and quantitative methods have identified many problems and fallacies in social media research. In this dissertation, I will systematically review three main projects I finished and published during my Ph.D. years and discuss how my works are able to rapidly test several different hypotheses that emerged among researchers and decision makers of each field.

Table of Contents

Dedication iv

Acknowledgments v

Abstract vi

List of Figures ix

List of Tables x

Chapter 1: Introduction 1
- 1.1 Communication Between Policymakers and The Public 1
- 1.2 The Impact of Social Media on Policy Implementation 3

Chapter 2: Problems Encountered by Scholars from Different Disciplines in Social Media Studies 5
- 2.1 Big Data vs. Small Data 5
- 2.2 Applying Topic Modeling to Social Media Study 7
- 2.3 Hypotheses of Social Media studies are often Biased 8
- 2.4 The Need for Rapid Testing Policymakers' Assumptions 10

Chapter 3: Human Papillomavirus Vaccine Discourse on Reddit 12
- 3.1 Introduction of the HPV Project 12
- 3.2 Key Literature Foundation of HPV Project 12
- 3.3 Methods of HPV Project 13
  - 3.3.1 Data Source 13
  - 3.3.2 Data Collection 14
  - 3.3.3 Data Analysis for Research Question 1 14
  - 3.3.4 Data Analysis for Research Question 2 15
  - 3.3.5 Estimate the Gender and Age 16
- 3.4 Results of HPV Project 17
  - 3.4.1 Are Reddit messages more related to cancer risks or to sexual behavior? 17
  - 3.4.2 What other topics characterize the discussion on HPV vaccination on Reddit? 19
- 3.5 Key Findings of HPV Project and their Implications to this dissertation 22

Chapter 4: Vaccine Hesitancy in China after 2018 Changchun Changsheng Vaccine Incident using Sina Weibo data 23
- 4.1 Introduction of the Sina Weibo Project 23
- 4.2 Key Literature Foundation of Sina Weibo Project 24
- 4.3 Methods of Sina Weibo Project 25
  - 4.3.1 Study 1: Qualitative Analysis of Messages During the 2018 Vaccine 26
  - 4.3.2 Study 2:Qualitative Analysis of Messages One Year After the Scandal 27

4.3.3 Study 3: Topic Model Analysis **27**
4.4 Results of Sina Weibo Project **28**
4.4.1 Result of Study 1: 2018 Data **28**
4.4.2 Result of Study 2: 2019 Data **29**
4.4.3 Result of Study 3: Results Replicate Across Entire Corpus using LDA **31**
4.5 Key Findings of Weibo Project and their Implications to this dissertation **34**
Chapter 5: Study "Yelp Effect" using Yelp and Google Maps data **36**
5.1 Introduction of the "Yelp Effect" Project and Literature Review **36**
5.2 Methods of "Yelp Effect" Project **39**
5.2.1 Data Collection **39**
5.2.2 Data Analysis using LDA **42**
5.2.3 Qualitative Annotation **41**
5.3 Key Results of "Yelp Effect" Project **43**
5.3.1 LDA result **43**
5.3.2 Annotation Results **46**
5.4 Key Findings of Yelp Project and their Implications to this dissertation **47**
Chapter 6: Systematic Comparison of the key methods and findings of the three projects **49**
6.1 A Systematic Comparison **49**
6.2 Limitations **52**
Chapter 7: Conclusions and Future works **54**
Chapter 8: Citations **57**
Appendix A: Qualitative codebook of the HPV Project **73**
Appendix B: Development Strategy of the Gender Classifier **74**
Appendix C: Guideline for Annotating the Weibo Vaccine Dataset **75**
Appendix D: Google Maps Reviews Collection Algorithm **79**

List of Figures

4-1 Trends of 6 categories in July 2018 **29**
4-2 Shift of Distribution of 6 categories, Pre-incident Data vs. 2019 Data **30**
4-3 6 Aggregated Topics from LDA Model **34**
5-1 Total number of reviews in US counties (Yelp data) **40**
5-2 Mean review stars in US counties (Yelp data) **41**
5-3 Total number of reviews in US counties (Google Maps data) **41**
5-4 Mean review stars in US counties (Google Maps data) **41**
5-5 Topics generated by the Latent Dirichlet Allocation algorithm based on the text of all Google reviews in our data set **44**

List of Tables

3-1 Summary of estimated demographic information of the subjects of the HPV project **17**
3-2 Summary of human papillomavirus (HPV) vaccine topics generated by the Latent Dirichlet Allocation (LDA) approach **20**
4-1 Time-period distribution of 2018 data **26**
4-2 LDA model summary of Weibo Project: Assigned Category, Top Keyword in Chinese, and their English Translation **31**
5-1 LDA model summary of Google Maps Data **45**
5-2 Primary Message Categories of the 200 Messages in 7 Categories **47**
6-1 Comparison of my three published projects during my Ph.D. years **49**
A-1 Qualitative codebook of the HPV Project **73**

Chapter 1: Introduction

1.1 Communication Between Policymakers and The Public

Policymakers in public health have been facing growing trouble recently: on the one hand, the public's trust in policymakers has been declining over the years (Cairney & Wellstead, 2021; Fairbrother, 2017). On the other hand, the turmoil of international politics and surge of populism has created an unpredictable environment for policy consistency (Algan et al., 2017; Rodrik, 2021; Staerklé & Green, 2018).

When analyzing the cause of the distrust between the public and the policymakers, the first reaction of some media and activist groups is to blame the entire situation on the growing populism movement and conspiracy theory (Lasco & Larson, 2020). Nevertheless, several scholars have argued that it is unconstructive or even detrimental to do so: Nyhan et al. (2016) found that many conspiracy beliefs were first developed based on the over-classified or redacted materials that were heard by the public, therefore arguing that insufficient government transparency is indeed a primary factor associated with several widespread conspiracies. Jasanoff and Simmet (2017) argued that public reasoning was imperative progress towards public acceptance of new technology even though those reasonings are often biased or erroneous; Therefore, scholars should not describe the current time as "post-truth" age. Similarly, Larson and Broniatowski (2021) argued that communication about vaccine hesitancy should be delivered in an empathic tone; an overly aggressive and critical doctrine may drive vaccine hesitatnt members of the public towards anti-vaccine positions. Wolf (2021) found that a recent conflict between Belgium authority and local residents on a proposed highway escalated quickly only after the authority labeled and dismissed objecting the public in an offensive stance. In recent years, many have

noticed the communication problem between the public and the policymakers. Some argue that the current democracy should embrace more hybrid and participatory involvement from the public (Callon et al., 2011; Guston, 2014). Several of these inclusive frameworks of democratic governance have been suggested or partially implemented to meet the emerging issues in nanotechnology (Macnaghten et al., 2005), crisis management (Fujigaki & Tsukahara, 2011), gene editing (Jasanoff et al., 2015), and climate change (Jordan et al., 2018).

However, the problems are still more significant than the remedies today: Gilens and Page (2014) found that wealthy elites and business interest groups have significant impacts on U.S. policy; meanwhile, average citizens and groups representing the working class have little influence on the policy making. One study of U.S. state legislative politicians found that both Republicans and Democrats have overestimated the constituents' support for conservative policies as the politicians are more likely to be directly contacted by conservative activists (Broockman & Skovron, 2018). Another study of the legislative staffer of U.S. Congress indicates that most staffers do not have the resource to investigate their constituents' preference of specific policy directly; instead, they often seek consultation from preferred think tanks and interest groups to draft the policy (Hertel-Fernandez et al., 2019). Another study of the politicians in Belgium, Canada, and Israel indicates that politician's perceptions of the accuracy and importance of political topics are less aligned with the majority of their constituents, yet more aligned with the "male, highly educated, and politically interested citizens" (Sevenans et al., 2021). The miscommunication often leads to tragic results in public health. One interview-based study of the opioid crisis in New York state found the policy makers of the state are confident

that the state has more than enough beds for treatment of overdosed patients, yet the public cannot access those spare beds because of health insurance limitations, admission criteria, and family and social constraints (Strach et al., 2020).

The communication is complicated in the other direction as well. Broockman et al. (2021) found that interest groups can also easily steer the voter's attention to selected topics. Peterson and Allamong (2021) found that media coverage can often set the tone and the frame of a policy issue, therefore influencing the voter's opinion. Sherman et al. (2021) found that activists are less likely to discuss the long-term benefits and costs of a policy; Instead, they are more likely to label the passing/rejecting of the policy as a victory. Phoenix et al. (2019) further argued that social researchers must present research results as "knowledge brokers" between the public and the policy makers to replace the biased activists.

How can policymakers listen to the public more efficiently? How can we target the desired population that is most relevant to a policy quickly and affordably? Many have attempted to study social media data directly to answer this question.

1.2 The Impact of Social Media on Policy Implementation

Social media has brought enormous volumes of information into everyone's life. In January 2020, Reddit reported that the platform has 52 million active daily users (Baumgartner et al., 2020). Facebook reported that it had 1.85 billion daily active Facebook users at the end of 2020 (Facebook, 2021). At the same time, Twitter has reported having 199 million daily active users as of April 2021(Twitter, 2021).

When I started my Ph.D. study in 2015, social media was still considered a supplement

of traditional media; not an equal. However, as of 2021, the power of social media has eclipsed traditional media. The impact of social media has reached the stock market (Costola et al., 2021; Jin et al., 2013), political elections (Grinberg et al., 2019), vaccine refusal (Dredze et al., 2016) , climate change and nuclear waste disposal (Kahan et al., 2017). While my research has been focused on public health, it is imperative to acknowledge that social media has the potential to impact the implementation of any policy nowadays. Even a country with a strong censorship has become aware that the policy resistance on social media can outbreak in a significant magnitude (Xu et al., 2011; Fu et al., 2013;). However, studying social media is not a trivial task. Social media study is neither omniscient nor a replacement of other traditional methods. Although scholars from both qualitative and quantitative backgrounds have acknowledge the value of social media study, they have also identified many problems and fallacies in social media research at the same time.

Chapter 2: Problems Encountered by Scholars from Different Disciplines in Social Media Studies

2.1 Big Data vs. Small Data – A Debate Originated by Qualitative Scholars

Scholars from qualitative studies were among the first to notice the most obvious problems. When social media studies are being introduced in earlier days, they were often intertwined with the concept of "big data" (Labrinidis & Jagadish, 2012). Although it often comes with different definitions, big data is commonly defined by having extremely high value in three dimensions: high-volume, high-velocity, and high-variety (Gandomi & Haider, 2015). In recent years, a new dimension: veracity, is sometimes added to describe the emergence of fake content within the information (Abbasi et al., 2016). As the concept of big data become popular, a comparative concept has also risen among literature: small data (Miller, 2010), a term intended to describe all studies whose data is collected in carefully controlled sampling techniques and smaller volumes.

Naturally, a debate has arisen in the literature on whether big data studies can already replace small data studies nowadays or sometime in the future (Mills, 2018). Kitchin and Lauriault (2015) argued that although big data is higher on velocity and variety, it does not necessarily guarantee that big data are more suited to answer all types of research questions, as most research questions in social science and psychology are often framed to be tested only in low-variety and low variety settings. boyd and Crawford (2011) argued that some analyses of earlier qualitative studies analyzing census data have already reached or surpassed the size of many of today's "big data" studies, and scientific and meticulous sampling used in earlier days are still relevant in "big data" era.

Grover et al. (2020) argue that most big data studies often better describe the landscape

of a situation, not answering why the situation has emerged. Similarly, Abbasi et al. (2016) has argued that big data is better suited to answering the questions of "what," "where," and "when" of the emerging problems, but not on "why" or "how" that problems emerged. Chandler (2015) has even strongly argued that the big data hype had oversimplified human society into "a world without causation," which should be a concerning development in the study of social science. It is essential to point out that the questions that cannot be solved by the big data method are often the critical contributions valued by qualitative scholars.

Furthermore, most big data is created, stored, and distributed by giant technology corporations using customizable APIs that favors certain groups of scholars more than others (Dalton et al., 2016). Although these API-based data are often accessible and affordable, most APIs do not describe their methods and biases thoroughly (Joseph et al., 2014). Their black-box nature has raised concerns about transparency and replicability (Lazer et al., 2014). A majority of pre-2014 big data paper has used Twitter data simply because of the API is convenient and affordable to use (Tufekci, 2014). A recent study by Tromble (2021) has also pointed out that studies based on earlier API can become inaccessible due to API upgrades or deprecation. If not fixed, the lack of transparency and accessibility of these APIs will divide qualitative researchers and hinder the possibility of re-analysis and collaboration.

Overall, although many qualitative scholars have acknowledged the potential and power of the big data approach in social media study, most of them have also realized the severe limitation of this approach and therefore have concluded that rigorous qualitative analysis cannot be replaced by existing big data techniques.

2.2 Applying Topic Modeling to Social Media Study - Searching for Meaning and Validity

Although the concern among qualitative scholars is valid and reasonable, researchers simply cannot afford to annotate every single message on social media. Therefore, to understand the whole landscape of social media, we still need to incorporate some computational methods. Topic modeling is one approach to do so.

Topic modeling is a family of probabilistic algorithms that aims to discover the underlying topics (themes) of unstructured text documents in large volumes (Blei, 2012). The most frequently used algorithm of this type is Latent Dirichlet Allocation (Blei et al., 2003). Although the field of Natural Language Processing has developed significantly in the last two decades, LDA algorithms can still be implemented together with the latest word embedding techniques like Word2Vec (Mikolov et al., 2013) and Global vectors for word representation (GLoVe; Pennington et al., 2014). Even result from multi-context language model that were pre-trained in deep learning like Bert (Devlin et al., 2018) can be used as the input for LDA algorithms (Grootendorst, 2022).

Application of topic modeling has been used on text analysis in many fields, including but not limited to legislative debate (Grimmer, 2010; Quinn et al., 2010), newspaper articles (DiMaggio et al., 2013), crypto market policy (Munksgaard & Demant, 2016). Naturally, it has also been used to analyze text messages on various social media like blogs (Dehghani et al., 2014), Twitter (Kim et al., 2016) and Sina Weibo (Wang et al., 2014).

While more and more studies are using LDA to understand social media messages, some scholars have become increasingly aware of the limitation of the LDA output. Grimmer and Stewart (2013) argued that some output from the LDA models may be misleading for

political science researchers, as the arguments from rival politicians looked the same in the LDA output. Chang et al. (2009) found that results from topic models can still have low interpretability even when the model has low perplexity, a key measurement for the LDA model. Boussalis and Coan (2016) argued that reconstructing the semantic structure of the original messages might improve the interpretability of the LDA model. Similarly, DiMaggio et al. (2013) have also argued that validation of the LDA model should include semantic validity besides statistical validity. In a recent study, Morstatter & Liu (2018) incorporated crowdsourcing techniques to ask annotators whether a topic is coherent by manipulating the order and content of LDA topic words.

Although these attempts have improved the interpretability of the LDA output, there is still limited consensus regarding the standard of validation of the LDA model. Boussalis and Coan (2016) researchers using the LDA model should consult experts of respective fields during validation if necessary. Xie et al. (2015) has suggested incorporating common-used word correlation to improve the performance of the LDA model.

2.3 The Misinformation on Misinformation, and the Echo Chamber of Echo Chambers – Hypotheses of Social Media studies are often Biased

Other than the limitation on the methods side, some in recent years have noticed that scholars studying the misinformation on social media are often under the influence of media and political activists in the first place (Nyhan, 2020; Anderson, 2021). Moreover, an ill-designed observational study on social media without rigorous design will only get findings that are biased (Kulshrestha et al., 2017), exaggerated (Guess et al., 2018) or even manipulated by malicious trolls and bots (Broniatowski et al., 2018).

There is no doubt that misinformation is present on social media (Kapantai et al., 2021). A 2019 survey from Pew reports that 50% of the subjects agree that "Made-up news" is "a very big problem" in the U.S. (Mitchell et al., 2019). However, many recent studies have discovered that misinformation on social media is not as significant as suggested by several media had reported; For example, Grinberg et al. (2019) found that 1% of Twitter users were exposed to 80% of fake news (the exposure being defined as viewing), and 0.1% of users were responsible for sharing 80% of fake news. Similarly, Guess et al. (2019) found that 91.5% of the studied subjects have never shared any fake news, and only 0.4% of the subjects have shared fake news more than five times. Another highly-cited paper in Science (Vosoughi et al., 2018) reported that news from verified sources would take more than 6 times to reach a certain level of popularity compared to unverified news. Nevertheless, this paper also reported that a majority of "unverified news" is merely unconfirmed celebrity gossip rather than misinformation (Kulshrestha et al., 2017) of scientific or political topics.

A report from Buzzfeed concluded that 20 fake news stories on Facebook have accumulated nearly 9 million during the 2016 U.S. presidential campaign (Silverman, 2016). However, Watts and Rothschild (2017) had argued that although the aforementioned 9 million interactions are entirely trivial compared to the daily volumes of the interaction of 1.5 billion Facebook users in 2016. Furthermore, Rogers (2020) found out that all news from "Fox News" was defined false outright in the Buzzfeed report, and the result would in fact favor true news if reanalyzing the same data by relabeling "Fox News" as true news. Wagner and Boczkowski (2019) pointed out that earlier studies using "sharing" and "commenting" of fake news as the approximate definition of "believing" or "supporting" is a very biased and risky measurement, as a majority of these comments and sharing were

actually non-believers actively ridiculing and arguing against the source of the fake news. Another study by Fletcher & Nielsen (2018) argue that users on social media are not gullible or unsophisticated, instead, they are usually skeptical of the content they encountered. Similarly, Glenski et al. (2019) found that Reddit users' understanding and discussion of cryptocurrency are ahead of most researchers before 2018.

2.4 The Need for Rapid Testing Policymakers' Assumptions

In the previous literature review section, we have witnessed that researchers often overestimate or underestimate certain effects or hypotheses. Yet, if we are able to incorporate topic modeling techniques correctly, we are able to rapidly estimate the scale and magnitude of certain topics. After we obtained the estimation of the model, we are able to decide whether it is worth the time and effort to further annotate the meaning of these messages using qualitative methods.

In this dissertation, I will systematically review three main projects I finished and published during my Ph.D. years and discuss my understanding of the necessary steps and strategies for this type of work. Furthermore, I want to argue that when researchers design social media studies, they must attempt to understand the following questions to understand the messages scientifically. Although answering them can be challenging or impossible in some situations, they are the foundational bricks of validity of these studies.

(1) Who/what the author of the message is? (Activists, regular user, business/organizational users, official bots, malicious bots, malicious actor)

(2) What is the context/vehicle of the message? (Post, Comment, Tweet, Review)

(3) Who/what the audience of the message is? (Other users, followers, all users within

a sub-forum/subreddit). These three questions are essential to ensure the messages studied are indeed from the subjects the researchers are aiming for, and the discourse is the right type.

I will perform a systematic review of my main research projects during my Ph.D. student years in the following chapters.

Chapter 3: Human Papillomavirus Vaccine Discourse on Reddit

The following chapter summarizes our paper published in 2019 regarding HPV vaccine hesitancy on Reddit (Lama et al., 2019), where I was listed as the second author. The content is slightly abridged compared to the published version. The qualitative codebook is presented in appendix A. The strategy of demographic classifier is presented in appendix B.

3.1 Introduction of the HPV Project

Long-lasting infections with high-risk human papillomaviruses (HPVs) can cause cancer in parts of the body where HPV infects cells, such as in the cervix, oropharynx, anus, rectum, penis, vagina, and vulva (Jena et al., 2015). HPV vaccination acceptance is a critical public health issue as low completion rates of the vaccine series place adolescents and young adults at risk for HPV-associated infection and cancers (Friedman & Shepeard, 2007). The multiple factors that influence vaccine acceptance have been examined on traditional mass media and social media, including the understudied platform Reddit. This study is the first to examine HPV discourse on Reddit, which is intended to provide insight about attitudes of HPV vaccination among a unique population that may not be readily accessible in traditional data surveillance methods (Dredze et al., 2016)

3.2 Key Literature Foundation of HPV Project

The previous conversation on HPV vaccine hesitation by the public had suggested that the main reason why parents rejected the HPV vaccine is the concern of promiscuity (Casciotti et al., 2014; Quintero Johnson et al., 2011). Specifically, media coverage has

often focused on one possible mechanism: adolescents who had HPV vaccine are more likely to engage in risky sexual behavior (Forster et al., 2010). Furthermore, several preliminary studies have found that parents with female adolescents are more likely to develop hesitation against HPV vaccine because of the concern of promiscuity, therefore less likely to vaccinate their children for the HPV vaccine (Johnson et al., 2017; Reiter et al., 2009). However, Daley et al. (2016) argued that the belief that female adolescents' promiscuity is the main barrier of HPV vaccine is exaggerated; parents of male adolescents have also shown increasing hesitancy. Parents have also used the internet to express other concerns that were often unmentioned by traditional surveys (McRee et al., 2012).

Therefore, the competing arguments presented by the literature motivated our investigation of the following two questions for this paper:

Research Question 1: Are Reddit messages more related to cancer risks or to sexual behavior?

Research Question 2: What other topics characterize the discussion on HPV vaccination on Reddit?

3.3 Methods of HPV Project

3.3.1 Data Source

We used Reddit as our main data source. Unlike microblogging social media platforms such as Twitter, Reddit does not limit message length, allowing users to express their opinions in depth. Communities can be organized by different topics through "subreddits," which are forums dedicated to a specific topic such as news (/r/news) or HPV (/r/HPV). Reddit also has a strong community-based moderation culture; repetitive messages or

unrelated messages will often be removed by moderators of each subreddit. Although it is possible that messages in our sample are due to malicious actors, such as bots and trolls (Ferrara et al., 2016; Cresci et al., 2019), Reddit's moderation culture also helps to mitigate these concerns (Bergstrom, 2011; Sanfilippo et al., 2017). Therefore, Reddit comments are a promising data source for understanding how a segment of the online community perceives the risks of HPV and the risks and benefits of the HPV vaccine.

3.3.2 Data Collection

We downloaded all recent Reddit comments from January 1, 2007, to September 30, 2015, using a platform named Pushshift to collect and share complete Reddit datasets for research purposes (Baumgartner et al., 2020). At the time of analysis in 2016, the most up-to-date data were culled, resulting in our use of data from 2007 to 2015. From the total set of all messages, we identified a subset of HPV-related messages by filtering all comments containing the text "hpv" or "human papillomavirus". From the set of hpv-related messages, we further identified a subset of vaccine-related messages by filtering messages for the token "vaccin."

3.3.3 Data Analysis for Research Question 1

Our aim was to determine whether HPV-related messages discussed sexual behavior or cancer risk more often. First, 2 annotators manually annotated 50 messages and selected an initial set of keywords associated with both topic areas. The same annotators then annotated another 350 messages, in rounds of 50 each, adjusting the keywords and refining coding techniques. We agreed upon a final set of keywords as follows: a single keyword, "cancer," was sufficient to identify messages pertaining to cancer. Messages pertaining to sexual behavior included keywords derived from the "purity" category in the Moral

Foundation Dictionary (Graham et al., 2009): "piety, pious, purity, pure, clean, sterile, sacred, chaste, saint, innocent, unclean, slut, whore, dirty, impiety, impious, profane, promiscuity, promiscuous, adultery, unchaste, sexual, sex, intercourse, coitus, lovemaking, and premarital." Using these specific keywords, annotators initially began annotating for promiscuity; however, owing to a limited number of messages, this topic was expanded to include all sexual activity related to HPV (including promiscuity). Through this iterative process, we created a codebook to identify messages as either pertaining to cancer risk, sexual activity, or others. There was moderate agreement on messages indexing sexual behaviors, Cohen kappa=0.62 (95% CI 0.54 to 0.69), and high agreement on messages indexing cancer risk, kappa=0.81 (95% CI 0.74 to 0.88).

Next, we employed Latent Semantic Analysis (LSA) (Deerwester et al., 1990), a commonly used technique in natural language processing, to measure the semantic similarity between each Reddit message and the keyword lists. Specifically, we rendered the set of all HPV-related Reddit messages into a term-document matrix and transformed this corpus into a "latent semantic space." When preparing the corpus for the LSA analysis, we removed all but the 500 most frequent unigrams. We next applied term frequency-inverse document frequency weighting to the corpus. Finally, consistent with the past literature, we retained dimensions 2 to 101 when computing the Singular Value Decomposition underlying LSA. We used the underlying LSA space to measure the cosine similarity between each Reddit post and each keyword list and calculated the average similarity score per month.

3.3.4 Data Analysis for Research Question 2

To further explore the topics discussed on Reddit, we used the aforementioned Latent

Dirichlet Allocation (LDA) model to automatically segment Reddit messages into 10 probabilistic "topics" (using unigram features; hyperparameter values were alpha=.1 and η=0.01). LDA assigns each word token to one of these topics, allowing one to summarize the content of each Reddit post. After examining models with 20 and 50 topics manually, we determined that a 10-topic model best captured the variety of discussion subjects with minimal redundancy.

3.3.5 Estimate the Gender and Age of Individual Reddit Users

Because the literature has raised specific concerns regarding the female and male difference in HPV vaccine hesitancy, we have to develop two classifiers to assess the gender and age of the user posting each message. I was the main designer and implementor of these two classifiers in this project.

This gender classifier helped us understand the overall difference between male and female user discussions related to the HPV vaccine. The classifier works by identifying gender-indicating keywords and expressions included in Reddit messages. For instance, statements such as "As a man…" or "I am a straight woman…" are used as indicators of a user's gender. In some instances, there was not enough information to identify a user's gender with confidence, so those users were classified as having an "unknown" gender.

We manually annotated over 100 messages to test the accuracy of the classifier. A total of 2 annotators read messages identified by the classifier and determined that 43 of 50 messages were correctly classified as "male," and 45 of 50 messages were correctly classified as "female." This indicates an overall accuracy of 88%

We also developed an "age classifier" to assess the age of users into 7 age categories (10 to 18, 18 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, and 65 and older) consistent with

Reddit surveys. When reading the gender-indicating expressions (e.g., "I am 27 years old"), we observed that the linguistic features related to age-indicating expressions were sometimes complicated and open to misclassification. For example, some expressions discussing time or gestational age were coded as user ages, resulting in the classifier achieving 67% accuracy. Nevertheless, the accuracy is much higher than the random baseline of 7 categories (14.28%). Table 3-1 summarized the estimated demographic information obtained by our classifier.

Table 3-1 Summary of estimated demographic information of the subjects of the HPV project

Age group (years)	Female users	Male users	Unspecified gender	Age subtotal
10 to 18	128	339	57	524
18 to 24	293	862	291	1446
25 to 34	376	1489	297	2162
35 to 44	103	549	65	717
45 to 54	42	195	26	263
55 to 64	7	37	2	46
65 and above	7	13	0	20
Unspecified age	187	3385	4202	7774
Gender subtotal	1143	6869	4940	12952

3.4 Results of HPV Project

3.4.1 Are Reddit messages more related to cancer risks or to sexual behavior?

Our final dataset included 22,729 messages filtered to include "hpv" or "human papillomavirus" and "vaccin". Mitigating concerns about behavior by malicious actors, such as bots, we found that no two messages in our dataset are exactly the same, a common

indicator of bot behavior (Gilbert, 2013). The number of messages varied by year and increased over time, from 162 messages in 2007 to a peak of 5311 messages in 2015. Among those posting HPV vaccine messages (n=12,952), 53.0% (6869/12,952) were male, 8.8% (1143/12,952) were female, and 38.1% (4940/12,952) were of unknown gender (Table 3-1). Nearly 17% of users were aged from 25 to 34 years, followed by 11.2% of users aged 18 to 24 years, although 60.0% of users did not specify their age.

In manual annotation, it was clear that messages describing cancer risk were more identifiable than messages discussing HPV-related sexual behavior. Cancer risk messages tended to focus on specific themes, including the relationship between HPV infection and cancer incidence (especially related to specific HPV strains), and on HPV vaccination as a resource to prevent future cancer incidence. Messages about sexual behaviors were more varied. Initially, we attempted to code for instances strictly related to the HPV vaccine and sexual promiscuity but found very few instances (<5 in the first 100 messages) that we were unable to do so. The widespread use of sarcasm and irony on Reddit also made it difficult to know the meaning of some messages out of context. Many of the messages we encountered incorporated the moral foundation keywords, but only to refute the idea that the HPV vaccine promotes promiscuity. To capture the complexity of this topic, we broadened our coding scheme to include any discussion of sexual activity and HPV vaccination and were able to identify a wider range of subjects, including safe sex practices, personal stories, and debates over sexual risk.

The LSA analysis showed that cancer is significantly more discussed than sexual activity. Discussions about sexuality and cancer did not typically co-occur: The cosine similarity score between the cancer risk and sexual activity keyword lists in the semantic

space was −0.01, indicating that the 2 lists of keywords (cancer vs. sexual activity) were essentially orthogonal. Discussions about the HPV vaccine were more similar to the keyword "cancer" than to the sexual purity keywords. However, the average monthly Reddit message discussed neither cancer nor sexuality (cosine similarity values were 0.11 and 0.08, respectively). These low similarity scores indicate that other topics must exist in the corpus that is neither about cancer nor sexual behavior. Furthermore, we observed a decrease in average similarity to both cancer and sexuality over time, indicating that other topics of discourse may have emerged within these forums.

In general, we did not notice a significant difference between genders in discussions of cancer and sexual activity.

3.4.2 What other topics characterize the discussion on HPV vaccination on Reddit?

In our close reading of messages for manual annotation, we recognized that although messages included keywords related to cancer risk or sexual activity, there were additional topics that extended beyond the scope of either topic (Table 3-2). Some of these conversations were narrowly focused on HPV vaccination, for instance, debates over whether vaccinating the entire population was a cost-effective prevention method or discussions of vaccine efficacy related to particular strains of HPV. Other conversations were focused on sexual activity and health, in general. We observed a subset of conversations that detailed personal experiences with STIs and described the various steps individuals had taken to avoid all types of STIs (e.g., condom use, HPV vaccination, and regular STI testing). Another major topic was related to circumcision, as some users felt that primary prevention measures such as HPV vaccination reduced the need to circumcise

(as circumcision is often cited as a method to reduce STI infection). Finally, another subset of conversations was focused on broader political and philosophical discussions that sometimes used HPV vaccination as a talking point. For instance, political discussions of the role of government sometimes questioned vaccine mandates as infringing on parental rights as part of a larger justification for libertarianism. Alternatively, users would discuss sexual education, access to contraceptives, and HPV vaccination as talking points in discussions of abortion rights (Table 3-2). Although these conversations were not specifically focused on HPV vaccination, we believe they reflect the broader discourse on the topic and were an essential area of study.

Table 3-2 Summary of human papillomavirus (HPV) vaccine topics generated by the Latent Dirichlet Allocation (LDA) approach

Topic name	Topic keywords	Description
General vaccine debate (Topic 0)	peopl, immun, diseas, get, like, flu, one, hpv, caus, children	Broadly concerned with vaccine decision making; Vaccine safety; Risks of side effects; Importance of vaccination; Personal stories
HPV[a]-related cancer and genital warts (Topic 1)	hpv, strain, cancer, caus, wart, protect, type, infect, genit, genit-wart	HPV strains related to cancer and genital warts; Strains covered by vaccine
Sharing government links (Topic 2)	http, www, http-www, gov, cdc, cdc-gov, nih, www-cdc, nih-gov, nlm	Vaccine concerns corroborated by links to CDC[b] and PubMed research articles
HPV cancer prevention (Topic 3)	cancer, cervic, cervic-cancer, hpv, get, women, prevent, caus, hpv-vaccin, men	Gender-related vaccine and HPV concerns; cervical cancer; other HPV related cancers; recommendations for men and women
HPV-specific vaccine debate (Topic 4)	effect, gardasil, studi, report, gt, hpv, death, year, side, hpv-vaccin	Adverse events linked to HPV vaccination; Side effect risk; Vaccine efficacy; Debating facts (fewer personal stories than topic 0).
HPV vaccine political debate (Topic 5)	peopl, hpv, make, gt, would, like, hpv-vaccin, think, one thing	Parental vaccine rights; Reach of federal power; Abortion debate; Sexual education policy; General political debate; Individual politicians; Corruption
Sharing general links on HPV (Topic 6)	http, com, www, http-www, org, amp, en, hpv, comment, wikipedia	HPV concerns with links to news stories; Links to other Reddit threads; Generally sharing single links rather than a list of links (Topic 2)

HPV vaccination schedule (Topic 7)	get, hpv, hpv-vaccin, year, got, doctor, go, shot, age, girl	Age of vaccination; Preventive screenings; Medical advice; Personal stories; Deciding whether to vaccinate
Circumcision debate (Topic 8)	circumcis, men, gt, risk, women, male, hiv, infect, hpv, benefit	Risks and benefits of circumcision; Likelihood of contracting STI[c]; Cancer prevention
Sexual behaviors including HPV (Topic 9)	get, hpv, sex, test, condom, peopl, like, know, partner, std	Discussions of sexual behaviors; STI prevention; Vaccination; Condom use; Relationship norms

[a]HPV: human papillomavirus. [b]CDC: Centers for Disease Control and Prevention. [c]STI: sexually transmitted infection.

Expanding upon the LSA findings, the LDA topics generated a more comprehensive list of all HPV vaccine–related topics of interest. The most widely discussed topic was HPV-related political debate (17.2%), wherein users posted messages on political topics including vaccine policy, abortion rights, and sex education policy. Cancer was the second most widely discussed (16.1%), after combining the two closely related topics of HPV strains causing cancer and genital warts and HPV cancer prevention. This was followed by sexual activity and preventive behaviors (14.5%), which focused on behaviors such as using condoms, receiving HPV vaccination, and regular testing and screenings for STIs. The fourth most popular topic was HPV disease and immunity (13.5%), which included broad concerns with diseases and infections associated with HPV. The fifth was the vaccine schedule related to HPV (11.5%), which included messages on age of vaccination initiation and requesting information on vaccination from physicians. The sixth, vaccine side effects and risks (9.7%) included messages on adverse events linked to HPV vaccination. The seventh topic was related to messages that included use of government or research (eg, CDC and PubMed) papers and general websites (eg, Wikipedia and news sites) to corroborate user opinions (9.1%). The final topic discussed circumcision, particularly the risks and benefits as related to STI and cancer prevention (8.5%)

3.5 Key Findings of HPV Project and their Implications to this dissertation

LSA findings show that HPV vaccine discussions are significantly more related to cancer (300 out of 400) compared with sexual behavior (199 out of 400) from 2008 to 2015 ($P<.001$). We did not find a significant difference between genders in discussions of cancer and sexual activity ($P>.05$). Our findings demonstrated that women and men both discussed HPV, highlighting that Reddit users do not perceive HPV as an issue that only pertains to women.

Furthermore, LDA analyses demonstrated that although topics related to cancer risk and sexual activity were both frequently discussed (16.1% and 14.5% of word tokens, respectively), the majority of online discussions featured other topics. The most frequently discussed topic was politics associated with the vaccine (17.2%). Other topics included HPV disease and/or immunity (13.5%), the HPV vaccine schedule (11.5%), HPV vaccine side effects (9.7%), hyperlinks to outside sources (9.1%), and the risks and benefits of HPV vaccination (8.5%).

The fact that our analysis discovered several key items weighted almost as crucial as cancer risk and sexual activity had rejected the previous hypotheses that women's rejection of the HPV vaccine is out of the fear of promiscuity. The main reason we are able to make that discovery is that we are able to identify the gender and age information of Reddit users. Moreover, we incorporate the gender and age classifiers in qualitative annotations and topic modeling.

Chapter 4: Vaccine Hesitancy in China after 2018 Changchun Changsheng Vaccine Incident using Sina Weibo data

The following sub-chapter is a summary of our paper published in 2020 regarding vaccine hesitancy in China following the 2018 Changchun Changsheng Vaccine Incident (Hu et al., 2020). The content is moderately abridged compared to the published version. The qualitative codebook is presented in Appendix C.

4.1 Introduction of the Sina Weibo Project

On July 15, 2018, several Chinese newspapers reported that government inspectors had determined that Changchun Changsheng Biotechnology, a prominent manufacturer of vaccines in China, had violated national regulations and standards when producing 250,000 rabies vaccine doses (Tan, 2018). Later that day, this incident was confirmed by The China Food and Drug Administration. The violation might have significantly undermined the effectiveness of the involved rabies vaccines. However, no one reported any adverse reactions or deaths from the involved vaccines before they were recalled in July 2018 (Hong & Wei, 2018). News of this incident slowly escalated on Chinese social media platforms. On July 21, 2018, one blog article critiquing Changsheng Biotechnology, "The King of Vaccines," became the most popular article of the day. On July 23, 2018, the article peaked with 2 million shares, in which the author directly referred to the vaccine incident as a major scandal. Meanwhile, Chinese people voiced their concerns on traditional media, advertisements, and patriotic protests during the incident. The Chinese government acted quickly after the incident (Hong & Wei, 2018), leading Changsheng Biotechnology to declare bankruptcy on June 27, 2019.

4.2 Key Literature Foundation of Sina Weibo Project

Vaccine acceptance is a crucial public health issue (Dubé, Vivion, & MacDonald, 2015), which has been exacerbated by the use of social media to spread content opposing vaccination (Betsch et al., 2012). However, data from social media have also provided new insights into the drivers of vaccine acceptance. Beyond supplementing traditional measures of vaccine hesitancy (Dredze, Broniatowski, Smith, & Hilyard, 2015), studies using data from Twitter (e.g., Broniatowski et al., 2018, who found evidence that state-sponsored and pecuniary interests amplify online vaccine debates to promote discord and distribute spam and malware), Facebook (e.g., Orr, Baram-Tsabari, & Landsman, 2016, who found that polio vaccine refusal was associated with religious rationales), and Reddit (e.g., Lama, Hu, Jamison, Quinn, & Broniatowski, 2019, who found that misconceptions regarding HPV prophylaxis, but not sexual promiscuity, were spread in online communities) have shown that social media can provide new information regarding the dynamics of vaccine communication online, potentially affecting real-world vaccine uptake behaviors. China has the world's largest population, with several cities among the world's top population densities. Thus, disease outbreaks in China have the potential to affect billions of people. However, little is known of vaccine acceptance in China on the national level. The vast majority of the Chinese population is vaccinated: Cao et al. (2012) conducted a national level survey in China (4,681 children from 32 counties) of the coverage of the Expanded Program on Immunization (EPI) vaccines. This study found that primary immunization coverage of original EPI vaccines was overall above 90％, and the coverage of new vaccines integrated into EPI after 2008 was overall above 85%. However, such studies do not directly measure public attitudes towards vaccination.

Social media may be used to measure public opinion regarding vaccination in a manner that is both fast and relatively uncorrupted by response bias due to the self-reported nature of social media messages (Dredze, Broniatowski, Smith, & Hilyard, 2016). However, few studies have been carried out examining Chinese social media due to practical and linguistic barriers: Most people residing in mainland China cannot access international social media without a Virtual Private Network (VPN) service. As a result, most Chinese citizens rely on China-based social media. Sina Weibo is well-known as the Chinese version of Twitter, although it has more users than Twitter does (Radcliffe, 2017).

Unlike the Reddit HPV project, we do not have a handful of literature with detailed hypotheses or competing arguments in the Sina Weibo Project. The existing literature had only suggested that Chinese people have generally accepted mandatory vaccines policy. Therefore, at the time, the project was intended to explore the Weibo data to explore whether Chinese people's attitudes towards vaccines had shifted due to the 2018 Changchun Changsheng Incident.

4.3 Methods of Sina Weibo Project

Discussion regarding the 2018 vaccine incident primarily unfolded on Weibo. Thus, we conducted a retrospective observational study using a convenience sample as our primary data source from this social media. We crawled public messages on Weibo.com with the keyword 疫苗 (Chinese for "vaccine") from July 15, 2018 to July 31, 2018, yielding 22 million messages. We downloaded 11,085 of these messages using Weibo's API. We performed a similar search in August, 2019, yielding 10.5 million messages. Due to changes in Weibo's API, we manually collected another 500 messages in 2019 to assess

the long-term effect of the scandal.

Messages were categorized into three time periods:

(1) Pre-Incident Period, 7/1/2018-7/10/2018. During this period, the incident was still under investigation, and had not yet been reported to the public. We assume that data in this period is representative the vaccine-related discussion on Weibo immediately prior to the incident. 2,721 messages were collected in total for this period.

(2) Escalating Period, 7/11/2018-7/20/2018. During this time period, an investigation was carried out and reported by several minor national newspapers. At this time, more and more people became aware of the severity of the vaccine incident. 3,808 messages were collected in total for this period.

(3) National Headline Period, 7/21/2018-7/31/2018. In this period, the incident became one of the most controversial topics in China. 4557 messages were collected in total for this period.

The following table (Table 4-1) has summarized the time-period distribution of the collected messages.

Table 4-1 Time-period distribution of 2018 data

	Total	Study 1, First Analysis	Study 1, Secondary Analysis
July 1 - July 10	2721	111 (4%)	168 (6%)
July 11- July 20	3808	174 (5%)	155 (4%)
July 21- July 31	4557	216 (5%)	177 (4%)

4.3.1 Study 1: Qualitative Analysis of Messages During the 2018 Vaccine

To answer our first research question, we assigned 500 randomly selected Weibo messages into five categories using qualitative annotation: (1) Animals/Pet, messages

referring to pet and animal vaccines; (2) Scandal/Incident, messages about the vaccine scandal; (3) Harm/Injury, reports on adverse vaccine reactions; (4) Trust, messages that debate societal or moral issues, or expresses frustration that faith in the government might have been misplaced; (5) Self-Reporting Vaccination Experience, explicit self-reports about vaccination or intention to vaccinate. The key difference between categories 3 and 5 is that the messages from category 3 report some forms of adverse reaction after receiving a vaccine and messages from category 5 do not. Messages that did not correspond to any of these categories were marked Unsure, with specific notes. This qualitative codebook was prepared and reviewed by all four authors.

Two authors, both of whom read Chinese fluently, categorized all 500 messages independently using the codebook. When messages were categorized as Unsure one of the paper's authors made notes describing this annotation. We also carried out a second analysis of another 500 messages in 2018 to verify several unexpected observations found in the first round of annotation.

4.3.2 Study 2: Qualitative Analysis of Messages One Year After the Scandal

To assess the long-term effects of the scandal, we carried out the same search from August 1, 2019 to August 10, 2019, yielding another 500 messages. Before we began this collection, Weibo's user agreement changed to require log-in credentials for the aforementioned web-based API. The new policy made it impossible for us to access the same amount of data in 2019 as in 2018. Therefore, we only collected 500 messages by manually downloading 25 web pages generated by the API. The same annotators categorized all 500 messages independently using the same codebook for Study 1.

4.3.3 Study 3: Topic Model Analysis

In order to verify that our qualitative coding results were representative of the full set of 11,585 messages extracted, we once again used Latent Dirichlet Allocation to segment all Weibo messages into 20 probabilistic "topics automatically". Specifically, we used unigram and bigram features with hyperparameter values of alpha=0.1 and η=0.01. When applied to a corpus of messages, LDA returns the probability that each message is about each of these topics. Although LDA output requires interpretation by humans, the result can be used to provide an overall description of the whole corpus of messages.

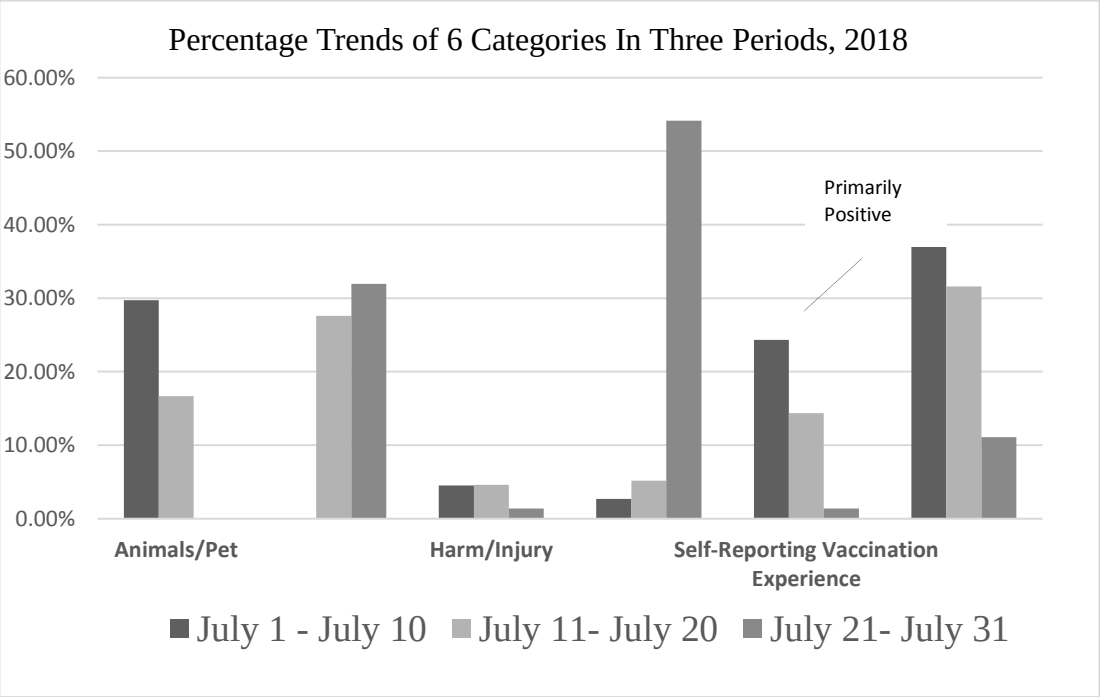
4.4 Results of Sina Weibo Project

4.4.1 Result of Study 1: 2018 Data

The two annotators independently reached the same labels in all but three cases out of 500 with high reliability (Cohen's Kappa = 0.98). The result of the first analysis is shown in Figure 4-1. There were significant differences between the first- and second periods of our analysis, χ^2 (5)=41.60, p<0.001, and between the second and third periods, χ^2 2 (5)=224.86, p<0.001. Figure 4-1 shows that, before the incident, 24% of messages described self-reported vaccination. However, this category dropped to 14% during the Escalating Period, and it eventually became 1% in the third period. In contrast, before the incident, very few people on Weibo had discussed trust in government in the context of vaccines, but this number reached a high of 54% by the third period. Messages not relevant to the incident, but expressing concerns of vaccine harm or injury were only observed five times during the Pre-Incident Period; however, after the incident, twelve messages that were mainly about trust also explicitly used the word "harm people", "harm baby" or "harmful". Finally, messages discussing animal/pet-related vaccines accounted for 30% of

all messages during the Pre-Incident Period but eventually vanished to 0% by the third period.

Figure 4-1. Trends of 6 categories in July 2018



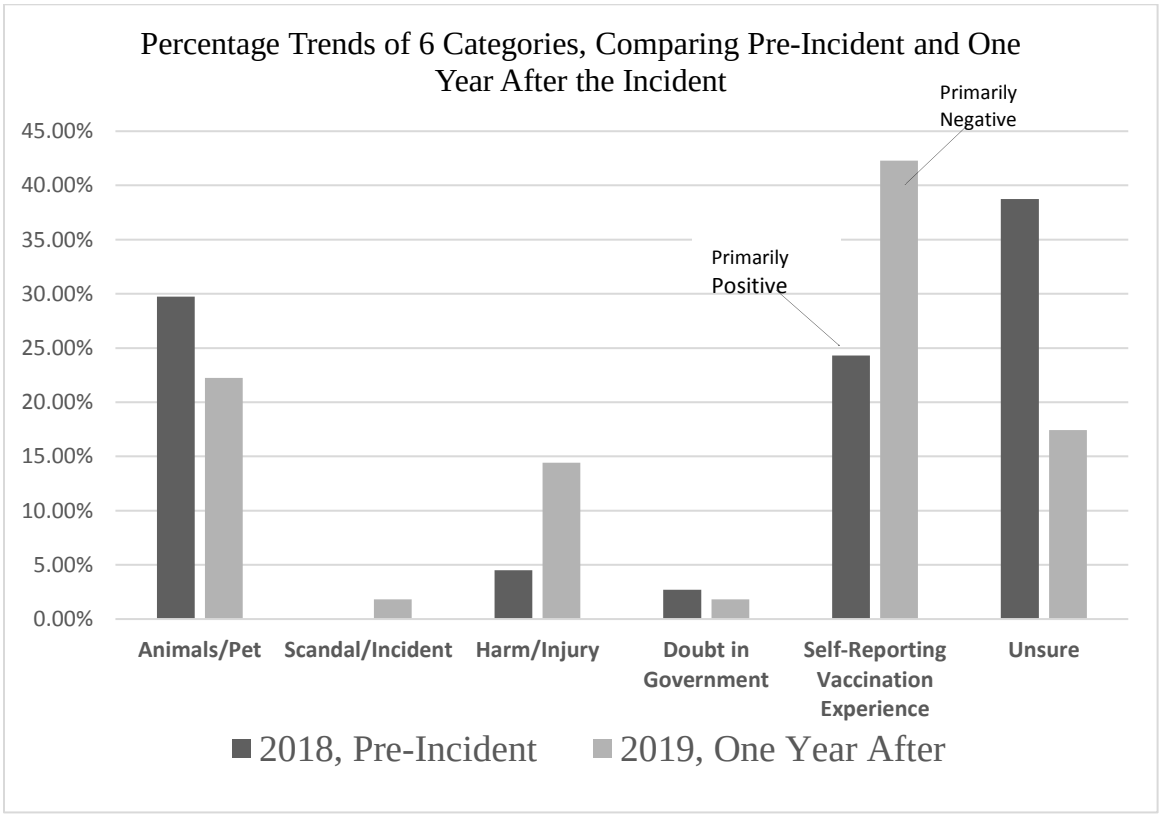
Among the 122 messages categorized as "unsure", we found that 85 messages contained pro-vaccine content, including advertisements for vaccines by local clinics, advertisements for non-mandatory vaccines, or educational material for the public about the vaccine schedule for newborns. Out of these 85 pro-vaccine messages, 14 of them promoted vaccination trips to Hong Kong, or identified themselves as vaccine providers located in Hong-Kong, where the healthcare standards are considered higher by most Chinese people.

4.4.2 Result of Study 2: 2019 data

The two annotators' agreement was moderate (Cohen's Kappa = 0.53). The disagreement resulted from two types of messages: one author interpreted ambiguous messages using the word "baby as referring to humans, whereas the other interpreted these

messages as referring to pets. The second source of disagreement came from discrepancies as to whether some self-reports should be categorized as harm/injury because of their negative sentiment about vaccines. We did not categorize a message as expressing vaccine harms unless it did so explicitly. These results (Figure 4-2) show that the proportion of messages reporting adverse reactions to vaccines or expressing concerns of potential vaccine harm jumped from 4.5% to 14.42% in the past year. Compared to the pre-incident period in Study 1, there was a significant change, $\chi^2(1) = 5.795$, p=0.016 in how messages pertaining to vaccination were distributed, with a marked increase in the proportion of messages self-reporting vaccination. Furthermore, the tone of these messages changed from primarily positive to primarily negative. However, the scandal itself was rarely discussed.

Figure 4-2. Shift of Distribution of 6 categories, Pre-incident Data vs. 2019 Data



4.4.3 Result of Study 3: Results Replicate Across Entire Corpus using LDA

After we fit a 20-topic model to the 11,515 messages in our corpus using LDA and manually aggregated the output into 6 overarching topics based on the most representative words in each topic. Table 4-2 displays the top 5 tokens (either unigram or bigram) of each topic in Chinese, their English translations, and to which of the 6 aggregate topics it was assigned to.

Table 4-2. LDA model summary of Weibo Project: Assigned Category, Top Keyword in Chinese, and their English Translation

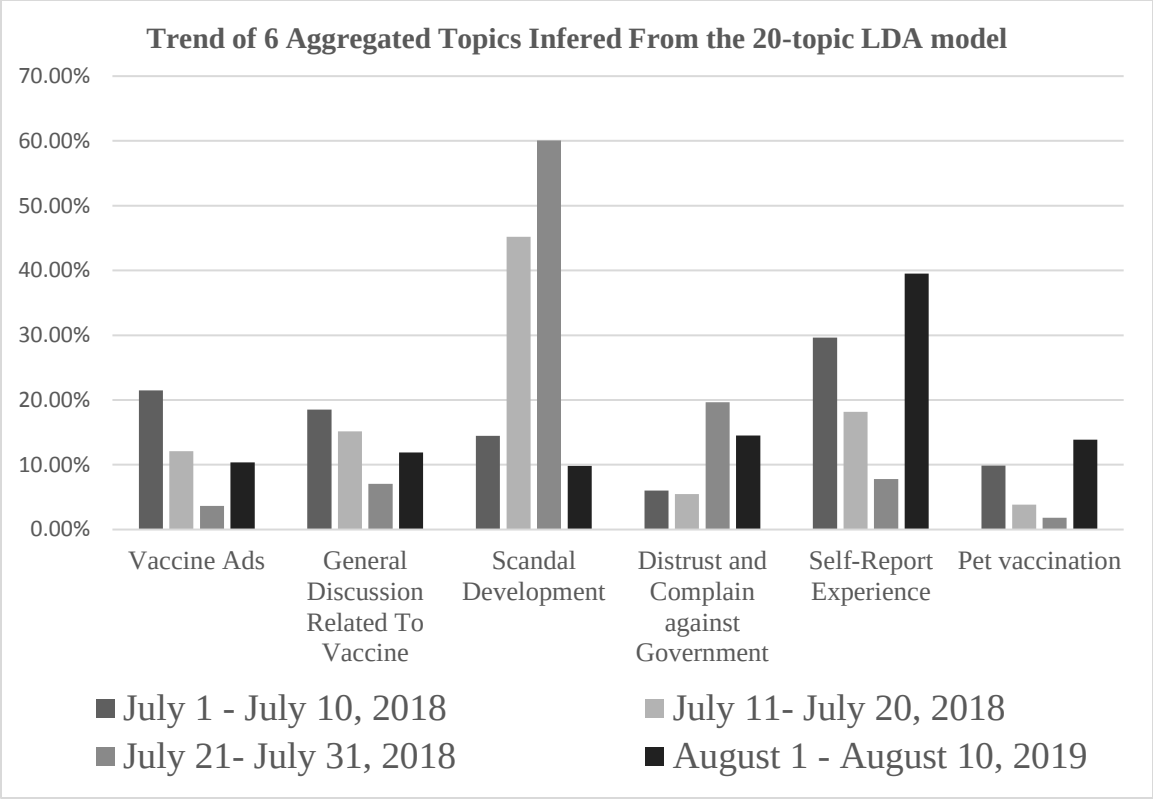
ID	Category	Top 5 Chinese Words					English Translation				
0	Vaccine Ads	hpv	疫苗	香港	hpv-疫苗	宫颈癌	Hpv	Vaccine	Hongkong	Hpv-vaccine	Cervical Cancer
1	Scandal Development	生产	疫苗	违法	长生	狂犬病	Production	Vaccine	Illegal	Changsheng	Rabies
2	General News/Education/Discussion Related To Vaccine	疫苗	接种	展开	展开-全文	全文	Vaccine	Inoculation	Expand	Expand-Full Article	Full Article
3	Scandal Development	疫苗	武汉	百白破	生物	武汉-生物	Vaccine	Wuhan	DPT	Biology	Wuhan-Bioloy
4	Self-Report Experience	疫苗	孩子	注射	注射-疫苗	感冒	Vaccine	Children	Injection	Inject-Vaccine	Flu
5	Scandal Development	事件	疫苗	疫苗-事件	关注	粉丝	Incident	Vaccine	Vaccine-incident	Follow	Fan
6	Scandal Development	cn	http	http-cn	疫苗	新闻	cn	http	http-cn	Vaccine	News
7	Pet vaccination	领养	疫苗	xe627	驱虫	展开	Adoption	Vaccine	Unicode xe627 (V)	Worm treatment	Expand

8	General News/Education/Discussion Related To Vaccine	伤口	狂犬	狂犬-疫苗	冲洗	疫苗	Wound	Rabies	Rabies-Vaccine	Wash	Vaccine
9	Scandal Development	生产	疫苗	长生	狂犬病	疫苗-生产	Production	Vaccine	Changsheng	Rabies	Vaccine-Production
10	Scandal Development	生物	长生	长生-生物	疫苗	造假	Biology	Changsheng	Changsheng-Biology	Vaccine	Fraud
11	Self-Report Experience	疫苗	今天	没有	自己	时候	Vaccine	Today	No	Self	Time
12	Scandal Development	长春	长生	长春-长生	疫苗	狂犬病	Changchun	Changshen	Changchun-Changshen	Vaccine	Rabies
13	Distrust and Complain against Government	疫苗	我们	没有	什么	一个	Vaccine	We	Don't have	What	One
14	Scandal Development	http	cn	http-cn	疫苗	狂犬	http	cn	http-cn	Vaccine	Rabies
15	Distrust and Complain against Government	疫苗	问题	中国	监管	企业	Vaccine	Problem	China	Regulation	Cooperation
16	Vaccine Ads	围观	免费	免费-围观	一起	问题	Broadcasting	Free	Free-Broadcasting	Together	Problem
17	Self-Report Experience	疫苗	出现	宝宝	可以	大家	Vaccine	Appear	Baby	Could	Everyone
18	Self-Report Experience	宝宝	疫苗	接种	检查	观察	Baby	Vaccine	Inoculation	Check	Observe

19	General News/Education/Discussion Related To Vaccine	疫苗	接种	接种-疫苗	有效	注射	Vaccine	Inoculation	Inoculation-Vaccine	Effective	Injection

We determined the final 6 topics to be: (1) Vaccine Ads, (2) General News/Education/Discussion Related To Vaccine, (3) Scandal Development, (4) Distrust and Complaints against Government, (5) Self-Report Experience, and (6) Pet vaccination. Figure 4-3 demonstrates how the distribution of these 6 topics developed overtime. Four of our main findings from prior sections are in qualitative agreement with the LDA model's output: First, we can see the rapid development of the scandal in late July, 2018. Second, pro-vaccine advertisements gradually decreased as the incident developed, and in 2019 it remained lower than the pre-incident level of 2018. Third, in 2019, people on Weibo were more likely to self-report vaccination (and potentially, vaccine harms) compared to the pre-incident level of 2018. Finally, we found that in 2019, people were more likely to mention the Chinese government when mentioning vaccines: a rarity before the incident.

Figure 4-3. 6 Aggregated Topics from LDA Model



4.5 Key Findings of Weibo Project and their Implications to this dissertation

Our results demonstrate a shift in the vaccine discourse on Weibo. Immediately after the incident, the frequency of self-reported vaccination declined significantly, while the discussion of trust in government increased. One year later, discussions of trust in government are no longer present and self-report rates increased significantly. However, the nature of these self-reports changed. Before the incident, self-reports were primarily positive in nature, whereas one year after the incident, vaccinators expressed more concerns. Furthermore, a higher proportion of messages reported harms from vaccination.

Our results indicate that vaccination experience has become a more conspicuous topic among Weibo users, with many actively seeking information on this issue. This may suggest the erosion of vaccine confidence in Chinese society, since prior research has shown that the very presence of a vaccine debate can lead to more vaccine hesitancy (Dixon

& Clarke, 2013).

In this project, our main contribution to the scientific community is that we are able to collect the 2018 data promptly as the scandal was developing despite the limitation of the Weibo API. And more importantly, both LDA and annotations have agreed on overall trends of key categories.

The fact that we are able to discover the shift of the vaccine sentiment in China using social media data with LDA and sample message annotation without a full survey is worth mentioning. Our findings were replicated and confirmed by at least four studies using survey methods (Du et al., 2020; He et al., 2020; Sun et al., 2021; Xu & Liu, 2021). All of them have cited our work.

This project first indicates that the prior hypothesis that vaccine hesitancy in China is no longer solid. And this project also indicates that the combination of LDA and qualitative annotation of social media messages can generate insightful hypotheses that can be tested by survey studies.

Chapter 5: Study "Yelp Effect" using Yelp and Google Maps data

The following chapter is a summary of our paper published in 2021 regarding patient-doctor communication on the issue of antibiotics usage (Hu et al., 2021). The literature review part of the paper is only slightly abridged compared to the published version. However, the visualization part and the details of data crawling of the paper are mostly abridged compared to the published version. The detail of the data collection process can be found in Appendix D.

5.1 Introduction of the "Yelp Effect" Project and Literature Review

Overuse of antibiotics is the primary factor leading to the development of antibiotic resistance, deemed by the World Health Organization (WHO, 2015) to be one of the world's most pressing health problems.

Approximately one-third of all antibiotics prescribed in the United States outpatient sector are for infections for which antibiotics are not indicated (Fleming-Dutra, JAMA, 2016). One study (Palms et al. 2018) found that in urgent care centers, antibiotics were prescribed for almost half (46%) of acute respiratory tract infections for which no antibiotics are indicated.

We aim to understand the precursors to overprescribe antibiotics in the urgent care setting. Providers often prescribe because they assume that patients desire antibiotics and will be unsatisfied if antibiotics are not prescribed (Mangione-Smith 1999; May 2014). However, physicians' assumptions about patient expectations frequently do not match actual patient expectations (Stivers 2003). Furthermore, prior literature on patient satisfaction and doctor-patient communication suggests that there are other factors at play

(Pew 2017). Welschen et al. (2004) found that receiving information or reassurance was more strongly associated with satisfaction than receiving an antibiotic prescription in primary care. Ong et al. (2007) found that patient satisfaction was not related to receipt of antibiotics but was related to the belief they had a better understanding of their illness. Stearns et al. (2009) found that patients at Veterans Administration hospitals generally had equal levels of visit satisfaction regardless of antibiotic treatment status.

Provider surveys querying the reasons for antibiotic overprescribing have identified fear of negative patient reviews and a perception of an association between antibiotic prescribing and patient satisfaction as drivers for inappropriate prescriptions (Zgierska 2014).

Patients often use online reviews when deciding whether and where to seek treatment (Zhang, Zhao, Cheung, & Lee, 2014). However, these reviews may not accurately reflect quality medical care. For example, Daskivich et al. (2018) found that data from 5 popular online platforms (Healthgrades, Vitals, Yelp, RateMDs, and UCompareHealth) showed no significant association between the scores rated by online consumers and scores generated by professional peer-review systems. Despite this lack of association, Yaraghi et al. (2018) found that when consumers choose primary care physicians, they tend to perceive nonclinical ratings on commercial reviews websites to be as important as the ratings from government websites. Thus, to the extent that online reviews drive provider behavior, they may do so in a manner that undermines public health.

Authors in several media outlets have coined the term “Yelp Effect,” proposing that urgent care and other providers of on-demand (e.g., walk-in visit) care prescribe medically unnecessary antibiotics to prevent receiving negative online reviews (Mckenna, 2016;

Lagasse, 2019). There is a perception by urgent care providers that many of their patients expect antibiotics, even when they are clinically unnecessary, and will leave negative reviews online if these expectations are not met. To avoid negative online reviews, which could impact provider pay and the performance of urgent care centers, providers are driven to prescribe antibiotics even when clinically unnecessary.

The "Yelp Effect" is reportedly magnified by the surging popularity of online reviews sites like Google reviews, Yelp, and Healthgrades (Mckenna, 2018). Despite prior studies analyzing online physician ratings show that the vast majority are favorable (Kadry 2011; Lopez 2011), at least one study has demonstrated that online physician reviews containing the words “medication” and “prescription” are mentioned in more negative contexts (Paul, Wallace, & Dredze, 2013). This perception of a connection between negative patient online reviews and antibiotics is hypothesized to be a significant driver of inappropriate antibiotic prescribing, which is evident in at least one petition that has received 40,000 signatures supporting the removal of online doctor reviews (Physicians Working Together, 2017).

To our knowledge, no previous studies have evaluated online reviews to assess study antibiotic antibiotics prescribing specifically. Yet, prior work has primarily focused on the value of mining online reviews to discover novel topics. For example, Cawkwell et al. (2015) used Yelp to track 137 hookah bars in New York State and found that Yelp can be used as a powerful tool when investigating sensitive habits. Similarly, Ranard et al. (2016) collected 18,058 reviews on Yelp to study patients’ assessments towards hospitals, and they found that Yelp reviews can often provide complementary topics that are not covered by survey items. This prior work emphasizes the value of analyzing online review data for understanding how physicians’ perceptions of the public’s concerns may diverge from the

reality of what is being expressed in online reviews.

We seek to determine whether there truly is a "Yelp Effect" by evaluating how frequently patients leave negative online reviews complaining that they did not receive antibiotics when they were told it was medically unnecessary.

To determine the prevalence of the "Yelp Effect", we analyzed a large sample of online reviews of urgent care centers in the United States, calculating the proportion of these reviews that relate to antibiotic prescription, and especially to no receipt of antibiotics when patients believed they were expected.

Specifically, we sought to (1) quantify what proportion of negative reviews were posted due to patients (reviewers) not receiving an antibiotic and (2) evaluate the content of reviews not related to antibiotic prescription.

5.2 Methods of "Yelp Effect" Project

5.2.1 Data Collection

We used the Yelp application programming interface (API) between October 1 and October 20, 2018, to obtain a list of facilities that were tagged as "urgent care" in the United States. We also retrieved the star ratings for all reviews that were associated with these facilities. We removed all facilities that did not have a US zip code. We were unable to obtain the full texts of these reviews—Yelp's terms of service prohibit web scraping, and Yelp declined to provide us with permission to use the text. Thus, by using the list of facilities that was obtained from the Yelp API, we ran the data collection algorithm from January 21 and February 10, 2019, and obtained all of the Google Maps reviews of each urgent care facility that were posted before January 21, 2019. Unlike Yelp, Google's terms

of service permit the collection of their reviews as long as it does not put undue burdens on Google's servers. To collect data from Google Maps reviews, we designed a novel computer program for automatically collecting these reviews based on their Google Maps URLs. The Google Maps URLs of the urgent care facilities were collected after we presented their titles and addresses (obtained from the Yelp API) to 20 workers on Amazon Mechanical Turk, a crowdsourcing marketplace. The detail of the data collection process can be found in Appendix D.

These 20 workers were selected based on their prior demonstrated ability to successfully collect a set of 100 known URLs. Workers received the US $0.09 for each URL collected. Figure 5-1 to 5-4 are a list maps demonstrating the geographical distribution of the aforementioned data.

Figure 5-1. Total number of reviews in US counties (Yelp data)

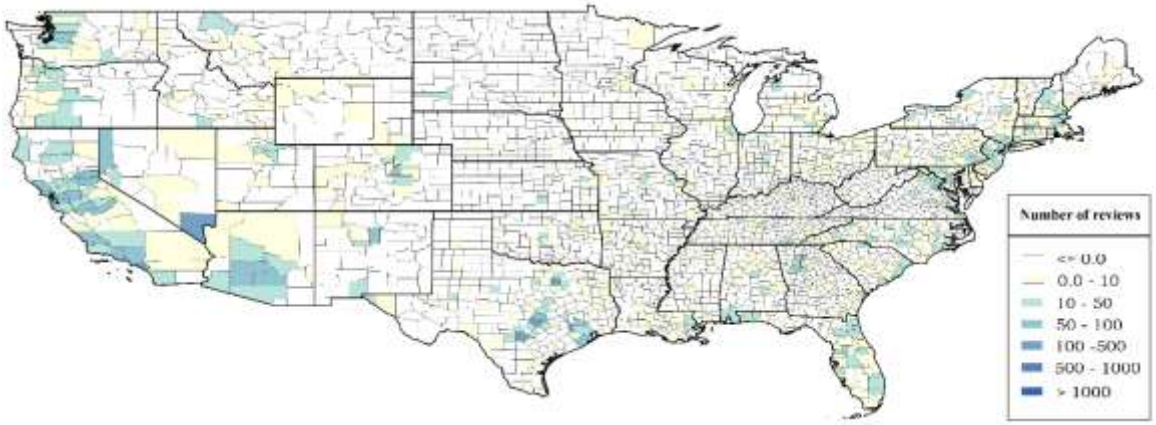


Figure 5-2. Mean review stars in US counties (Yelp data)

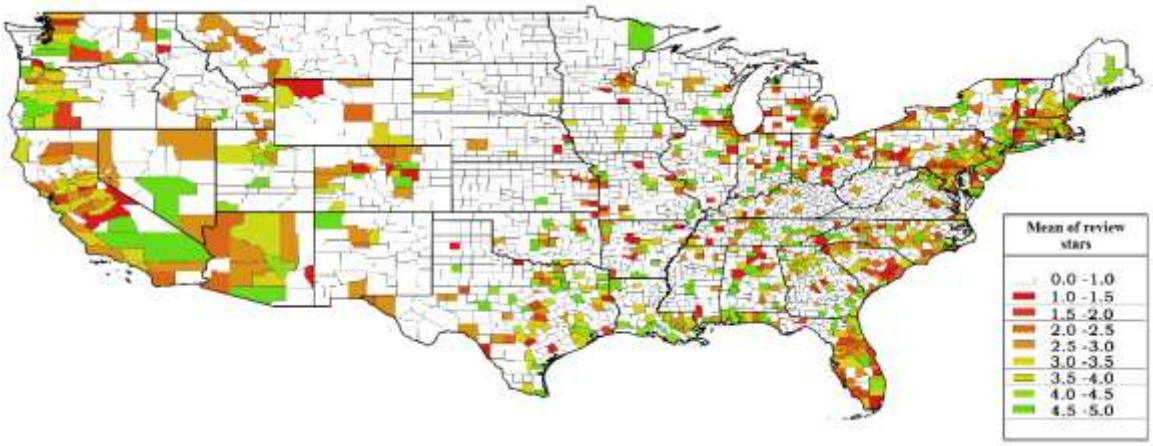


Figure 5-3. Total number of reviews in US counties (Google Maps data)

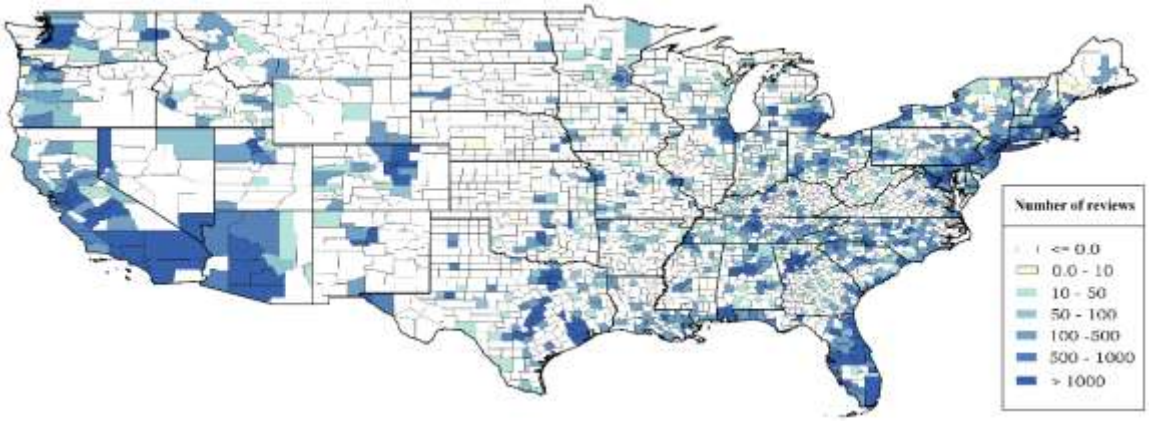
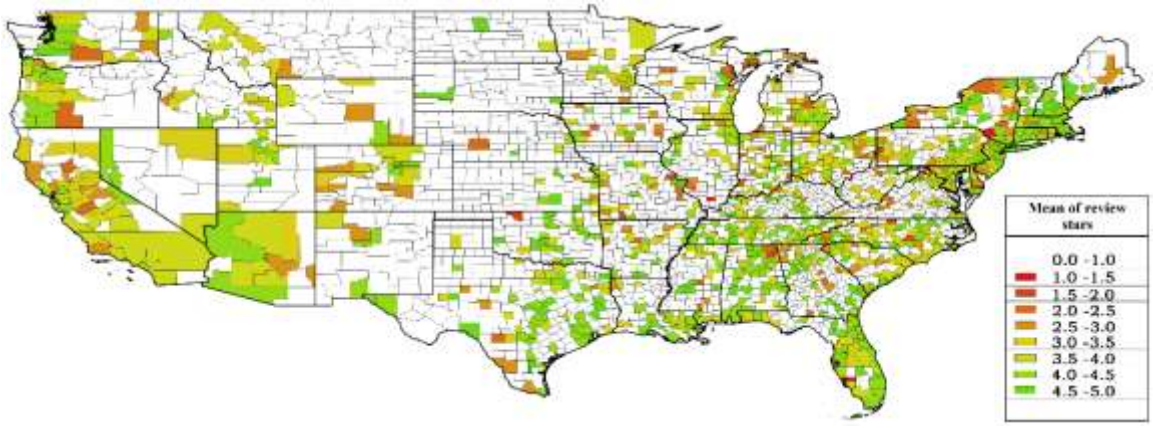


Figure 5-4. Mean review stars in US counties (Google Maps data)



5.2.2 Data Analysis using LDA

To summarize the contents of web-based reviews, we once again used the latent

Dirichlet allocation (LDA) topic model (Blei & Lafferty, 2006). This algorithm was implemented in the LDA python package (default settings were used) to fit the model to the Google Maps data set. For each topic, we calculated the total number of word tokens that were found in positive reviews (4 or 5 stars) and negative reviews (1 or 2 stars). We fitted 3 topic models (one model with 10 topics, another with 20 topics, and another with 50 topics) to the data. We selected the model that generated the most coherent topics without a large increase in perplexity (a measure of model goodness of fit that is commonly used in natural language processing) (Blei, 2012). Afterward, we extracted antibiotic-related reviews by using a list of keywords that was generated by one of the authors, a pediatrician who specializes in antibiotic stewardship.

For each review, we examined the proportion of words in each topic. We applied the same procedure to the subset of reviews containing antibiotic-related keywords. We then developed a qualitative codebook to determine the content and sentiment of 200 reviews that were sampled at random from all reviews containing these keywords.

5.2.3 Qualitative Annotation

Four authors collectively reviewed a subset of the 200 randomly sampled reviews to determine the types of messages that were present in the reviews (Krippendorff, 2013). After this initial review and the development of inductive codes, the reviews were categorized into one of the following categories in the codebook: (1) the Yelp effect category (the patient wanted antibiotics but did not receive them); (2) the opposite of the Yelp effect category (the patient received antibiotics but did not want them); (3) the convenience, inconvenience, and wait times category; (4) the staff competence or

incompetence, courtesy and attitude, and satisfaction of care category; (5) the cost and price of drugs per visit (including sticker shock) category; (6) the other prescription-related complaints category; and (7) the other or none of the above category. Additionally, all reviews were annotated as "positive" (e.g., the patient was satisfied with their care, and the review had 4 or 5 stars) or "negative" (e.g., the patient was dissatisfied with care, and the review had fewer than 4 stars).

The four annotators then independently reviewed the same 200 randomly sampled reviews to assign them to 1 of the 7 categories. The final categories for the reviews were assigned based on the majority category among annotators. If there was no majority category, disagreements were resolved discursively until a consensus category was agreed upon.

Some reviews mentioned that the reviewer did not receive an antibiotic when it was expected, even if that was not the main message. Thus, after assigning each message a primary category, the same four annotators revisited all 200 reviews to determine if they mentioned the Yelp effect in passing (i.e., whether the review mentioned an unfulfilled expectation for antibiotics). A code for mentioning the presence or absence of the Yelp effect—even if it was mentioned in passing—was then assigned to each review as a secondary code. By using these 200 annotated samples, we inferred population proportions for each category in our codebook and calculated 95% CIs

5.3 Key Results of "Yelp Effect" Project

5.3.1 LDA result

From the Yelp API, 8662 unique urgent care facilities were identified, containing 84127

unique reviews. From Google Maps, 481825 U.S.-based reviews were collected. Among them, 340,328 contains some text. The average star ratings for Yelp reviews were significantly lower than for Google reviews $t(565950) = 82.38$, $p = .0000$.

Of all 481825 reviews with text from Google Maps, we categorize reviews in Google Maps into three groups: (1) Reviews without any text (n=141497, mean=4.53); (2) Reviews with text but not mentioning antibiotics (n=332566, mean=3.94); (3) Reviews with text but not mentioning antibiotics (n=7762, mean=2.40). We found significant differences in the average star ratings across these three groups, $F1, 481824= 12480.31$, $p = .0000$. A post-hoc Tukey HSD test for multiple comparisons shows that reviews with antibiotics keywords had the lowest ratings compared to reviews without text, Mean Difference=-2.13, p=0.001, and reviews with text but not keywords, Mean Difference=-1.54, p=0.001. Reviews without text were significantly more positive than reviews with text but no keywords, Mean Difference=0.59, p=0.001.

The LDA topic modeling analysis (Figure 5-1) yielded 20 topics. The most negative topics pertained to rude staff, wait times, billing, callbacks, and other aspects of customer experience. Although the "infections and symptoms" topic was also predominantly negative, this topic did not make up a plurality of the reviews (481760/6795468, 7.09%). The graphical representation is displayed in Figure 5-2 and the word summary and topic representation is in Table 5-1

Figure 5-5. Topics generated by the latent Dirichlet allocation algorithm based on the text of all Google reviews in our data set. *Note*: The size of each topic is proportional to the number of words that were assigned to each topic by the algorithm. Words are further segmented according to the sentiment of each review. Reviews with 4-5 stars are positive, reviews with 1-2 stars are negative, and reviews with 3 stars are neutral.

Table 5-1. LDA model summary of Google Maps Data

ID	Topic Weight	Topic Interpretation	Tokens
Topic 0	421980	Billing	insur bill pay visit charg would go call get told even place offic paid servic never compani money doctor receiv
Topic 1	373851	Facilties	care urgent urgent-care facil need best go staff care-facil ever visit experi center medic one thank time use receiv profession
Topic 2	329896	Knowledgeable Staff	time question dr took care staff answer profession explain visit medic listen concern answer-question took-time treatment thorough doctor experi patient
Topic 3	339340	Feelings	feel make care made better like thank dr sure comfort staff help manner made-feel realli much feel-like well make-sure bedsid
Topic 4	309548	Children	son us daughter old year took year-old kid child even dr take care children time close sick get brought nurs
Topic 5	325847	Appointments	minut wait 30 10 within appoint seen 15 check min hour time 20 walk 30-minut less arriv room got onlin
Topic 6	297999	Reccomendations	recommend highli would highli-recommend staff would-recommend definit anyon care profession friendli place definit-recommend recommend-place dr would-highli facil would-definit recommend-anyon need
Topic 7	260402	Wait Times	wait time long wait-time staff friendli short long-wait wait-long nice doctor clean short-wait profession good alway littl experi visit staff-friendli
Topic 8	232727	Friendly Staff	friendli staff fast profession staff-friendli clean quick help friendli-staff servic effici friendli-help friendli-profession facil fast-friendli knowledg super nice easi courteou

Topic 9	232129	Positive Reviews	great staff experi great-experi place good care great-staff nice friendli doctor staff-great alway dr great-place thank help great-care good-experi quick
Topic 10	481760	Infections and symptoms	doctor went infect day antibiot go told test ear flu prescrib said get throat symptom back strep time came cough
Topic 11	459841	Call back	call back day get told would test appoint result call-back said ask phone work time prescript offic next doctor need
Topic 12	196527	Quality of Care	servic great excel care custom custom-servic fast great-servic staff thank good profession quick googl excel-care origin provid receiv peopl attent
Topic 13	503598	Rude Staff	go place never doctor like rude time ask would get one said say give peopl want know even ever back
Topic 14	269311	Would go back	back go right need definit nice got help go-back staff away get quickli right-away come friendli seen care everyon would
Topic 15	444898	Emergencies	pain go doctor emerg er room hospit went ray get could would told need emerg-room medic even day care came
Topic 16	256630	Medical Staff	doctor nurs staff nice offic place practition go help nurs-practition love care see nurs-doctor best doctor-offic realli doctor-nurs alway kind
Topic 17	421106	Wait Times	wait hour room time wait-hour wait-room peopl seen doctor back one go come appoint get place see still minut half
Topic 18	273218	Front Desk	front desk front-desk ladi nurs help nice staff rude doctor person peopl work patient girl one check experi assist like
Topic 19	364860	Patient Needs	care dr time alway doctor primari patient year staff medic famili everi take primari-care physician best go clinic need offic

5.3.2 Annotation Results

Four annotators annotated 97 of the first 100 reviews, achieving moderate reliability (Fleiss' k = 0.44). All annotators agreed on sentiment. After disagreements were resolved discursively, the second round of review on the next set of 100 reviews yielded substantial agreement (Fleiss' k = 0.66) with disagreements again resolved by consensus among reviewers. Table 5-2 summarizes the result of these annotations. Out of a total of 200 reviews, we found that only five (2.5%, 95% CI: 0.3%-4.7%) exhibited the "Yelp effect" as the primary message category.

Table 5-2 Primary Message Categories of the 200 Messages in 7 Categories

(1) Yelp Effect – the patient expected antibiotics but didn't receive an antibiotic.	5
(2) Counter Yelp effect – the patient received antibiotics but didn't want them.	1
(3) Convenience/Inconvenience/Wait times (+/-).	36
(4) Staff competence/incompetence, Courtesy/Attitude, Satisfaction of care (+/-).	138
(5) Cost/Price (of drugs/visit) {includes sticker shock} (+/-).	13
(6) Other Complaints (+/-).	3
(7) Other/None of the Above.	4
Total	200

5.4 Key Findings of Yelp Effect Project and their Implications to this dissertation

Our data suggest that the "Yelp Effect" is quite rare. Out of a set of almost half a million reviews, less than 1 in 1000 seemed to contain the Yelp Effect as the primary message with 95% confidence. Furthermore, with 95% confidence, less than 1 in 250 reviews seemed to contain the Yelp Effect, even as the secondary message.

In contrast, we found that 69% of the reviews in our dataset focused primarily on assessments of staff competence and quality of personal interaction. This suggests that, to the extent that a "Yelp Effect" exists, patients express it by questioning the expertise or personal qualities of urgent care staff. This may put urgent care providers in a bind: although they should not prescribe inappropriately, failure to explain to patients why the patients' preferred treatment is ineffective may lead to reviews designed to undermine providers' credibility, expertise, and personal qualities. Thus, it is of paramount importance that providers and urgent care staff both provide high-quality care, and also leave patients with a meaningful understanding for why they received the treatment that they did.

In this data, both our LDA model and the qualitative annotation have confirmed that

only a small fraction of the total reviews has mentioned antibiotics. The estimate from the LDA model suggests it is only 7.09%, while the annotation of the 200 reviews suggests it is only 2.5% (95% CI: 0.3%-4.7%)) of all messages. The results are very similar.

Chapter 6: Systematic Comparison of the key methods and findings of the three projects

6.1 A Systematic Comparison

I designed the following Table 6-1 to summarize the highlights of the three projects

Table 6-1 Comparison of my three published projects during my Ph.D. years

	Reddit HPV Project	**Weibo Project**	**Yelp Effect Project**
Data Source	Reddit Data 2009-2015	Weibo Data, 2018-2019	Yelp and Google Maps, pre-2019
Data Collection Methods	Pushshift API	Weibo API and Manual Crawling	Crowdsourcing and massive crawling
Who/what is the author of the message	Reddit Users/Might Contain Bots Undetected by Community Moderators	Personal Weibo Users/Unlikely Bot due to because of the Cell-phone Registration Policy	Yelp/Google Users, although Malicious Users are Detected
What is the context/vehicle of the message	Text only Comment/Post	Weibo (Tweet) (Broadcasting)	Reviews
Who/what is the audience of the message?	Other Reddit Users browsing the same Forum	Followers/Future Followers	Urgent Care Owner Other customers
The Theoretical Population	All People who speak English	All Chinese People	All People who have posted urgent care reviews on Yelp/Google Maps
The Study Population	All English-Speaking Reddit Users collected Via Pushshift API	Chinese Users on Weibo	All People who have posted urgent care reviews on Yelp/Google Maps
The Actual Sample	All English-Speaking Reddit Users collected Via Pushshift API	Randomly sampled personal Chinese Users who was active on Weibo during the July to July 2018, and August 2019	All People who have posted urgent care reviews on Yelp/Google Maps

Sample Thoroughness	Very Thorough, almost all public data (as described by Pushshift)	Limited in volume, a small fraction compared to all Weibo data at that time	Very Thorough, all available public data pre-2019
LDA Implementation	Gensim LDA	Python LDA package	Python LDA package
LDA evaluation methods	Annotator's judgment of samples messages / Interpretability	Annotator's judgment of samples messages / LDA model Perplexity	Annotator's judgment of the quality samples messages / LDA model Perplexity
Qualitative Annotators	2 (I was not one of the main result annotators, only developed the codebook)	2 (I was the main result annotator, also developed the codebook)	4 (I was not one of the main result annotators, only developed the codebook)
LDA Results vs. Qualitative Annotation Results	**Cancer/Sexual Comparison Ratio:** LDA: 1.11 (16.1% / 14.5%) Annotation: 1.51 (300/199)	**Messages of Vaccine Distrust: pre-incident /one-year-after incident ratio:** LDA: 2.42 (14.52% / 6.0%) Annotation: 3.2 (14.42% / 4.5%)	**Estimate of messages displaying Yelp effect of all messages with antibiotics keyword, which is 2.28% of total messages** LDA: 7.09% Annotation: 2.58%
Samples Representativeness of the small data	We have most of the users covered. Reddit users are young and liberal based on literature; however, we are able to document their gender and age using our classifier.	Totally randomized data collection within the scope of the API, difficulty to assess the degree of bias	We have most of the users covered. Since our target population is people who left online reviews, the selection bias is minimal

The hypothesis suggested by prior literature, which motivates this study in the first place	1. Promiscuity is the main reason why parents do not vaccinate their children with the HPV vaccine (rejected by our work). 2. Women are more likely to develop hesitancy against HPV vaccine (rejected by our work).	1. China has a relatively high vaccine acceptance, and the government has a mandatory vaccine policy (rejected by our work).	1. Many patients left negative reviews online because they did not receive the antibiotics they desired (rejected by our work).
Preliminary findings/new hypothesis generated by the work	1. Promiscuity is not the central concern of parents. 2. Gender difference is minimal for HPV vaccine hesitation 3. Many other factors are at the table	1. Vaccine Hesitancy in China increased significantly after the 2018 incident, which are later confirmed by several surveys	1. Antibiotics are merely one of those controversial topics between doctor-patient communication 2. Most negative reviews left by those patients regarding antibiotics are due to lack of communication, but because they receive it.

Overall, our work has shown that a four-step combination works well: (1) a thorough data collection, (2) a randomized sampling from the big data to generate the small target sample, and (3) a professional and critical qualitative annotation of the small data and (4) a state-of-art LDA implementation that can describe the entire big data.

These four steps are the central framework we followed in all three projects. And all three times, they prove that they can discover new content relevant to traditional literature in public health from social media. More importantly, they can provide insights that can

challenge immature preliminary hypotheses and generate new hypotheses that are worth investigating.

6.2 Limitations

All three studies have focused on the specific platforms, and we have claimed in each paper that we do not intend to generalize these findings to other platforms. Instead, we encourage other researchers to test our findings on the different populations and platforms.

Although some level of selection bias is inevitable due to the inherited nature of social media platforms, we see more and more people using social media across the globe every day. Furthermore, people not using social media have become the minority in most developed countries. Therefore, if scholars can thoroughly collect social media data from cross platforms, the selection bias of social media studies is less terrible than it was ten years ago.

Nevertheless, our work is still limited in several key factors. Several studies have suggested that bots and malicious "users" exist on Twitter (Broniatowski et al., 2018; Cresci et al., 2019; Luceri et al., 2020). Reddit has mainly incorporated a subreddit specific policy which allows moderators of each Subreddit to control access and write "official" bot to remove and ban suspicious users (Jhaver et al., 2019), however, researchers are still able to find bot-like behavior occasionally at some subreddits (Hurtado et al., 2019). Meanwhile, although very few have evaluated the bot/malicious actor on Yelp, or Google Maps, we have encountered one suspicious bot ruring our annotation of the Google Maps data. In our dataset, one user named "Ryan M" posted the one exacted same message ("*NON AFFILIATED / CHARGES DOUBLE: I went in for a routine visit, and all they did*

was took vitals, asked a couple questions and prescribed me an antibiotic. I was charged double for having insurance vs being a ""self-pay"" patient. This facility is also not affiliated with Providence Hospital in which the Medical building they are located inside of. After talking with their billing supervisor, I was told that there was nothing that they could do. I will never visit this location again, nor will I ever recommend their facility to anyone else.") 32 times across different states. And we suspect more might be hidden in the online review systems. Therefore, we cannot rule out the possibility that a considerable portion of the result of these studies is not from genuine users. Secondly, due to the limitation of the annotation budget, we have only applied pre-trained the standard natural language processing model in tokenizing data, which can be a problem in detecting culture community-specific "lingos." Thirdly, although our sampling techniques are relatively unbiased and representative compared to the default API, our data collecting framework and analysis are time-consuming to develop and expensive to deploy. Therefore, our analysis cannot be scaled quickly.

Chapter 7: Conclusions and Future work

Policymakers must act rapidly to grasp the voice of a swiftly-changing society; however, yet even the best algorithms cannot directly provide insights that can be translated into policy. Meanwhile, public health researchers must have the courage to dive into various data sources to explore the understudied population and communities. In this dissertation, we have seen that social media data can be meticulously sampled and studied based on demographic information (Reddit), sensitive time periods (Weibo), and geographic information (Yelp/Google Maps).

The founders of Pushshift API cited our Reddit work in their project introduction paper (Baumgartner et al., 2020), acknowledging that we are among the first groups of scholars using Reddit data in a large volume. Although we can only identify the gender information of 61.8% of Reddit users with 88% accuracy and the age information of 40% of Reddit users with 67% accuracy, this was still the state-of-the-art published demographic classifiers of Reddit users in 2019. During my fellowship at FDA, I have applied an improved version of the demographic classifier and have obtained meaningful results of women on Reddit (paper in preparation). Therefore, I am confident that the procedure and strategy used in our Reddit work can be incorporated into other public health research projects, especially those interested in the difference between the various demographic groups on Reddit.

After we collected enough data in our Weibo study in 2018 using a combination of API and crawling, Weibo abruptly changed its API policy in 2019 and tightened the API access for international researchers. Although we could still collect 500 datapoints in 2019 by keyword-based crawling, performing similar analysis needs the support and collaboration

of scholars or companies residing within China. Nevertheless, our study has demonstrated that social media in China can still provide meaningful insights into Chinese people's thoughts and focus, despite possible censorship and interference. Therefore, social media is still one of the best windows for international scholars to understand the genuine concern for public health concern in China. More health concerns will emerge in China in the future with the rapid aging population; therefore, we must keep listening to the true concerns of Chinese people on health issues and fight the improper propaganda hiding those issues.

Meanwhile, we are probably the very first group studying Yelp and Google Maps reviews in such a volume. Before our work, one literature review paper identified that prior research had studied hospital reviews on HealthGrades, RateMDs, Vitals, and Yelp on smaller scales, but not Google Maps (Hong et al., 2019). We have also noticed that the convenient Yelp academic dataset (11310 reviews) only has data in 10 states and the most data are from Arizona (53.79%) and Nevada (38.30%). Our study has overcome the significant geographical bias and has ensured our finding is fair and comprehensive to describe the entire U.S. Moreover, although the current data is only collected on urgent cares, the same methods can be applied easily to other types of facilities that are interesting for public health researchers (e.g., hospitals, schools, restaurants) if given enough funding. Our approach is specifically powerful to investigate the primary health concern of the health facilities in rural America at county levels, an area that is often underrepresented in public health research.

To ensure that policymakers are studying genuine citizens, the direction of identifying bots and malicious actors on different platforms is my very next step. I plan to estimate the level of activity of bot and malicious users on Reddit and Google Maps soon. Last but not

least, more pre-trained natural language models are being rapidly developed. I can also work on the direction to improve the performance of existing LDA models by combining those new pre-trained models.

Social media is a product of human innovations and an extension of society. It gives researchers a unique lens to understand humanity, but it also inherits the problems of humanity at the same time. Researchers should always pay attention to whether the data gathered on social media is representative of the people they aim to study or help, whether the data analysis methods have instrumental bias, and whether the literature has in-depth discussed the same problem in the pre-social media era.

Social media might be virtual, but social media's impact is real. We need to keep the study of social media rapid and rigorous simultaneously.

Citations

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems*, *17*(2), 3.

Adams, J., Bledsoe, G. H., & Armstrong, J. H. (2016). Are pain management questions in patient satisfaction surveys driving the opioid epidemic? *American Journal of Public Health*, *106*(6), 985.

Agarwal, A. K., Guntuku, S. C., Meisel, Z. F., Pelullo, A., Kinkle, B., & Merchant, R. M. (2021). Analyzing Online Reviews of Substance Use Disorder Treatment Facilities in the USA Using Machine Learning. *Journal of General Internal Medicine*, 1–4.

Ahmad, F., Rossen, L., & Sutton, P. (2021). *Provisional drug overdose death counts*. National Center for Health Statistics.

Aletras, N., & Stevenson, M. (2013). Evaluating topic coherence using distributional semantics. *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)–Long Papers*, 13–22.

Algan, Y., Guriev, S., Papaioannou, E., & Passari, E. (2017). The European trust crisis and the rise of populism. Brookings Papers on Economic Activity, 2017(2), 309–400.

Anderson, C. (2021). Fake news is not a virus: On platforms and their effects. Communication Theory, 31(1), 42–61.

Antibiotic resistance: Multi-country public awareness survey. (2015). *World Health Organization*.

Attwell, K., Navin, M. C., Lopalco, P. L., Jestin, C., Reiter, S., & Omer, S. B. (2018). Recent vaccine mandates in the United States, Europe and Australia: A comparative study. *Vaccine*, *36*(48), 7377–7384.

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., & Blackburn, J. (2020). The pushshift reddit dataset. Proceedings of the International AAAI Conference on Web and Social Media, 14, 830–839.

BBC. (2016). China Shandong illegal vaccine scandal sparks anger. *BBC*. https://www.bbc.com/news/world-asia-china-35859927.

Bergstrom K. "Don't feed the troll": Shutting down debate about community expectations on Reddit. com. First Monday. Published online 2011.

Betsch, C., Brewer, N. T., Brocard, P., Davies, P., Gaissmaier, W., Haase, N., Leask, J.,

Renkewitz, F., Renner, B., Reyna, V. F., & others. (2012). Opportunities and challenges of Web 2.0 for vaccination decisions. *Vaccine*, *30*(25), 3727–3733.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84.

Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. *Proceedings of the 23rd International Conference on Machine Learning*, 113–120.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022.

Borgman, C. L. (2016). *Big data, little data, no data: Scholarship in the networked world*. MIT press.

Boussalis, C., & Coan, T. G. (2016). Text-mining the signals of climate change doubt. *Global Environmental Change*, *36*, 89–100.

Boyd, D., & Crawford, K. (2011). Six provocations for big data. A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society.

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. *American Journal of Public Health*, *108*(10), 1378–1384.

Broniatowski, D. A., Klein, E. Y., May, L., Martinez, E. M., Ware, C., & Reyna, V. F. (2018). Patients' and clinicians' perceptions of antibiotic prescribing for upper respiratory infections in the acute care setting. *Medical Decision Making*, *38*(5), 547–561.

Broniatowski, D. A., Klein, E. Y., & Reyna, V. F. (2015). Germs Are Germs, and Why Not Take a Risk? Patients' Expectations for Prescribing Antibiotics in an Inner-City Emergency Department. *Medical Decision Making*, *35*(1), 60–67. https://doi.org/10.1177/0272989X14553472

Broniatowski, D. A., & Reyna, V. F. (2019). To illuminate and motivate: A fuzzy-trace model of the spread of information online. *Computational and Mathematical Organization Theory*, 1–34.

Broockman, D. E., Kaufman, A. R., & Lenz, G. S. (2021). Heuristic Projection: How Interest Group Cues Can Harm Voters' Judgements. Working Paper. https://osf. io/6yskq.

Broockman, D. E., & Skovron, C. (2018). Bias in perceptions of public opinion among political elites. American Political Science Review, 112(3), 542–563.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, *6*(1), 3–5.

Cao, L., Wang, H.-Q., Zheng, J.-S., Yuan, P., Cao, L.-S., Zhang, G.-M., & others. (2012). National immunization coverage survey in China after integrated more vaccines into EPI since 2008. *Chinese Journal of Vaccines and Immunization*, *5*, 007.

Cairney, P., & Wellstead, A. (2021). COVID-19: Effective policymaking depends on trust in experts, politicians, and the public. Policy Design and Practice, 4(1), 1–14.

Callon, M., Lascoumes, P., & Barthe, Y. (2011). Acting in an uncertain world: An essay on technical democracy. MIT press.

Casciotti, D. M., Smith, K. C., Tsui, A., & Klassen, A. C. (2014). Discussions of adolescent sexuality in news media coverage of the HPV vaccine. *Journal of Adolescence*, *37*(2), 133–143.

Cawkwell, P. B., Lee, L., Weitzman, M., & Sherman, S. E. (2015). Tracking hookah bars in New York: Utilizing yelp as a powerful public health tool. *JMIR Public Health and Surveillance*, *1*(2), e19.

Chandler, D. (2015). A world without causation: Big data and the coming of age of posthumanism. *Millennium*, *43*(3), 833–851.

Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Neural Information Processing Systems*, *22*, 288–296.

China.org.cn. (2010). Shanxi vaccine scandal. *China.Org.Cn*. http://www.china.org.cn/china/2010-04/08/content_19769756.htm.

Compton, W. M., Jones, C. M., Stein, J. B., & Wargo, E. M. (2019). Promising roles for pharmacists in addressing the US opioid crisis. *Research in Social and Administrative Pharmacy*, *15*(8), 910–916.

Connell, C. M., Gilreath, T. D., Aklin, W. M., & Brex, R. A. (2010). Social-ecological influences on patterns of substance use among non-metropolitan high school students. *American Journal of Community Psychology*, *45*(1), 36–48.

Conway, K. P., Vullo, G. C., Nichter, B., Wang, J., Compton, W. M., Iannotti, R. J., & Simons-Morton, B. (2013). Prevalence and patterns of polysubstance use in a nationally representative sample of 10th graders in the United States. *Journal of Adolescent Health*, *52*(6), 716–723.

Costola, M., Iacopini, M., & Santagiustina, C. R. (2021). On the "mementum" of Meme Stocks. Economics Letters, 207, 110021.

Cresci S, Lillo F, Regoli D, Tardelli S, Tesconi M. Cashtag piggybacking: Uncovering spam and bot activity in stock microblogs on Twitter. ACM Transactions on the Web (TWEB). 2019;13(2):1-27.

Daley, E. M., Vamos, C. A., Zimet, G. D., Rosberger, Z., Thompson, E. L., & Merrell, L. (2016). The feminization of HPV: reversing gender biases in US human papillomavirus vaccine policy. *American Journal of Public Health*, *106*(6), 983.

Dalton, C. M., Taylor, L., & Thatcher, J. (2016). Critical data studies: A dialog on data and space. *Big Data & Society*, *3*(1), 2053951716648346.

Dasgupta, N., Beletsky, L., & Ciccarone, D. (2018). Opioid crisis: No easy fix to its social and economic determinants. *American Journal of Public Health*, *108*(2), 182–186.

Daskivich, T. J., Houman, J., Fuller, G., Black, J. T., Kim, H. L., & Spiegel, B. (2018). Online physician ratings fail to predict actual performance on measures of quality, value, and peer review. *Journal of the American Medical Informatics Association*, *25*(4), 401–407.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, *41*(6), 391–407.

Dehghani, M., Sagae, K., Sachdeva, S., & Gratch, J. (2014). Analyzing political rhetoric in conservative and liberal weblogs related to the construction of the "Ground Zero Mosque." *Journal of Information Technology & Politics*, *11*(1), 1–14.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. ArXiv Preprint ArXiv:1810.04805.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics*, *41*(6), 570–606.

Dixon, G., & Clarke, C. (2013). The effect of falsely balanced reporting of the autism–vaccine controversy on vaccine safety perceptions and behavioral intentions. *Health Education Research*, *28*(2), 352–359.

Dredze, M., Broniatowski, D. A., & Hilyard, K. M. (2016). Zika vaccine misconceptions: A social media analysis. *Vaccine*, *34*(30), 3441.

Du, F., Chantler, T., Francis, M. R., Sun, F. Y., Zhang, X., Han, K., Rodewald, L., Yu, H., Tu, S., Larson, H., & others. (2020). The determinants of vaccine hesitancy in China: A cross-sectional study following the Changchun Changsheng vaccine incident. *Vaccine*, *38*(47), 7464–7471.

Dubé, E., Vivion, M., & MacDonald, N. E. (2015). Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: Influence, impact and implications. *Expert Review of Vaccines*, *14*(1), 99–117.

Eversman, M. H. (2015). "We want a living solution": Views of harm reduction programs in black US communities. *Journal of Ethnicity in Substance Abuse*, *14*(2), 187–207.

Facebook Reports Fourth Quarter and Full Year 2020 Results. (2021). Facebook. https://s21.q4cdn.com/399680738/files/doc_financials/2020/q4/FB-12.31.2020-Exhibit-99.1.pdf

Fairbrother, M. (2017). Environmental attitudes and the politics of distrust. Sociology Compass, 11(5), e12482.

Ferrara E, Varol O, Davis C, Menczer F, Flammini A. The rise of social bots. Communications of the ACM. 2016;59(7):96-104.

Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. New Media & Society, 20(7), 2450–2468.

Fleming-Dutra, K. E., Hersh, A. L., Shapiro, D. J., Bartoces, M., Enns, E. A., File, T. M., Finkelstein, J. A., Gerber, J. S., Hyun, D. Y., Linder, J. A., & others. (2016). Prevalence of inappropriate antibiotic prescriptions among US ambulatory care visits, 2010-2011. *Jama*, *315*(17), 1864–1873.

Forster, A., Wardle, J., Stephenson, J., & Waller, J. (2010). Passport to promiscuity or lifesaver: Press coverage of HPV vaccination and risky sexual behavior. *Journal of Health Communication*, *15*(2), 205–217.

Friedman, A. L., & Shepeard, H. (2007). Exploring the knowledge, attitudes, beliefs, and communication preferences of the general public regarding HPV: findings from CDC focus group research and implications for practice. *Health Education & Behavior*, *34*(3), 471–485.

Fu, K., Chan, C., & Chau, M. (2013). Assessing censorship on microblogs in China: Discriminatory keyword analysis and the real-name registration policy. *IEEE Internet Computing*, *17*(3), 42–50.

Fujigaki, Y., & Tsukahara, T. (2011). STS implications of Japan's 3/11 crisis. East Asian Science, Technology and Society: An International Journal, 5(3), 381–394.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144.

Gilens, M., & Page, B. I. (2014). Testing theories of American politics: Elites, interest groups, and average citizens. Perspectives on Politics, 12(3), 564–581.

Glenski, M., Saldanha, E., & Volkova, S. (2019). Characterizing speed and scale of cryptocurrency discussion spread on reddit. The World Wide Web Conference, 560–570.

Gollust, S. E., Dempsey, A. F., Lantz, P. M., Ubel, P. A., & Fowler, E. F. (2010). Controversy undermines support for state mandates on the human papillomavirus vaccine. *Health Affairs*, *29*(11), 2041–2046.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029.

Grimmer, J. (2010). A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases. *Political Analysis*, *18*(1), 1–35.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on twitter during the 2016 US Presidential election. *Science*, *363*(6425), 374–378.

Grover, V., Lindberg, A., Benbasat, I., & Lyytinen, K. (2020). The perils and promises of big data research in information systems. *Journal of the Association for Information Systems*, *21*(2), 9.

Gryczynski, J., Jaffe, J. H., Schwartz, R. P., Dušek, K. A., Gugsa, N., Monroe, C. L., O'Grady, K. E., Olsen, Y. K., & Mitchell, S. G. (2013). Patient perspectives on choosing buprenorphine over methadone in an urban, Equal-Access system. *The American Journal on Addictions*, *22*(3), 285–291.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1), eaau4586.

Guess, A., Nyhan, B., Lyons, B., & Reifler, J. (2018). Avoiding the echo chamber about echo chambers. *Knight Foundation*, *2*.

Guston, D. H. (2014). Understanding 'anticipatory governance.' Social Studies of Science, 44(2), 218–242.

He, Q., Wang, H., Ma, Y., Wang, Z., Zhang, Z., Li, T., & Yang, Z. (2020). Changes in parents' decisions pertaining to vaccination of their children after the Changchun Changsheng vaccine scandal in Guangzhou, China. *Vaccine, 38*(43), 6751–6756.

Hertel-Fernandez, A., Mildenberger, M., & Stokes, L. C. (2019). Legislative staff and representation in Congress. American Political Science Review, 113(1), 1–18.

Hollingsworth, A., Ruhm, C. J., & Simon, K. (2017). Macroeconomic conditions and opioid abuse. *Journal of Health Economics*, *56*, 222–233.

Hong, J., & Wei, D. (2018, August 7). How Faulty Vaccines Sparked Consumer Panic in China. *Bloomberg*. https://www.bloomberg.com/news/articles/2018-08-07/how-faulty-vaccines-sparked-consumer-panic-in-china-quicktake

Hong YA, Liang C, Radcliff TA, Wigfall LT, Street RL. What do patients say about doctors online? A systematic review of studies on patient online reviews. Journal of medical Internet research. 2019;21(4):e12521.

Hu, D., Liu, C. M. H., Hamdy, R., Cziner, M., Fung, M., Dobbs, S., ... & Broniatowski, D. A. (2021). Questioning the Yelp Effect: Mixed Methods Analysis of Web-Based Reviews of Urgent Cares. Journal of medical Internet research, 23(10), e29406.

Hu, D., Martin, C., Dredze, M., & Broniatowski, D. A. (2020). Chinese social media suggest decreased vaccine acceptance in China: An observational study on Weibo following the 2018 Changchun Changsheng vaccine incident. *Vaccine*, *38*(13), 2764–2770.

Huhn, A. S., Strain, E. C., Tompkins, D. A., & Dunn, K. E. (2018). A hidden aspect of the US opioid crisis: Rise in first-time treatment admissions for older adults with opioid use disorder. *Drug and Alcohol Dependence*, *193*, 142–147.

Hurtado S, Ray P, Marculescu R. Bot detection in reddit political discussion. In: Proceedings of the Fourth International Workshop on Social Sensing. ; 2019:30-35.

Hyun, D. (n.d.). *What Drives Inappropriate Antibiotic Use in Outpatient Care?* Pew Charitable Trust. https://www.pewtrusts.org/en/research-and-analysis/issue-briefs/2017/06/what-drives-inappropriate-antibiotic-use-in-outpatient-care

Isoaho, K., Gritsenko, D., & Mäkelä, E. (2021). Topic modeling and text analysis for qualitative policy research. *Policy Studies Journal*, *49*(1), 300–324.

Jasanoff, S., Hurlbut, J. B., & Saha, K. (2015). CRISPR democracy: Gene editing and the need for inclusive deliberation. Issues in Science and Technology, 32(1), 25–32.

Jasanoff, S., & Simmet, H. R. (2017). No funeral bells: Public reason in a 'post-truth'age. Social Studies of Science, 47(5), 751–770.

Jena, A. B., Goldman, D. P., & Seabury, S. A. (2015). Incidence of sexually transmitted infections after human papillomavirus vaccination among adolescent females. *JAMA Internal Medicine*, *175*(4), 617–623.

Jhaver S, Birman I, Gilbert E, Bruckman A. Human-machine collaboration for content regulation: The case of reddit automoderator. ACM Transactions on Computer-Human Interaction (TOCHI). 2019;26(5):1-35.

Jin, F., Dougherty, E., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on twitter. *Proceedings of the 7th Workshop on Social Network Mining and Analysis*, 8.

Johnson, K. L., Lin, M.-Y., Cabral, H., Kazis, L. E., & Katz, I. T. (2017). Variation in human papillomavirus vaccine uptake and acceptability between female and male adolescents and their caregivers. *Journal of Community Health*, *42*(3), 522–532

Jordan, A., Huitema, D., Van Asselt, H., & Forster, J. (2018). Governing climate change: Polycentricity in action? Cambridge University Press.

Joseph, K., Landwehr, P. M., & Carley, K. M. (2014). Two 1% s don't make a whole: Comparing simultaneous samples from Twitter's streaming API. International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction, 75–83.

Kadry, B., Chu, L. F., Kadry, B., Gammas, D., & Macario, A. (2011). Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. *Journal of Medical Internet Research*, *13*(4), e95.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54–86.

Kapantai, E., Christopoulou, A., Berberidis, C., & Peristeras, V. (2021). A systematic literature review on disinformation: Toward a unified taxonomical framework. New Media & Society, 23(5), 1301–1326.

Kennedy, J. (2019). Populist politics and vaccine hesitancy in Western Europe: An analysis of national-level data. *European Journal of Public Health*, *29*(3), 512–516.

Kim, E. H.-J., Jeong, Y. K., Kim, Y., Kang, K. Y., & Song, M. (2016). Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. *Journal of Information Science*, *42*(6), 763–781.

Kitchin, R., & Lauriault, T. P. (2015). Small data in the era of big data. *GeoJournal*, *80*(4), 463–475.

Klein, E. Y., Martinez, E. M., May, L., Saheed, M., Reyna, V., & Broniatowski, D. A. (2017). Categorical Risk Perception Drives Variability in Antibiotic Prescribing in the Emergency Department: A Mixed Methods Observational Study. *Journal of General Internal Medicine*, *32*(10), 1083–1089. https://doi.org/10.1007/s11606-017-4099-6

Krippendorff, K. (2013). *Content Analysis: An Introduction to Its Methodology*. SAGE.

Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., & Karahalios, K. (2017). Quantifying search bias: Investigating sources of bias for political searches in social media. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 417–432.

Labrinidis, A., & Jagadish, H. V. (2012). Challenges and opportunities with big data. Proceedings of the VLDB Endowment, 5(12), 2032–2033.

Lagasse, J. (2019, December 20). Doctors' online reviews show more polarization, have more impact than those of other industries. *Healthcare Finance*. https://www.healthcarefinancenews.com/news/doctors-online-reviews-show-more-polarization-have-more-impact-those-other-industries

Lama, Y., Hu, D., Jamison, A., Quinn, S. C., & Broniatowski, D. A. (2019a). Characterizing Trends in Human Papillomavirus Vaccine Discourse on Reddit (2007-2015): An Observational Study. *JMIR Public Health and Surveillance*, *5*(1), e12480.

Larson, H. J., & Broniatowski, D. A. (2021). Volatility of vaccine confidence. In Science (Vol. 371, Issue 6536, pp. 1289–1289). American Association for the Advancement of Science.

Lasco, G., & Larson, H. J. (2020). Medical populism and immunisation programmes: Illustrative examples and consequences for public health. Global Public Health, 15(3), 334–344.

Lau, J. H., Newman, D., & Baldwin, T. (2014). Machine reading tea leaves: Automatically evaluating topic coherence and topic model quality. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 530–539.

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google Flu: Traps in big data analysis. *Science*, *343*(6176), 1203–1205.

López, L., Weissman, J. S., Schneider, E. C., Weingart, S. N., Cohen, A. P., & Epstein, A. M. (2009). Disclosure of Hospital Adverse Events and Its Association With Patients' Ratings of the Quality of Care. *Archives of Internal Medicine*, *169*(20), 1888–1894. https://doi.org/10.1001/archinternmed.2009.387

Luceri L, Giordano S, Ferrara E. Detecting troll behavior via inverse reinforcement

learning: A case study of russian trolls in the 2016 us election. In: Proceedings of the International AAAI Conference on Web and Social Media. Vol 14. ; 2020:417-427.

Macnaghten, P., Kearnes, M. B., & Wynne, B. (2005). Nanotechnology, governance, and public deliberation: What role for the social sciences? Science Communication, 27(2), 268–291.

Madras, B. K. (2017). The surge of opioid use, addiction, and overdoses: Responsibility and response of the US health care system. *Jama Psychiatry*, *74*(5), 441–442.

Mangione-Smith, R., McGlynn, E. A., Elliott, M. N., Krogstad, P., & Brook, R. H. (1999). The relationship between perceived parental expectations and pediatrician antimicrobial prescribing behavior. *Pediatrics*, *103*(4), 711–718.

May, L., Gudger, G., Armstrong, P., Brooks, G., Hinds, P., Bhat, R., Moran, G. J., Schwartz, L., Cosgrove, S. E., Klein, E. Y., & others. (2014). Multisite exploration of clinical decision-making for antibiotic use by emergency medicine providers using quantitative and qualitative methods. *Infection Control and Hospital Epidemiology*, *35*(9), 1114.

McCallum, A. K. (2002). *MALLET: A Machine Learning for Language Toolkit*.

Mckenna, M. (2018, March 22). The Yelping of The American Doctor. *Wired*. https://www.wired.com/story/the-yelping-of-the-american-doctor/

McRee, A.-L., Reiter, P. L., & Brewer, N. T. (2012). Parents' Internet use for information about HPV vaccine. *Vaccine*, *30*(25), 3757–3762.

Meldrum, M. L. (2016). The ongoing opioid prescription epidemic: Historical context. *American Journal of Public Health*, *106*(8), 1365.

Menendez, M. E., Mellema, J. J., & Ring, D. (2015). Attitudes and self-reported practices of hand surgeons regarding prescription opioid use. *Hand*, *10*(4), 789–795.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. ArXiv Preprint ArXiv:1301.3781.

Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science*, *50*(1), 181–201.

Mills, K. A. (2018). What are the threats and potentials of big data for qualitative research? *Qualitative Research*, *18*(6), 591–603.

Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 262–272.

Mitchell, A., Gottfried, J., Stocking, G., Walker, M., & Fedeli, S. (2019). Many Americans say made-up news is a critical problem that needs to be fixed. Pew Research Center, 5, 2019.

Morstatter, F., & Liu, H. (2018). In search of coherence and consensus: Measuring the interpretability of statistical topics. *Journal of Machine Learning Research*, *18*(169), 1–32.

Moshier, S. J., McHugh, R. K., Calkins, A. W., Hearon, B. A., Rosellini, A. J., Weitzman, M. L., & Otto, M. W. (2012). The role of perceived belongingness to a drug subculture among opioid-dependent patients. *Psychology of Addictive Behaviors*, *26*(4), 812.

Munksgaard, R., & Demant, J. (2016). Mixing politics and crime–The prevalence and decline of political discourse on the cryptomarket. *International Journal of Drug Policy*, *35*, 77–83.

Nyhan, B. (2020). Facts and myths about misperceptions. Journal of Economic Perspectives, 34(3), 220–236.

Nyhan, B., Dickinson, F., Dudding, S., Dylgjeri, E., Neiley, E., Pullerits, C., Seog, M., Simpson, A., Szilagyi, H., & Walmsley, C. (2016). Classified or coverup? The effect of redactions on conspiracy theory beliefs. Journal of Experimental Political Science, 3(2), 109–123.

O'callaghan, D., Greene, D., Carthy, J., & Cunningham, P. (2015). An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, *42*(13), 5645–5657.

Ong, S., Nakase, J., Moran, G. J., Karras, D. J., Kuehnert, M. J., Talan, D. A., Group, Emerge. I. N. S., & others. (2007). Antibiotic use for emergency department patients with upper respiratory infections: Prescribing practices, patient expectations, and patient satisfaction. *Annals of Emergency Medicine*, *50*(3), 213–220.

Orhurhu, M. S., Salisu, B., Sottosanti, E., Abimbola, N., Urits, I., Jones, M., Viswanath, O., Kaye, A. D., Simopoulos, T., & Orhurhu, V. (2019). Chronic pain practices: An evaluation of positive and negative online patient reviews. *Pain Physician*, *22*(5), E477-86.

Orr, D., Baram-Tsabari, A., & Landsman, K. (2016). Social media as a platform for health-related public debates and discussions: The Polio vaccine on Facebook. *Israel Journal of Health Policy Research*, *5*(1), 34.

Palms, D. L., Hicks, L. A., Bartoces, M., Hersh, A. L., Zetts, R., Hyun, D. Y., & Fleming-Dutra, K. E. (2018). Comparison of antibiotic prescribing in retail clinics,

urgent care centers, emergency departments, and traditional ambulatory care settings in the United States. *JAMA Internal Medicine*, *178*(9), 1267–1269.

Paul, M. J., Wallace, B. C., & Dredze, M. (2013). What affects patient (dis) satisfaction? Analyzing online doctor ratings with a joint topic-sentiment model. *AAAI Workshop on Expanding the Boundaries of Health Informatics Using AI*, 53–58.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 1532–1543.

Peterson, E., & Allamong, M. B. (2021). The Influence of Unknown Media on Public Opinion: Evidence from Local and Foreign News Sources. American Political Science Review, 1–15

Phoenix, J. H., Atkinson, L. G., & Baker, H. (2019). Creating and communicating social research for policymakers in government. Palgrave Communications, 5(1), 1–11.

Physicians Working Together. (2017). *Remove online reviews of doctors!* https://www.change.org/p/yelp-remove-online-reviews-of-doctors

Pierce, J. R., & Schott, P. K. (2020). Trade liberalization and mortality: Evidence from US counties. *American Economic Review: Insights*, *2*(1), 47–64.

Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science*, *54*(1), 209–228.

Quintero Johnson, J., Sionean, C., & Scott, A. M. (2011). Exploring the presentation of news information about the HPV vaccine: A content analysis of a representative sample of US newspaper articles. *Health Communication*, *26*(6), 491–501.

Radcliffe, D. (2017). Understanding social media in China. *International Journalists' Network*. https://ijnet.org/en/story/understanding-social-media-china.

Ranard, B. L., Werner, R. M., Antanavicius, T., Schwartz, H. A., Smith, R. J., Meisel, Z. F., Asch, D. A., Ungar, L. H., & Merchant, R. M. (2016). Yelp reviews of hospital care can supplement and inform traditional surveys of the patient experience of care. *Health Affairs*, *35*(4), 697–705.

Ratner, N. B., & Gleason, J. B. (2004). Psycholinguistics. In L. R. Squire (Ed.), *Encyclopedia of Neuroscience* (pp. 1199–1204). Academic Press. https://doi.org/10.1016/B978-008045046-9.01893-3

Reiter, P. L., Brewer, N. T., Gottlieb, S. L., McRee, A.-L., & Smith, J. S. (2009). Parents' health beliefs and HPV vaccination of their adolescent daughters. *Social Science &*

Medicine, *69*(3), 475–480.

Riddell, A., Hopper, T., Luo, S., Leinweber, K., & Grivas, A. (n.d.). *Lda-project*. https://doi.org/10.5281/zenodo.1412135

Roberts, M. E., Stewart, B. M., Tingley, D., Airoldi, E. M., & others. (2013). The structural topic model and applied social science. *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*, *4*, 1–20.

Rodrik, D. (2021). Why does globalization fuel populism? Economics, culture, and the rise of right-wing populism. Annual Review of Economics, 13, 133–170.

Rogers, R. (2020). Research note: The scale of Facebook's problem depends upon how 'fake news' is classified. The Harvard Kennedy School (HKS) Misinformation Review.

Russell, B. S., Trudeau, J. J., & Leland, A. J. (2015). Social influence on adolescent polysubstance use: The escalation to opioid use. *Substance Use & Misuse*, *50*(10), 1325–1331.

Sanfilippo M, Yang S, Fichman P. Managing online trolling: From deviant to social and political trolls. Published online 2017.

Scholl, L., Seth, P., Kariisa, M., Wilson, N., & Baldwin, G. (2019). Drug and opioid-involved overdose deaths—United States, 2013–2017. *Morbidity and Mortality Weekly Report*, *67*(51–52), 1419.

Sevenans, J., Soontjens, K., & Walgrave, S. (2021). Inequality in the public priority perceptions of elected representatives. West European Politics, 1–24.

Sherman, D. K., Shteyn, M. F., Han, H., & Van Boven, L. (2021). The exchange between citizens and elected officials: A social psychological framework for citizen climate activists. Behavioural Public Policy, 5(4), 576–605.

Stearns, C. R., Gonzales, R., Camargo, C. A., Jr, Maselli, J., & Metlay, J. P. (2009). Antibiotic prescriptions are associated with increased patient satisfaction with emergency department visits for acute respiratory tract infections. *Academic Emergency Medicine*, *16*(10), 934–941.

Staerklé, C., & Green, E. G. (2018). Right-wing populism as a social representation: A comparison across four European countries. Journal of Community & Applied Social Psychology, 28(6), 430–445.

Stivers, T., Mangione-Smith, R., Elliott, M. N., McDonald, L., & Heritage, J. (2003). Why do physicians think parents expect antibiotics? What parents report vs what

physicians believe. *Journal of Family Practice*, *52*(2), 140–147.

Stoicea, N., Costa, A., Periel, L., Uribe, A., Weaver, T., & Bergese, S. D. (2019). Current perspectives on the opioid crisis in the US healthcare system: A comprehensive literature review. *Medicine*, *98*(20).

Strach, P., Zuber, K., & Pérez-Chiqués, E. (2020). Why policies fail: The illusion of services in the opioid epidemic. Journal of Health Politics, Policy and Law, 45(2), 341–364.

Strand, M. A., Eukel, H., & Burck, S. (2019). Moving opioid misuse prevention upstream: A pilot study of community pharmacists screening for opioid misuse risk. *Research in Social and Administrative Pharmacy*, *15*(8), 1032–1036.

Sun, R., Wang, X., Lin, L., Zhang, N., Li, L., & Zhou, X. (2021). The impact of negative emotional reactions on parental vaccine hesitancy after the 2018 vaccine event in China: A cross-sectional survey. *Human Vaccines & Immunotherapeutics*, 1–10.

Sundling, R. A., Logan, D. B., Tawancy, C. H., So, E., Lee, J., & Logan, K. (2020). Opioid prescribing habits of podiatric surgeons following elective foot and ankle surgery. *The Foot*, *45*, 101710.

Tan, R. (2018). *Vaccine scandal gripping China could cause serious problems for the government*. https://web.archive.org/web/20180816083403/https://www.washingtonpost.com/news/worldviews/wp/2018/07/26/vaccine-scandal-gripping-china-could-cause-serious-problems-for-the-government/.

The Lancet. (2018). Vaccine scandal and confidence crisis in China. *Lancet (London, England)*, *392*(10145), 360.

Traxler, M., & Gernsbacher, M. A. (2011). *Handbook of psycholinguistics*. Elsevier.

Tromble, R. (2021). Where have all the data gone? A critical reflection on academic digital research in the post-API age. *Social Media+ Society*, *7*(1), 2056305121988929.

Tucker, J. S., Pollard, M. S., de la Haye, K., Kennedy, D. P., & Green Jr, H. D. (2013). Neighborhood characteristics and the initiation of marijuana use and binge drinking. *Drug and Alcohol Dependence*, *128*(1–2), 83–89.

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. Eighth International AAAI Conference on Weblogs and Social Media.

Twitter. (2021). Investor Fact Sheet.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151.

Wagner, M. C., & Boczkowski, P. J. (2019). The reception of fake news: The interpretations and practices that shape the consumption of perceived misinformation. Digital Journalism, 7(7), 870–885.

Wang, S., Paul, M. J., & Dredze, M. (2014). Exploring health topics in Chinese social media: An analysis of Sina Weibo. *AAAI Workshop on the World Wide Web and Public Health Intelligence*, *23*, 20–23.

Watts, D. J., & Rothschild, D. (2017). Don't blame the election on fake news. Blame it on the media. Columbia Journalism Review. https://www.microsoft.com/en-us/research/publication/dont-blame-the-election-on-fake-news-blame-it-on-the-media/

Welschen, I., Kuyvenhoven, M., Hoes, A., & Verheij, T. (2004). Antibiotics for acute respiratory tract symptoms: Patients' expectations, GPs' management and patient satisfaction. *Family Practice*, *21*(3), 234–237.

Wenghofer, E. F., Wilson, L., Kahan, M., Sheehan, C., Srivastava, A., Rubin, A., & Brathwaite, J. (2011). Survey of Ontario primary care physicians' experiences with opioid prescribing. *Canadian Family Physician*, *57*(3), 324–332.

Wilson, N. (2020). Drug and opioid-involved overdose deaths—United States, 2017–2018. *MMWR. Morbidity and Mortality Weekly Report*, *69*.

Wolf, E. E. A. (2021). Dismissing the "vocal minority": How policy conflict escalates when policymakers label resisting citizens. Policy Studies Journal, 49(2), 640–663.

Xie, P., Yang, D., & Xing, E. (2015). Incorporating word correlation knowledge into topic modeling. *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 725–734.

Xu, J., & Liu, C. (2021). Acceptance of a COVID-19 Vaccine Before it is Available in China During the Pandemic. *International Journal of Public Health*, *66*.

Xu, X., Mao, Z. M., & Halderman, J. A. (2011). Internet censorship in China: Where does the filtering occur? *International Conference on Passive and Active Network Measurement*, 133–142.

Yaraghi, N., Wang, W., Gao, G. G., & Agarwal, R. (2018). How online quality ratings influence patients' choice of medical providers: Controlled experimental survey study. *Journal of Medical Internet Research*, *20*(3), e99.

Zgierska, A., Rabago, D., & Miller, M. M. (2014). Impact of patient satisfaction ratings

on physicians and clinical care. *Patient Preference and Adherence*, *8*, 437.

Zhang, K. Z., Zhao, S. J., Cheung, C. M., & Lee, M. K. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, *67*, 78–89.

Appendix A: Qualitative codebook of the HPV Project

Table A-1: Qualitative codebook of the HPV Project

Codebook	Description
Cancer in General	Related to HPVs relationship to cancer. Concerns related to cancer. HPV vaccine prevents cancer. Awareness of cancer risk. See child at risk for cancer. Can be any realted type of cancer: cervical, throat, penile, anal, etc.
General Sexual Behavior	Related to HPV and sexual activity . Promiscuity is part of it, but it is broader to include - discussions of sex, sexual debut, risky sexual behavior, age, multiple sex partners, etc. Not enough to be discussing sex, needs to have an explicit connection to the HPV and/or vaccine. Acknowledgement that sex is related to HPV infection. Directionless/Judgement free.
Positive Vaccine Sentiment	View vaccine favorably, believe it is effective, view it as necessary, ultimately view it as a positive preventive measure.
Negative Vaccine Sentiment	View vaccine negatively, not "worth it", risky, unnecessary, side effects discussion. Also people economic arguements. Parents talking about parental rights.
Neither	Does not mention either cancer or promiscuity risk
Girl	Use female pronouns, "my daughter", discussion of sex organs or gendered medical treatments (e.g. cervix, pap smear. etc.)alone is not enough to indicate female gender, must be explicitly connected to the individual recieving the vaccine.
Boy	Use male pronouns, "my son", men.
Unspecified	"My child", Unclear, No people specified.
*** For both**	Tag both girl and boy if people discuss "everyone" or discuss both genders needing vaccine.
Other	Not include any links or images.

Appendix B: Development Strategy of the Gender Classifier

Step 1: Reconstruct the Reddit data based on the username of the messages.

Step 2: For each user, we use linguistic-explicit self-reporting regular expressions to scan through all messages contributed from that user to identify certain demographic traits like age, gender or race. For example, if one user has used the expression once like “I’m 25.”, “I am a woman,” “my husband” The algorithm would record the expression and give the user a temporal evaluation of “25”, “Female.” To ensure maximum precision of the method, we would only consider message with self-reporting grammar, e.g., “I am,” “I’m” and “my”. During this step, each regular expression is tested and modified to filter out the most common noise.

Step 3: After going through all messages of all users, we calculate the overall evaluation of each user. Because of our strict restriction in regular expressions. Some users would remain unknown after step 2 and step 3.

Appendix C: Guideline for Annotating the Weibo Vaccine Dataset

Category 1: Messages mainly about the vaccination of pets.

This category should be messages in which people are mentioning vaccine, but only about vaccinating the animals.

Example 1: “两个宠物级蓝双妹妹 三个半月疫苗两针已经做完 找有爱的家长 喜欢的问我价格惊喜！是真的惊喜#广州猫舍##卖猫##今天你有猫了吗#

Explanation: This is a very typical Weibo behavior: people advertise pets for sale. In this case, the kitten is vaccinated. In this message, both the keyword “cat” and “vaccine” are explicitly mentioned.

Example 2: “我在闲鱼发布了【松狮，八个多月，打了四针疫苗，】#退坑回血便宜出！”

Explanation: Another advertisement, however, in this case, the keyword “dog” is not found. However, the dog breed name is mentioned. Reading this breed name can be difficult, but Google image would help.

Example 3: “#全国科普日# #生活百科# 【给鱼打针你听说过吗？吃了以后会不会对人体有害？】市场上有些价格较高的鱼确实存在注水增重的现象，不过视频中给鱼打针主要是防病治病和催产。一般来说，青鱼、草鱼、鲮鱼等都需要助产，催产剂主要为糖蛋白，高温煮熟后会分解，不必担心。一般给小鱼注射的疫苗是安全有效的 ?...展开全文 c”

Explanation: Popular Science message about how to vaccinate fish.

Category 2: Messages about the vaccination scandal development.

This category should be messages in which people try to spread the vaccine scandal.

Example 1: “ 不只狂犬病疫苗造假！长春长生因百白破疫苗质量不合格再被处罚_搜狐新闻_搜狐网 http:\/\/t.cn\/RgRSY0T ? “

Explanation: The Chinese word “长春长生” is the name of the company involved in the scandal. ”狂犬病疫苗” (Rabies Vaccine) and “百白破疫苗” (DPT vaccine) are the two vaccines that went wrong.

Example 2: “ 先是狂犬疫苗，然后是百白破疫苗，也就是一类疫苗他们都没放过，那么有理由怀疑他们所有产品。那么，我现在能怎么办？#长生生物被罚 344 万# ?”

Explanation: Another similar commentary of the scandal, where the Chinese word “长生生物” is similar to “长春长生”, but the former added the word “biological company” after the name.

Category 3: Messages mainly about the harm/injury that might be caused by vaccination.

This category should be messages in which people express their concern for the potential injury and harm of the vaccine.

Example 1: “ 他们生产的假疫苗 正源源不断注入你孩子身体中

Explanation: The message is saying that “the fake vaccine is hurting your baby.”

Example 2: “ 下个月我家小宝就是打疫苗了，可以现在出现疫苗造假问题，真心不敢给孩子打了，怕怕怕怕怕怕怕怕”

Explanation: The message is saying that “ My baby will need to have a vaccine next month,

now I am really afraid of the possible problem of the fake vaccine.

Category 4: Messages mainly about the Trust of Government/Society/Moral Value.

This category is where people speak about the scandal in anger and disappointment. However, because of the censorship mechanism, people do not directly use the word "government" or "president."

Example 1: "从毒奶粉，再到疫苗事件，真的让人心寒！企业没良心，国家监管不到位！说明什么? ?"

Explanation: The message is saying that "the poisonous formulas and the vaccine incident has made me heartbroken."

Example 2: " 毒奶粉，假疫苗，我们的孩子究竟要经历多少风暴，才能长大成人?"

Explanation: The message is saying that "our baby must survive poisonous formulas and fake vaccine now."

Example 3: " 疫苗无毒人有毒，为官不知百姓苦；当官不为民做主，不如回家卖红薯。?"

Explanation: A sarcastic Chinese poem ridicules the government.

Category 5: Self-Reporting Vaccination Experience

This category is where the message explicitly self-reports vaccination or intention to vaccinate.

Example 1: " 104day~上午去打了疫苗 这么小的 baby 已经打过一针卡介苗 两针乙肝疫苗 两针 13 价肺炎 一针四联 一针脊灰灭活好多针"

Explanation: The message is saying that "my baby had taken many different vaccines so

far.”

Example 2: ” 今天米拉去打疫苗，体检时大夫说她超重了，不是吧，这么小就得减肥**”**

Explanation: The message is saying that “My baby took the vaccine today and the doctor said that she is overweight and need lose some weight.”

Appendix D: Google Maps Reviews Collection Algorithm

For each urgent care in our dataset, we retrieved the unique hash value assigned by Google Maps. For example, the School of Media and Public Affairs of GW has the following address on Google Maps:

https://www.google.com/maps/place/George+Washington+University+School+of+Media+and+Public+Affairs/@38.8998658,-77.0482864,19z/data=!4m5!3m4!1s0x89b7b7b08af3296f:0x3e32dedd82d144b3!8m2!3d38.899823!4d-77.0461864

In this case, the string "1s0x89b7b7b08af3296f:0x3e32dedd82d144b3" is the aforementioned hash value.

Hiring Mturk workers with high performance to collect the URLs

Since Google periodically updates these hash values, we hired human annotators to manually gather google URLs for our task. We conducted a pilot run in which we assessed Mturk workers' performance. Specifically, we incorporated 100 URLs for which we already knew the correct response into each task. Workers who failed to submit the URLs in correct format were rejected. All Mturk workers are reimbursed for their work if they correctly submitted all URLs in the correct format. Ultimately, 20 workers were selected based on their performance on this pilot task. Each worker received $0.09 for each URL collected.

Pseudocode for the crawling algorithms

Each URL collected by an Mturk worker was used as input to a software program that we wrote to retrive the reviews from corresponding urgent cares. Here, we presented the pseudocode of our crawling algorithms. The code is composed of 4 main parts: crawl, parse, count and scroll. The core strategy is to expand the comment section thoroughly using a function that mimics human's scroll and click action. After that, we parse and save all

```
def crawl(self, url):
        try:
            Assess the validity of the URL
        except when the URL is not valid:
            Print Exception
            Ignore this URL
            Continue on next URL
        try:
            Assess the validity of the review section :
            Using SeleniumAPI the_driver.find_element_by_css_selector
        except when the URL is not valid:
            Print Exception
            Ignore this URL
            Continue on next URL
        count_before_scroll ← Run Count function
        Run Scroll Function
        count_after_scroll ← Run Count function (#again after scroll#)
        while count_after_scroll > count_before_scroll:
                (# The scrolling will not stop until no more comment is available#)
                count_before_scroll ← count_after_scroll

                        (#We updated the count_before_scroll#)
                         Run Scroll Function
                        (# Keep scrolling#)

                        count_after_scroll ← Run Count function
                        (#We updated the count_after_scroll#)
        Click All more button using SeleniumAPI:execute_script
        Run parse Function

def parse(self):
    parse review_text
    parse review_star
    parse review_username

def count(self):
    Count the current total of comments.
    return num

def scroll(self):
    Scroll the comment section Using SeleniumAPI: ActionChains
    Pause Time for 15 seconds (# Timing Control)
```

comment into an encrypted hard drive. All usernames are hashed during the crawling process to protect the privacy of all users. For each review, we obtained the following information: business index (to which urgent cares the review belongs), review text (if the user writes anything), stars (minimal star rating 1, maximum star rating is 5), state, and zip code. We also randomly generated a unique ID for each URL.