Quantifying & Characterizing Information Diets of Social Media Users

A dissertation submitted towards the degree
Doctor of Engineering (Dr.-Ing.)
of the Faculty of Mathematics and Computer Science
of Saarland University

by

Juhi Kulshrestha

Saarbrücken
July 2018

Date of Colloquium:	20^{th} December 2018
Dean of Faculty:	Prof. Dr. Sebastian Hack
Chair of the Committee:	Prof. Dr. Holger Hermanns
Reporters	
First Reviewer:	Prof. Dr. Krishna P. Gummadi
Second Reviewer:	Prof. Dr. Karrie G. Karahalios
Third Reviewer:	Dr. Ingmar Weber
Academic Assistant:	Dr. Balakrishnan Chandrasekaran

ABSTRACT

An increasing number of people are relying on online social media platforms like Twitter and Facebook to consume news and information about the world around them. This change has led to a paradigm shift in the way news and information is exchanged in our society – from traditional mass media to online social media.

With the changing environment, it's essential to study the information consumption of social media users and to audit how automated algorithms (like search and recommendation systems) are modifying the information that social media users consume. In this thesis, we fulfill this high-level goal with a two-fold approach. First, we propose the concept of *information diets* as the composition of information produced or consumed. Next, we quantify the *diversity and bias in the information diets* that social media users consume via the three main consumption channels on social media platforms: (a) *word of mouth* channels that users curate for themselves by creating social links, (b) *recommendations* that platform providers give to the users, and (c) *search* systems that users use to find interesting information on these platforms. We measure the information diets of social media users along three different dimensions of topics, geographic sources, and political perspectives.

Our work is aimed at making social media users aware of the potential biases in their consumed diets, and at encouraging the development of novel mechanisms for mitigating the effects of these biases.

KURZDARSTELLUNG

Immer mehr Menschen verwenden soziale Medien, z.B. Twitter und Facebook, als Quelle für Nachrichten und Informationen aus ihrem Umfeld. Diese Entwicklung hat zu einem Paradigmenwechsel hinsichtlich der Art und Weise, wie Informationen und Nachrichten in unserer Gesellschaft ausgetauscht werden, geführt – weg von klassischen Massenmedien hin zu internetbasierten Sozialen Medien.

Angesichts dieser veränderten (Informations-) Umwelt ist es von entscheidender Bedeutung, den Informationskonsum von Social Media-Nutzern zu untersuchen und zu prüfen, wie automatisierte Algorithmen (z.B. Such- und Empfehlungssysteme) die Informationen verändern, die Social Media-Nutzer aufnehmen. In der vorliegenden Arbeit wird diese Aufgabenstellung wie folgt angegangen: Zunächst wird das Konzept der *"Information Diets"* eingeführt, das eine Zusammensetzung aus produzierten und konsumierten Social Media-Inhalten darstellt. Als nächstes werden *die Vielfalt und die Verzerrung (der sogenannte "Bias") der "Information Diets"* quantifiziert die Social Media-Nutzer über die drei hauptsächlichen Social Media- Kanäle konsumieren: (a) *persönliche Empfehlungen und Auswahlen*, die die Nutzer manuell pflegen und wodurch sie soziale Verbindungen (social links) erzeugen, (b) *Empfehlungen*, die dem Nutzer von der Social Media-Plattform bereitgestellt werden und (c) *Suchsysteme* der Plattform, die die Nutzer für ihren Informationsbedarf verwenden. Die "Information Diets" der Social Media-Nutzer werden hierbei anhand der drei Dimensionen Themen, geographische Lage und politische Ansichten gemessen.

Diese Arbeit zielt zum einen darauf ab, Social Media-Nutzer auf die möglichen Verzerrungen in ihrer "Information Diet" aufmerksam zu machen. Des Weiteren soll diese Arbeit auch dazu anregen, neuartige Mechanismen und Algorithmen zu entwickeln, um solche Verzerrungen abzuschwächen.

PUBLICATIONS

Parts of this thesis have appeared in the following publications.

- "Search Bias Quantification: Investigating Political bias in Social media and Web search". Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna Gummadi, Karrie Karahalios. Under review for 'Social Media for Personalization And Search' special issue of the 'Information Retrieval Journal' (Springer).

  (Full paper)

- "Quantifying Search Bias: Investigating Sources of Bias for Political Searches in Social Media". Juhi Kulshrestha, Motahhare Eslami, Johnnatan Messias, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna Gummadi, Karrie Karahalios. In *Proceedings of ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, Portland, USA, February 2017.

  (Full paper)

- "Measuring and Managing Information Diets of Social Media Users". Juhi Kulshrestha. CSCW Doctoral Colloquium, *In Companion - Proceedings of ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, San Francisco, USA, February 2016.

  (Doctoral Colloqium Extended Abstract)

- "Characterizing Information Diets of Social Media Users". Juhi Kulshrestha, Muhammad Bilal Zafar, Lisette Espin Noboa , Krishna P. Gummadi, Saptarshi Ghosh. In *Proceedings of International AAAI Conference on Web and Social Media (ICWSM)*, Oxford, UK, May 2015.

  (Full paper)

- "Evaluating Social Recommendation Systems". Juhi Kulshrestha, Muhammad Bilal Zafar, Saptarshi Ghosh, Krishna P. Gummadi. In *Poster Session, Conference on Online Social Networks (COSN)*, Dublin, Ireland, October 2014.

(Poster)

- "Geographic Dissection of the Twitter Network". Juhi Kulshrestha, Farshad Kooti, Ashkan Nikravesh, Krishna P. Gummadi. In *Proceedings of International AAAI Conference on Web and Social Media (ICWSM)*, Trinity College Dublin, Ireland, June 2012.

  (Full paper)

Additional publications during doctoral studies.

- "Media Bias Monitor: Quantifying Biases of Social Media News Outlets at Large-Scale". Filipe N. Ribeiro, Lucas Henrique, Fabricio Benevenuto, Abhijnan Chakraborty, Juhi Kulshrestha, Mahmoudreza Babei, Krishna P. Gummadi. In *Proceedings of AAAI International Conference on Web and Social Media (ICWSM)*, Stanford, USA, June 2018.

  (Full paper)

- "Purple Feed: Identifying High Consensus News Posts on Social Media". Mahmoudreza Babaei, Juhi Kulshrestha, Abhijnan Chakraborty, Fabricio Benevenuto, Krishna P. Gummadi, Adrian Weller. In *Proceedings of AAAI/ACM Conference on Artificial Intelligence, Ethics & Society (AIES)*, New Orleans, USA, February 2018.

  (Full paper)

- "The Road to Popularity: The Dilution of Growing Audience on Twitter". Przemyslaw A. Grabowicz, Mahmoudreza Babaei, Juhi Kulshrestha, Ingmar Weber. In *Proceedings of International AAAI Conference on Web and Social Media (ICWSM)*, Cologne, Germany, May 2016.
  (Short paper)

- "Deep Twitter Diving: Exploring Topical Groups in Microblogs at Scale". Parantapa Bhattacharya, Saptarshi Ghosh, Muhammad Bilal Zafar, Juhi Kulshrestha, Mainack Mondal, Niloy Ganguly, Krishna P. Gummadi. In *Proceedings of ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW)*, Baltimore, USA, February 2014.

  (Full paper)

Dedicated to my family, especially Baba-Amma.

ACKNOWLEDGEMENTS

I feel incredibly fortunate to have had a fantastic support system during my graduate life, and there are many people I would like to thank for helping and supporting me along the way.

First of all, I would like to thank my adviser, Krishna Gummadi, for challenging me when I needed the push and supporting me when I needed the encouragement. I have learned so much from him, not just about doing research but also about having clear priorities in life and working towards them. I am also grateful to Karrie Karahalios and Ingmar Weber for their valuable guidance and support.

I had the privilege of collaborating with some remarkable researchers during my graduate studies and this PhD could not have been successful without my amazing co-authors and collaborators. In particular, I would like to thank Saptarshi Ghosh for being a fantastic mentor and a friend, Motahhare Eslami for challenging and supporting me throughout our collaborations, Farshad Kooti for showing me the ropes when I was just starting out, Johnnatan Messias for always being enthusiastic and proactive, Lisette Espin for her excitement and hard work, and Abhijnan Chakraborty for the endless whiteboard discussions. I would also like to thank Fabricio Benevenuto and Niloy Ganguly for their advice and encouragement.

I am thankful to the current and former members of MPI-SWS, especially, Bimal, Pramod, Ezgi, Riju, Anjo, Oana, Aastha, Arpan, Jisun, Eslam, Reinhard, Nina, and Till, for making my time here exciting and fun. Particularly, I want to thank Mainack for the numerous introspective discussions we had, Reza for being a great teammate and friend, Manohar for being the best friend I could have asked for and Bilal for helping me keep my sanity in crazy times and always cheering me up with his jokes. My experience at MPI would not have been the same without them.

I am grateful to all the technical support staff at MPI-SWS, and in particular, Carina, and Chris for helping me till the very last minute. I would also like to thank Mary-Lou for providing all the information regarding submitting the dissertation and scheduling the defense, and Rose for helping me communicate my research more effectively. I can not thank Claudia Richter, Brigitta Hansen, and

Annika Meiser enough for always looking out for me and making my life at MPI fun, relaxed and comfortable.

A special thanks also goes to Michelle Carnell, Holger Hermanns and Raimund Seidel for helping me get a great start at the graduate school, mentoring me along the way and for helping me make the most of the opportunities I was given.

Last but not the least, I would like to thank my family. They have always been my strength in tough times, and I am immensely grateful to them – Mumma-Papa for always inspiring me to be the best version of myself, Ila for believing in me and keeping me sane and motivated, Baba for teaching to always keep faith, Mom-Dad for trusting me and encouraging me to go the extra mile, and Aparna for managing to lighten my mood and keeping me calm. Prabhav, Nanda and Manohar have been my family away from home and without their unwavering support, I could not have come out successful. Esha, Sarang, and Akriti have been there for me for nearly all of my life, and the distance did not change that one bit. Finally, I couldn't be more thankful to Mittul for continually believing in me and motivating me to keep going, but most of all, for not giving up on me even when I was ready to give up.

TABLE OF CONTENTS

LIST OF TABLES xvii

LIST OF FIGURES xix

1 Introduction 1

1.1 Rising popularity of online social media as a source of news and information 1

1.2 Challenges for studying the paradigm shift from mass media to online social media 2

1.3 Thesis Research: Quantifying & Characterizing Information Diets of Social Media Users 3

1.4 Background & related work 4

1.4.1 Information consumption via mass media 4

1.4.2 Information production and consumption on social media 5

1.4.3 Role of retrieval algorithms in shaping information consumption 6

1.5 Overview of thesis contributions 7

1.5.1 Information diets - The concept 7

1.5.2 Word of mouth: Information diets of social media users 8

1.5.2.1 Topical diversity in information diets: Social media vs. mass media 8

1.5.2.2 Geographical source diversity of information diets 9

1.5.3 Recommendations: Impact on diversity in consumption 10

1.5.4 Search: Quantifying the bias in social media search 11

1.6 Organization of the thesis 12

2 Information diets: The concept 14

2.1 Information diets – A descriptive metric 15

2.2 Prior work related to information diets 16
2.2.1 Normative approach to information diets 16
2.2.2 Journalism and communication's repertoire approach to information diets 17
2.2.2.1 Within-platform repertoires 17
2.2.2.2 Cross-platform repertoires 17
2.3 Limitations & future directions 18
3 Word of mouth: Information diets of social media users 20
3.1 Topical diversity in information diets: Social media vs. mass media 21
3.1.1 Related work 22
3.1.1.1 Inferring topics of social media posts 22
3.1.1.2 Assigning topics to social media users 23
3.1.2 Methodology: Quantifying topical information diets 24
3.1.2.1 Inferring topic of a keyword 24
3.1.2.2 Constructing information diets from topics of keywords 30
3.1.2.3 Public deployment of information diets 30
3.1.2.4 Limitations 31
3.1.3 Production: How are the diets produced on social media and mass media different? 32
3.1.3.1 News organizations: Social media vs. mass media 32
3.1.3.2 Popular social media accounts vs. mass media 36
3.1.3.3 Random sampling of social media vs. mass media posts 38
3.1.4 Consumption: Are the social media users consuming balanced diets? 39
3.1.5 Summary: Topical diversity in information diets of social media users 42
3.2 Geographical source diversity of information diets 42
3.2.1 Related work 43
3.2.1.1 Inferring geolocations of social media users 43
3.2.1.2 Analyzing the impact of geography and language on online user interactions 44

3.2.2 Methodology: Quantifying geographical information diets ... 44
3.2.2.1 Twitter dataset ... 44
3.2.2.2 Inferring users' geolocations ... 45
3.2.2.3 Evaluating geolocation inference methodology ... 46
3.2.2.4 Constructing information diets from source geolocations ... 46
3.2.2.5 Limitations ... 47
3.2.3 Production: Are the sources geographically diverse? ... 47
3.2.3.1 Geo-distribution of Twitter users ... 47
3.2.3.2 Geo-distribution of elite Twitter users ... 48
3.2.4 Consumption: Is there geographic source diversity in the diets that different countries consume? ... 49
3.2.4.1 Geography of consumption channels ... 49
3.2.4.2 Geographic source diversity in consumed diets of countries ... 54
3.2.5 Summary: Geographical source diversity of information diets ... 55
3.3 Conclusion ... 56
4 Recommendations: Impact on diversity in consumption ... 57
4.1 Background & related work ... 59
4.1.1 Background: Twitter's personalized recommendations ... 59
4.1.2 Evaluation metrics for recommender systems ... 60
4.1.3 Impact of recommendations on diversity ... 60
4.2 Data collection & methodology ... 62
4.2.1 Experiment design ... 63
4.2.2 Data collected to construct recommended diets ... 64
4.2.3 Limitations & ethical considerations ... 67
4.3 Recommended diets vs. consumed diets ... 67
4.3.1 Are the recommendations personalized for each user? ... 68
4.3.2 To what extent do the recommendations given to a user match their consumed diet? ... 69

4.3.3 Do the recommendations mitigate or exacerbate the imbalances in the users' consumed diets? 70

4.4 Conclusion 72

5 Search: Quantifying political bias in social media and Web search 73

5.1 Related work 76

5.1.1 Measuring political bias 76

5.1.1.1 Measuring political bias in news media 77

5.1.1.2 Measuring political bias of social media users 78

5.1.2 Characterizing bias on social media 80

5.1.3 Characterizing bias in Web search systems 81

5.2 Search bias quantification framework 84

5.2.1 Bias score of an individual data item 85

5.2.2 Input Bias 86

5.2.3 Output Bias 86

5.2.4 Ranking Bias 87

5.2.5 Time-averaged Bias 88

5.3 Investigating sources of bias for political searches on social media 88

5.3.1 Collecting Twitter search data 88

5.3.1.1 Selecting search queries 88

5.3.1.2 Data collection from Twitter 90

5.3.2 Measuring political bias of an individual search result 91

5.3.2.1 Source bias - Inferring political bias of Twitter users 91

5.3.2.2 Public deployment of the source bias inference methodology 93

5.3.2.3 Evaluation of political bias inference methodology 93

5.3.2.4 Match between source bias & tweet bias 97

5.3.3 Characterizing the bias for political searches on Twitter social media 99

5.3.3.1 Where does the bias come from – Input data or ranking system? 99

5.3.3.2 Collective contribution of the input data and the ranking system 103

5.3.3.3 Variation of bias over time ... 106

5.4 Comparing relative bias in political searches on the Web and social media ... 108

5.4.1 Query selection & data collection ... 109

5.4.1.1 Collecting Google Web search data ... 109

5.4.1.2 Collecting Twitter news search data ... 109

5.4.2 Measuring political bias of a search result ... 109

5.4.2.1 Measuring bias of Google search results ... 110

5.4.2.2 Measuring bias of Twitter news search results ... 110

5.4.3 Comparing relative biases of Google search and Twitter news search ... 111

5.4.3.1 Temporal variation in search bias ... 111

5.4.3.2 Higher Democratic bias in Twitter news search results ... 112

5.4.3.3 Favorable bias on Google search via candidate-controlled sources ... 112

5.5 Comparing Relative Bias of Twitter's Different Ranking Systems ... 114

5.6 Generalizability of the search bias quantification framework ... 114

5.6.1 Extending to multiple perspectives scenario ... 115

5.6.2 Extending to limited data availability scenario ... 115

5.6.3 Extending to other search systems ... 116

5.7 Signaling bias in search results ... 116

5.8 Limitations & future directions ... 118

5.9 Conclusion ... 119

6 Concluding discussion ... 121

6.1 Future research directions ... 123

Appendices ... 126

A Screenshots & videos demonstrating the functionalities of publicly deployed Twitter applications ... 127

A.1 Making users aware of their information diets ... 127

A.2 Inferring political leaning of Twitter users ... 129

A.3 Twitter search split by the political leaning 129

BIBLIOGRAPHY 133

LIST OF TABLES

3.1 Number of nodes at different levels of the topical hierarchy. 25

3.2 Topics in the top two levels of our topical hierarchy. 26

3.3 Sample topical experts along with a sample of their expertise List-tags and the inferred topic (out of the 18 topical categories). 27

3.4 The 18 topic categories to which keywords and tweets are mapped, and some terms related for each topic. The terms are matched with expertise-tags of Twitter users, to map expert users to different topics. 27

3.5 Comparing the proposed author-based crowdsourced topic inference methodology with AlchemyAPI (which uses content-based NLP techniques) in terms of coverage and accuracy. 29

3.6 Mass media information diets of three news organizations, where the topics of the news-articles were judged by AMT workers (top topics highlighted). 33

3.7 Examples of topic-specific Twitter accounts of news organisations, along with the contribution of their topics of specialization in their produced diet. 34

3.8 Match between the different sources for geolocation resolution. 45

3.9 How much do countries rely on other countries for the information they consume?: The fraction of trans- and intra-national following links for the 10 countries with the most users (ranked by their fraction of transnational followings). . 50

3.10 Closest 5 follower (who seek information from them) and following (from whom they seek information) countries for a few example countries around the world. 52

3.11 Percentage of closest follower and following country pairs that share a geographic boundary, or a common language, or lie within the same continent. 53

3.12 Contribution of other countries in the consumed diets of the 10 countries with the most users. 55

4.1 Characteristics of the 15 test accounts created, which mimic the followings of 15 real Twitter users. 65

4.2 Range of contributions of different topics in the recommended diets given to the test accounts. 68

5.1 Explaining the bias metrics. 85

5.2 Confusion matrix of the match between AMT bias scores and Inferred bias scores 95

5.3 Coverage and accuracy of the political bias inference methodology for (i) current US senators, and (ii) common users who have declared their political ideology in their Twitter account profiles. . . . 96

5.4 Confusion matrix for source bias classification – gold standard tweet bias (based on AMT workers' judgement) vs. source bias. . . . 97

5.5 Confusion matrix for content bias classification (Support Vector Machine (SVM) classifier) – gold standard tweet bias (based on AMT workers' judgement) vs. content bias. . . . 98

5.6 Confusion matrix for content bias classification (Gradient boosted decision tree (GBDT) classifier) – gold standard tweet bias (based on AMT workers' judgement) vs. content bias. . . . 98

5.7 Time-averaged bias in Twitter search "top" results, for selected queries (related to political candidates and debates) – output bias TOB, input bias TIB, and ranking bias TRB. . . . 100

5.8 Time-averaged ranking bias for different ranking strategies: (i) Twitter's ranking (Twitter search "top" results), (ii) Most retweeted tweet first ranking, and (iii) Most favorited tweet first ranking. . . . 102

5.9 Randomly selected tweets from the search results for the queries *Hillary Clinton* and *Donald Trump*, which are posted by a user with an opposite bias as compared to the candidate. . . . 105

5.10 Statistical analysis of temporal variation of output bias for the query *dem debate* – (i) variation in output bias across 3 hours before and 3 hours after the Republican debate on December 15, 2015, (ii) variation in output bias across 3 hours before and 3 hours after the Democratic debate on December 19, 2015, (iii) variation in output bias across the 3 hour time periods during the Republican debate (on December 15, 2015) and the Democratic debate (on December 19, 2015) . . . 108

5.11 Comparing time-averaged output bias TOB in (i) Google search results, (ii) Twitter news search results. . . . 113

LIST OF FIGURES

3.1 The different steps of our author-based crowdsourced methodology for inferring the topic of a keyword. 28

3.2 Constructing information diets by extracting keywords from each tweet and inferring it's topic and then aggregating and normalizing the contributions of all keywords, such that the weights in the final topic-vector sum up to one. A sample tweet is also depicted to exemplify the process of generating the topic-vector from the tweets. 30

3.3 Screenshots of our Twitter application for making users aware of their information diets. 31

3.4 Comparing the information diet posted by news organizations in mass media (news articles in print editions) and social media (tweets posted by their primary Twitter accounts) for the same days in December 2014. 35

3.5 Mean contribution of the top topic (on which a user posts the highest fraction of their diet) in the produced diets of popular users, grouped according to their top topic of production. 36

3.6 Information diets produced by four selected verified users – @darrenrovell (ESPN Sports Business Reporter, ABC News Business Contributor.), @drsanjaygupta (Staff Neurosurgeon, Emory Clinic; CNN Chief Medical Correspondent), @maddow (American television host and political commentator), @jamieoliver (British celebrity chef and restaurateur). All of them produce focused diets on their topics of expertise. 37

3.7 Distribution of the 500 randomly selected verified users, according to the topic on which they produce the maximum fraction of their diet. 38

3.8 Comparing the information diet of the Twitter 1% random-sample with the mass media diet of news organizations (NYTimes and Washington Post). 38

3.9 Mean contribution of the top topic in the consumption diets (on which a user consumes the highest fraction of their diet) of random users grouped according to their top topic of consumption. 40

3.10 Mean contribution of the bottom 12 least dominant topics in the consumption diets of random users grouped according to their top topic of consumption. 40

3.11 Distribution of the 500 randomly selected users, according to the topic on which they consume the maximum fraction of their diet. 41

3.12 Distribution of Twitter population and world population of 20 countries with the most Twitter users. 47

3.13 Distribution of number of elites in the 20 countries with the most users. 49

3.14 Fraction of intra-national followers and followings in comparison to the Twitter user population share in different countries, ranked by their Twitter user population. 51

3.15 Groups of countries whose users are closely connected with one another via consumption channels. 53

4.1 Percentage of common items in the set of tweets recommended to two exactly identical accounts at the same instant of time 64

4.2 Over-representation of tweets that are tweeted/retweeted by at least ten users within u's 2-hop neighborhood, in the recommendations provided by Twitter to u. 66

4.3 Comparing the consumed diet and recommended diet of two test accounts – (i) the one with the minimum KL divergence, and (ii) the one with the maximum KL divergence of the recommended diet from the consumed diet. 70

4.4 The contribution of the (i) top 3 consumed topics and (ii) bottom 12 consumed topics in the consumed, recommended and combined diets for the test accounts. 71

4.5 KL divergence of the consumed and combined diets of test accounts from the baseline diets of (i) NYTimes mass media diet, and (ii) Twitter's 1% random sample diet. 71

5.1 Overview of our search bias quantification framework. 84

5.2 Screenshots of our Twitter application for inferring political leaning of users showing the inferred political leaning for me (the logged-in user) and 'Barack Obama'. 93

5.3 CDF of AMT bias scores and Inferred bias scores for politically interested common users. 95

5.4 The time-averaged output bias TOB in Twitter "top" search results for the Republican candidates – candidates are listed left to right from highest to lowest popularity. 104

5.5 Temporal variation of output and input bias for the query *dem debate* – (a) variation across the full duration over which we collected data (December 14–22, 2015), (b) variation during a 9-hour window around the Republican debate on December 15, 2015, (c) variation during a 9-hour window around the Democratic debate on December 19, 2015. 106

5.6 Screenshot of our Twitter-based split search service showing the results for the search term 'abortion'. A widget adjacent to each result shows its bias measure. . . . 118

A.1 Screenshots of our Twitter application for making users aware of their information diets. 128

A.2 Screenshots of our Twitter application for inferring political leaning of users. 130

A.3 Screenshots of our Twitter-based split search service showing the results for the search terms ‘guns’ and ‘abortion’. A widget adjacent to each result shows its political bias measure. 131

CHAPTER 1

Introduction

1.1 Rising popularity of online social media as a source of news and information

Traditionally, people relied on *mass media* organizations to acquire news and information. These mass media organizations used broadcast media like print (*e.g.,* NYTimes or Economist), radio (*e.g.,* NPR, BBC radio), or television (*e.g.,* CNN, ESPN) to disseminate news and information to a large number of people. However, in the last couple of decades, the financial fortunes and number of subscribers of the traditional mass media organizations have been on the decline. The US daily newspaper circulation in 2017 went down by about 11% from the previous year [114]. Only a minority of people in the US get news from radio (25%) and print newspapers (18%) [112], and the gap between the US adults who get news online (43%) and the ones who get it via television (50%) has also been narrowing consistently, and went down to 7% in 2017 from 19% in 2016 [112].

Today, an increasing number of people are using *online social media* sites like Twitter, Facebook, and Youtube to get information on recent events and topics of their interest and these social media platforms have become curators and gateways to news and information [136, 152, 226] in the society. A 2016 survey conducted by Pew Research Center estimated that about 62% of the US adults consume news primarily from social media sites [152], and this fraction increased to 67% by 2017 [113]. Additionally, in 2013 about 50% of Twitter users reported that they got news on the platform, while this number increased to 74% by 2017 [113].

With people relying more and more on social media sites for news and information rather than traditional mass media, there has been a profound change in the way that information is produced and consumed in our society. Next, we briefly discuss the challenges that this paradigm shift poses for studying the news and information consumption in the society.

1.2 Challenges for studying the paradigm shift from mass media to online social media

The paradigm shift in the information production and consumption behavior in our society – from traditional mass media to online social media – has led to many fundamental changes [148, 206]. Here we briefly enumerate three crucial differences between these two kinds of media.

Firstly, in traditional mass media, there are typically a small number of news organizations which produce news and information that millions of users consume. In contrast, the millions of users of online social media are not just passive consumers of information. Instead, they actively publish content, contribute their knowledge and expertise and spread ideas over these social media sites.

Secondly, while mass media organizations use the *broadcast channels* of communication (*i.e.,* all subscribers receive the same content), individual users on online social media consume information via *personalized channels*. Every individual user selects (*e.g.,* by establishing social links) their preferred sources of information from the millions of individual producers.

And finally, most social media platforms also provide *personalized information retrieval systems* like recommendation [10, 7] and search systems [1, 230] to their users, which form an additional channel of information. Therefore, an individual social media user might receive information that is very different from what other users on the platform receive, making it very challenging to study them at scale.

Thus, a user on online social media typically has three channels of consuming information – (i) via *word of mouth*, *i.e.,* the content posted by their social contacts, (ii) the content that they receive as *recommendations* from the platform provider, and (iii) the content that they *search* for on the social media. On one hand, users may choose to focus their attention selectively by connecting to only a few sources that reflect their preferences and topics of interest (also known as selective exposure [208, 219, 111, 75]) leading to a lack of diversity in the information being consumed via word of mouth channels and formation of echo chambers [23, 75] around themselves. While on the other hand, the retrieval algorithms (like recommender and search systems) also exert an influence on the information users consume, and these personalized algorithms could lead to undesirable outcomes like users being trapped in filter bubbles [23, 75, 177] where their preexisting preferences may keep getting reinforced by these algorithms. Therefore, it is essential to study the bias and diversity in the information that social media users are consuming via their personalized word of mouth channels, and how the algorithmic channels further personalize and impact the information they are exposed to.

1.3 Thesis Research: Quantifying & Characterizing Information Diets of Social Media Users

The high-level goal of this thesis is two-fold: (i) proposing *information diets*, a conceptual construct to measure and reason about the bias and diversity of the information being produced and consumed by social media users, and (ii) quantifying the diversity and bias in the information diets of social media users via the three main channels of consumption on social media: (a) *word of mouth* channels that the users choose for themselves by creating social links, (b) *recommendations* provided by the social media platform, and (c) the *search* systems on these platforms that the users use to find information of interest to them. We leverage our findings to propose systems that help make the users aware of the biases in their diets.

Based on our high-level goal, the central question we address in this thesis is 'what is the bias and diversity in the word of mouth information diets of social media users and how are they impacted by algorithmic channels of recommendation and search?'. Empirical measurements of word of mouth diets of social media users performed in this thesis reveal that both produced and consumed diets are topically and geographically focussed and exhibit low diversity. Furthermore, controlled experiments to audit the impact of algorithmic channels show that social recommendations mitigate the topical imbalances in social media users' consumed diets by adding topical diversity; while search bias presented to a user originates both from the user-generated data input to the ranking system as well as the ranking system itself.

In the rest of the chapter, we will first briefly review the related work and then give a high-level overview of the specific contributions of this thesis.

1.4 Background & related work

More and more people are relying on the online social media to discover information about the world around them [136]. Traditional mass media like newspapers, radio, and television, are rapidly being replaced by social media sites like Facebook or Twitter. With this paradigm shift in news and information consumption, an increasing number of users are using these social media platforms to exchange information on recent events and topics of their interest. In this thesis, we define the concept of information diets to study the bias and diversity in the information consumed by social media users via the self-curated word of mouth channels, as well as, algorithmically curated channels of recommendations and search. Being aware of their diets can help social media users to selectively consume information in a more balanced and healthy way [110].

In this section, we begin by building upon this context to provide background on the work that has been done on studying the bias and diversity in information disseminated via mass media and social media. We finish with a brief overview of prior work in the area of algorithm auditing, with a focus on the impact of the retrieval algorithms on the bias and diversity in the information consumed by social media users.

1.4.1 Information consumption via mass media

The active field of media studies has long been responsible for analyzing the content coverage of mass media and its effects on the society. There are many 'media watchdog organizations' (*e.g.,* FAIR (**http://fair.org/**), AIM (**http://www.aim.org/**)) which judge the content covered by news organizations based on fairness, balance, and accuracy of coverage.

A number of prior studies have investigated political bias in traditional news media [41, 79, 86, 161]. For instance, Groseclose & Milyo [86] linked media sources to members of the US Congress based on their co-citation of think tanks, and then assigned political bias scores to media sources based on the ADA scores of these Congress members.[1] Gentzkow & Shapiro [79] determined the similarity of the language used by a media source to the language of congressional Republicans or Democrats to infer the 'media slant'. Whereas, Budak *et al.* [41] combined machine-learning and crowdsourcing techniques to study the selection and framing of political issues by news organizations. As online news

[1]Americans for Democratic Action: **www.adaction.org**.

sources have gained popularity, such studies have also been extended to them, as in the case of the Balance study [161], which assigns political bias scores to many of the popular news websites based on the political leanings of the websites, blogs and Digg users that link to or vote for the news website.

It is easier to perform these studies on mass media because of its broadcast nature ensuring that all the users receive the same information. The task of studying information consumption on social media is more challenging than studying mass media since each user shapes their own self-curated channel of consumption by following a chosen set of users. In Chapter 3, we examine the word of mouth diets of social media users and compare them with the mass media diets of well-known news publications. Additionally, algorithmic channels like search and recommendation further add a layer of personalization, making it much harder to perform such studies at scale. Next, we briefly describe prior attempts at studying the information exchanged on social media.

1.4.2 Information production and consumption on social media

With the ever-increasing popularity of social media sites like Twitter and Facebook, more and more users are relying on these platforms to obtain news [136], real-time information about ongoing events and crowd opinion on public figures [226].

A significant portion of prior work on studying information production and consumption on social media [44, 131, 250] has been focused on examining the *amount* of information being exchanged between different groups of users.

However, there has been a limited effort towards analyzing the *composition* of information produced or consumed on social media along different dimensions such as topics or geography of sources. Some prior studies have analyzed the differences in the link creation and usage patterns of social media users from different geographic regions [109, 123, 222] and language communities [104, 117]. Ramage *et al*. [194] used topic models to characterize users and content on Twitter along the dimensions of substance, status, style and social characteristics of posts. In this thesis, we study topical and geographic source diversity not only of the diets that social media users are producing or consuming (Chapter 3), but also analyze whether personalized recommendations exacerbate or mitigate the topical imbalances in their diets (Chapter 4).

On the other hand, there has been substantial work on investigating political polarization on social media and whether users receive multiple perspectives on a specific event or topic [14, 35, 53], including studying the cross-ideological interaction of social media users [103, 140]. These studies show that political talk on Twitter is highly partisan and users are unlikely to be exposed to cross-ideological content through their friendship network. In this thesis (Chapter 5), we focus on examining the bias in the diets that users consume via political queries on social media and web search.

1.4.3 Role of retrieval algorithms in shaping information consumption

Algorithms have become ubiquitous in curating and presenting information to users on online platforms [228]. These algorithms affect users' online experience significantly, sometimes even in undesirable manners – for instance, by creating discriminatory ads based on gender [57] or race [221], or showing different prices for the same products/services to different users [91]. Such issues have lead researchers, organizations, and governments to pursue a new line of research of *auditing algorithms* which attempts to analyze if and how algorithmic systems cause biases, especially when they end up misleading or discriminating against some users [69, 203, 204, 214].

We are particularly interested in the subclass of the algorithm audit studies which study the role that information retrieval algorithms like recommendation and search systems play in shaping the information that social media users are consuming and next we give a brief overview of such studies.

In recent years significant attention has been paid to algorithms that filter, rank and personalize content on the web, with studies focusing on investigating the influence of recommendation systems and their undesirable outcomes like social media users getting trapped in echo chambers and filter bubbles that limit their exposure to ideologically cross-cutting information [23, 38, 75, 177]. Several studies [66, 155] have examined how the personalized recommendations impact the diversity in the information that users are exposed to.

Lately, there has also been a growing interest in studying the bias in web search engine results [76, 159, 237, 107, 85]. Much of this work focuses on the politics of search engines and tries to examine if dominant search engines like Google favorably ranks certain websites over the others.

A parallel line of prior work [90, 119, 199, 38] has focussed on studying the biases in web search due to personalization, *i.e.,* the differences in the search results seen by different users for the same query. While another set of prior work examined how search systems exaggerate stereotypes and propagate gender [115] and racial biases [172]. Finally, several studies have also investigated the political bias in web search queries [241] and results, with Epstein & Robertson [63] conducting a field study which shows that the voting preferences of undecided voters in an election can be affected up to 20% by manipulating and biasing the search results. These studies highlight the considerable impact biased search rankings can have on the political opinions of users, motivating us to study the biases in the information diets that users are consuming via the search on social media.

Given this background, in this thesis, we define the concept of information diets (Chapter 2) which we use to quantify and characterize the diversity and bias in the information that users consume via the main channels of information consumption on social media – user-curated word of mouth channels (Chapter 3), and algorithmically curated channels of recommendations (Chapter 4) and search (Chapter 5).

1.5 Overview of thesis contributions

In this section, we provide a brief overview of our thesis contributions. Our contributions can be divided into two broad parts. First, we propose the concept of information diets of social media users. Second, having defined the idea of information diets, we focus on measuring users' information diets via three main channels of consumption on social media — (i) word of mouth, (ii) recommendations, and (iii) search. In the rest of this section, we describe our methodology for quantifying the diversity and bias in the users' consumption via these three channels in brief, and the insights we gain from these analyses.

1.5.1 Information diets - The concept [127]

The widespread adoption of social media platforms like Twitter and Facebook has lead to a paradigm shift in the way our society is producing and consuming information – from the broadcast mass media to online social media. To study the effects of this paradigm shift, we define the concept of information diet as the composition of a set of information items being produced or consumed. Information diets can be constructed along many dimensions like *topics* (*e.g.,* politics, sports, science *etc.*), or *perspectives* (*e.g.,* politically left leaning or right leaning), or *geography of sources* (*e.g.,* information published from different parts of the world).[2] We use the descriptive metric of information diets to measure the diversity and bias in the information produced or consumed on online social media and to study how the automated retrieval algorithms like recommendation and search systems, that are provided by most

[2]We use the terms 'information diets' and 'diets' interchangeably in the rest of the thesis.

social media platforms, are shaping the diets of social media users. We use Twitter as a substrate to study the information diets of social media users. We leverage the insights we gain from analyzing social media users' diets to reason about better information discovery and exchange systems over social media.

1.5.2 Word of mouth: Information diets of social media users [125, 127]

To study and characterize information diets of social media users, we begin by examining the most organic way of information exchange on social media, *i.e.,* via the social links that users create by following or friending other users. We investigate the word of mouth diets of social media users along two dimensions – *topical diversity* and *geographical source diversity* – which we briefly describe next.

1.5.2.1 Topical diversity in information diets: Social media vs. mass media [127]

We begin by applying the concept of information diets to study the topical diversity in the information produced and consumed by users on a popular social media platform – Twitter. In other words, we measure the fraction of the information that users are producing and consuming in different topical categories (*e.g.,* information on politics, sports, entertainment *etc.*). We then proceed to examine the effect of the paradigm shift from mass media to online social media by contrasting these social media diets with the topical diets of mainstream mass media organizations.

The critical challenge for constructing these topical information diets is the lack of a reliable methodology for inferring the topics of short social media posts, which also scales up to work for millions of tweets. More traditional content-based schemes do not perform well for social media posts due to the posts' short length, informal language and fast evolving vocabulary. So we propose a novel *author based crowdsourced technique to infer the topics of social media posts*, which relies on the topical characteristics of the users who are discussing a piece of information. The basic intuition behind our methodology is that if many users interested in a particular topic are discussing a specific information piece, then that information is most likely related to that topic. For instance, if multiple politicians or political journalists are posting a particular keyword, then that keyword is very likely to be related to the topic of politics. We show that our method for topic inference works better for social media posts than a state-of-the-art content based tool – AlchemyAPI.[3]

Our findings indicate that the social media users are exceptionally topically focused in their production and consumption behaviors. We observe an unbundling of content creation on social media where each account (including accounts of news organizations) produces a specialized diet on a particular topic, with 20% to 50% of their production focused on just one topic. In contrast, the subscribers of these news organizations would get a much more topically balanced diet from their mass media editions. Moreover, we also observe that the consumed diets of the social media users are even more skewed with just one topic contributing at least 50% of the consumption for more than 50% of Twitter users.

System for measuring information diets: We also developed a system to make users aware of the diets they are consuming, in the form of a Twitter app [106] (deployed at **http://twitter-app.mpi-sws.org/information-diets/**). Using this service, the users can check the diet they are consuming on Twitter, as well as search for other users and see the diets they are producing. Additionally, they can also use this service to selectively read the tweets in their consumed diet from different topics.

1.5.2.2 Geographical source diversity of information diets [125]

Having explored the topical information diets, we turn our attention to another dimension of studying social media users' diets, namely geographic source diversity. In this case, we consider a country's diet as a whole, which is constructed from all the tweets consumed by all the users of that country. The questions that we want to answer are "what is the geographic diversity of the sources of information that the different countries are consuming from, and are there any similarities between the country under consideration and the countries who are the largest contributors to its consumed diet?".

To construct the geographical diets of the different countries, we need to infer the geolocations for a large number of Twitter users at the granularity of countries. For making this inference, we make use of two profile fields of the users: the location field (free text string), and the timezone field (location name + UTC offset). We use two map APIs (Bing maps and Yahoo maps) to resolve the location field

[3]See **https://en.wikipedia.org/wiki/AlchemyAPI**.

string into a country, and also convert the timezones to the corresponding countries and only retain those users for whom two of these three sources match in the resolved locations.

Our results show that on an average across all the countries' diets, the country itself contributes about two-thirds of the diet it consumes, indicating that a considerable amount of consumed tweets are produced locally within the same country. A non-trivial third of the tweets are produced internationally and cross national boundaries to be consumed in other countries. Moreover, when we investigate the highest contributors to the consumed diet of a country (excluding the country itself), we find that many of them share geographic proximity or linguistic similarity with the country, emphasizing the role that these offline characteristics play in shaping the online information diets on social media.

1.5.3 Recommendations: Impact on diversity in consumption [126, 127]

Most online social media platform providers deploy automated recommendation systems to help their users discover interesting information out of the deluge of content being generated and shared on these platforms. Often these algorithms are also personalized to individual users making it even harder to study them at scale. In this part of the thesis, we focus on quantifying the impact of these algorithms on the diversity in the consumed diets of social media users.

What makes our job especially challenging is the fact that the social recommendations provided by Twitter [88] are personalized for each user and thus are unavailable to us directly. We take the alternative approach of setting up carefully controlled experiments where we create test Twitter accounts which mimic real Twitter users by following the same set of users as the real user. For these test accounts, we could measure the topical diversity in both the word-of-mouth consumed diet (constructed using tweets they receive from the users they follow) as well as the recommended diet (constructed using tweets recommended to them). Therefore we could study the impact of recommendations on the consumed diets of users by answering the question: "do personalized recommendations provided by Twitter mitigate or exacerbate the imbalances in the users' consumed diets?". We surprisingly found that the social recommendations provided by Twitter do end up adding some topical diversity to the information being consumed by most of our test accounts directly via word of mouth.

1.5.4 Search: Quantifying the bias in social media search [124]

Search systems on the Web as well as online social media are frequently used to find information about ongoing events and public figures. These form another vital channel of information consumption which impacts the diets of users. The results returned by these search engines, especially the top-ranked results, have been shown to affect the opinion formed by the users about the topic (*e.g.,* an event or person) being searched [63]. When there are multiple competing perspectives on a topic, such as a political event or political candidate, bias in the top search results can play an essential role in shaping public opinion towards (or away from) certain perspectives.

In the last part of the thesis, we propose a novel generalizable search bias quantification framework, which not only measures the amount of search bias but also decouples the bias from different components of the search system to identify the sources of bias. We then conduct two case studies in the context of political searches about 2016 US Presidential primaries, which highlight the advantages of using our search bias quantification framework in two different use case scenarios.

First, we apply the framework to study the sources of bias in political searches on Twitter social media. To use our search bias quantification framework, we need a methodology to measure the political bias of an individual search result – a tweet in this case. We operationalize the political bias of a tweet as its source bias, *i.e.,* we approximate the bias of the tweet as the political bias of the author of the tweet. We have developed a scalable and accurate *crowdsourced methodology for inferring the political bias of a Twitter user* u based on (i) inferring the topical interests of u based on the users whom u follows, and (ii) examining how closely u's interests match the interests of two representative sets of users, one set comprising of users who are known to have a democratic bias, and the other set consisting of users who are known to have republican bias. We show that the bias in the search results does not only originate from the ranking system, but the input data (that is input to the ranking system) is also a significant contributor to the overall search bias. Moreover, we observe that the top Twitter search results display varying degrees of political bias. This variation in bias depends on several aspects, such as the topic (event/person) being searched for, the exact phrasing of the query (even for semantically similar queries), and also the time at which the query is issued.

Second, we use our search bias quantification framework to compare the relative bias of two different search systems - Twitter social media search and Google web search. Our analysis shows three interesting ways in which the search bias for political queries on Google web search differs from that for Twitter social media search: (i) first, when we investigate the temporal dynamics of the bias in the search results on the two systems, we find the bias in the social media search results to be significantly more dynamic over time, (ii) next, when we compare their time-averaged output bias values to capture the overall trend, we observe that for Google search the bias for most queries matches the leaning of the person or event being queried for, while the bias of Twitter news search for most queries is democratic leaning, and (iii) finally, we notice that on Google search, a much higher fraction of search results are candidate controlled sources (*e.g.,* candidate's website or social media accounts), leading to more favorable results for the candidates on web search than on social media search.

Our work is aimed towards making users aware of the potential biases of social media search and how it compares with the biases in web search, and towards encouraging the development of novel mechanisms for presenting search results which could represent multiple competing perspectives on the same event or person.

System to infer political bias of Twitter users: We have also developed a Twitter app which lets users log in with their Twitter credentials and see their inferred political leaning, as well as, search for other users to view their inferred political leanings [184] (deployed at **http://twitter-app.mpi-sws.org/search-political-bias-of-users/**).

1.6 Organization of the thesis

The rest of the thesis is organized as follows:

In Chapter 2, we introduce the concept of information diets and briefly discuss how it can be applied along the different dimensions. We also provide a short introduction to the prior work most related to the concept of information diets.

In Chapter 3, we examine the word of mouth information diets of social media users along the dimensions of topical diversity and geographic source diversity.

In Chapter 4, we investigate the impact that personalized recommendations have on the topical diversity of information that users consume on social media platforms.

In Chapter 5, we present a bias quantification framework for search systems and apply it to study the bias for political searches on social media and Web search platforms.

In Chapter 6, we conclude the thesis with a short discussion of the main findings of the thesis and their implications, and a brief description of some directions of future work.

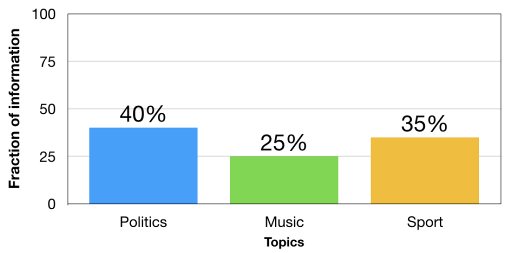
CHAPTER 2

Information diets: The concept

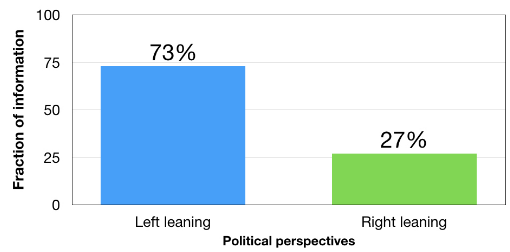
As we described in the previous chapter, the widespread adoption of social media platforms like Twitter and Facebook has lead to a paradigm shift in the way our society is producing and consuming information — from the broadcast mass media to online social media. Earlier, the only producers of information were traditional news organizations, which broadcast the same carefully-edited information to all the consumers over mass media broadcast channels. Whereas now, on online social media, any user can be a producer of information, and every user selects which other users they connect to, thereby choosing the information they consume. In addition, most social media platforms also employ various automated retrieval systems like personalized search and recommendation systems which also contribute towards the information that an individual user consumes on social media platforms.

In this scenario, we want to understand better what information users are producing and consuming on online social media via both the organic word of mouth channels and the algorithmic channels of recommendations and search. We define the concept of *information diets* as the composition of information produced or consumed. It forms a useful metric to measure the diversity or bias in the information produced and consumed by social media users.

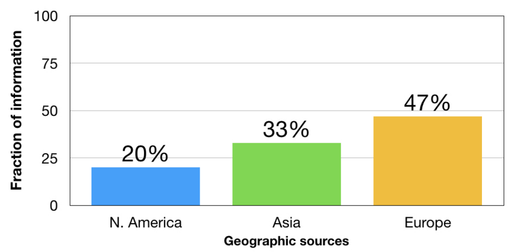
In this chapter, we begin by elaborating on our definition of information diets (Section 2.1) and how it can be used to measure the bias and diversity in the information exchanged on social media. We then discuss briefly prior work related to information diets and other related concepts (Section 2.2). And we end with a discussion of the limitations of our current operationalization of information diets and the possible future extensions (Section 2.3).



(a) **Topics**



(b) **Political perspectives**



(c) **Geographic sources**

Figure 2.1: Examples of the different dimensions on which information diets can be constructed: (a) topics, (b) political perspectives, and (c) geographic sources.

2.1 Information diets – A descriptive metric

To study the information consumption and production behavior of social media users, we define the concept of *information diet*. Similar to diet in nutrition, information diet is the composition of a set of information items being produced or consumed by a user. Information diets can be computed as the distribution of information items across many different dimensions, for instance, topics, perspectives on societal issues/debates, political leanings or opinions, geographic origins of sources, and many more. In this thesis, we focus on measuring information diets along three dimensions – *topics* (*e.g.,* user consumes 40% on politics, 25% on music and 35% on sports), *perspectives* (*e.g.,* user consumes 73% politically left-leaning news and 27% right-leaning news), and *sources* (*e.g.,* user consumes 33% news originating from Asia, 47% from Europe and 20% from North America), as shown in the toy examples depicted in Figure 2.1.

Information diets form a useful metric to measure the diversity and bias in the information exchanged on online social media platforms. In this thesis, we use the construct of information diets to not only study the information produced or consumed via word of mouth (Chapter 3) but also use it to examine the role automated algorithmic systems like recommendation (Chapter 4) and search systems (Chapter 5) play in shaping the diets of social media users.

In this thesis, we aim to use the metric of information diets in a descriptive manner to measure the bias and diversity in the information produced or consumed by social media users and to quantify the impact of automated retrieval algorithms on the information they consume. Making normative judgments about what is a "good" or "bad" diet is beyond the scope of this thesis, though we hope that in the future this measurement can lead to design and development of mechanisms for defining and promoting "better" information diets for social media users.

2.2 Prior work related to information diets

2.2.1 Normative approach to information diets

Johnson in his book titled "The Information Diet: A Case for Conscious Consumption" [110] states "*We know we're products of the food we eat. Why wouldn't we also be products of the information we consume?*". Through his book, he draws parallels between the industrialization of food and the on-going industrialization of information, both regarding production and consumption. He compares news and information which reaffirms our beliefs to the food we crave eating like salt, sugars and fat and postulates that media companies have also discovered what food companies already knew, that if you want to sell a lot of cheap calories pack them in with things that people crave, *i.e.,* selling affirmation and sensationalism is easier than selling balanced information since who would prefer to hear the truth than to hear that they are right. The part which makes the situation even grimmer is that unlike food, information diet of an individual not only impacts them but also has serious societal consequences. In his book, he attempts to describe good information consumption habits and suggests consuming more of the right stuff and consciously developing healthy habits.

While Johnson makes *normative* claims about what comprises a good information diet in his book, our definition of information diets is purely *descriptive* and does not include judgments of what is good or bad. While a descriptive information diet metric like ours tells what is the information that users are producing or consuming is, a normative outlook tells what information should users be consuming. Thus, we define information diets as a metric which can be used to measure the bias and diversity in the information produced or consumed by users. This measurement could in the future lead to the design of mechanisms for identifying and promoting what "*good*" information diet should look like. However, it is beyond the scope of this thesis.

2.2.2 Journalism and communication's repertoire approach to information diets

2.2.2.1 Within-platform repertoires

Channel repertoires: Heeter *et al*. [95] coined the term channel repertoires to describe the set of TV channels that an individual or a household watch regularly. In spite of a larger set of channels being available, most individuals and household maintained a smaller set in their channel repertoires as a way of coping with the large and varied media environment. People typically watch a small portion of all available channels, similar to what we observe with the focused consumption diets of social media users along topics and geography of sources (Chapter 3). Researchers have shown that channel repertoires vary in size and composition [96, 244, 71, 259] for different users.

Web repertoires: Ferguson & Perse [72] applied the repertoire approach to study the websites that users access on the world wide web, and found that most users maintain repertoires of websites that they visit very frequently and only consume a small fraction of the content available online.

Social media repertoires: Perhaps, the most related work to our study of information diets of social media users is the study conducted by Schmidt [207] where he examines the personalized set of sources that social media users aggregate for themselves, which he refers to as the Twitter friend repertoires. While in this thesis we construct information diets by examining the topics, geography of sources, and political perspectives of the information items that a social media user consumes via his chosen set of sources, the friends repertoires are constructed by comparing the set of sources (*i.e.,* friends) that a social media user chooses for themselves against a list of previously identified accounts of public figures and organizations.

2.2.2.2 Cross-platform repertoires

Information repertoires: With the model of information repertoires, Reagan [196] proposed that individual people choose different media sources (such as newspapers, radio, television or interpersonal communication) to get information on their topic of interest, according to their circumstances and needs.

Media repertoires: The concept of channel repertoires was extended to the idea of media repertoires to account for the cross-platform media use by the people. The media repertoires refer to the set of media sources that are part of individuals' regular media usage behavior [235, 92, 258, 223, 243]. The main principles characterizing media repertoires [92] – user-centered perspective (originating in audience research in media studies) attempts to capture the entirety of media components a user picks, and relationality of repertoire components (*i.e.,* their relative shares) – also hold true for our formulation of information diets.

2.3 Limitations & future directions

In this thesis, we have predominantly focused on quantifying and characterizing single-platform (within Twitter) information diets of social media users. However, to fully understand users' information consumption in the increasingly complex contemporary media environment, it is essential to extend our work to perform cross-platform studies of information diets in the future, where we examine the information that users get on the different digital and social media platforms that they subscribe to.

Moreover, in this thesis, we do not provide any normative guidelines for what constitutes a "good" or "balanced" information diet. I believe that to answer this question we need a concentrated effort from multiple disciplines, with the social scientists and technological scientists working together. While the technologists can help develop methods and mechanisms for capturing and measuring diets of users, the social scientists can help answer many confounding questions about not only how to define a balanced information diet, but also who should prescribe or enforce it – a central body like government or should each individual determine what's optimal for them. I believe that while a prescription of what's a generally desirable diet by a central body may be alright, but just as with food diet, each individual may need to decide based on their requirements how precisely to tweak general prescription to make it optimal for themselves.

And finally, even though we propose a descriptive view of information diets, making the users aware of them can potentially have a prescriptive effect on them. For instance, when activity tracking apps show users descriptive statistics about how many steps they have taken and how many stairs they have climbed, it may motivate them to change their daily behavior to include more activity [183]. Therefore making the users aware of their information diets may itself be the first step towards motivating and helping them towards correcting or improving their information consumption behaviors.

CHAPTER 3

Word of mouth: Information diets of social media users

With the widespread adoption of social media sites like Twitter and Facebook, there has been a shift in the way information is produced and consumed in our society. Earlier, the only producers of information were traditional news organizations, which broadcast the same carefully-edited information to all consumers over mass media channels. Whereas, now on online social media, any user can be a producer of information, and every user selects which other users they connect to, thereby choosing the sources they consume information from via word of mouth. By choosing the sources from which they consume information, users have overthrown the role of gatekeeping that journalists traditionally fulfilled [213].

We begin the study of information diets of social media users by examining the most natural channel of information exchange on social media, *i.e.,* the *word of mouth channel* via the social links created by the users. We investigate the word of mouth diets of social media users along two different dimensions – topical diversity and geographical source diversity.

In the first part this chapter (Section 3.1), we apply the concept of information diets to study the topical characteristics of users' information production and consumption on a popular social media platform – Twitter. We contrast these social media diets with the mass media diets of some mainstream news organizations on mass media (namely New York Times, Washington Post and The Economist), to study the impact of the paradigm shift in news production and consumption.

In the second part of this chapter (Section 3.2), we focus on studying the information diets of social media users along the dimension of geographical source diversity. Here we consider a country as a geographical unit, and instead of studying individual user's diets, we investigate the aggregated diet that all the users within a country are consuming as a whole. Doing so, we examine the geographic diversity of the sources of information that different countries are consuming from and investigate the similarities between the country and the most significant contributors to its diet.

3.1 Topical diversity in information diets: Social media vs. mass media

Having defined the concept of information diets in Chapter 2, we now focus on the *topical composition* of users' diets, *i.e.,* the fractions of their information diets that correspond to different topical categories of information (*e.g.,* information on politics, sports, entertainment, and so on). Using these topical diets we characterize the information produced and consumed by various types of users in the popular Twitter social media and compare it to the diets produced by news organizations on mass media.

One of the key goals of this thesis is to understand better how the differences in information production and consumption processes between mass media and online social media affect users' diets. To address this goal, we conducted a comparative analysis of the topical compositions of the information diets produced and consumed on social media with those on mass media. Our investigation focused on the following high-level questions:

1. **Production**: What is the topical composition of the information published on mass media (*e.g.,* NYTimes print edition)? How does the information produced on social media compare with the information published on mass media?

2. **Consumption**: How topically balanced or unbalanced are consumption diets of social media users (relative to the mass media diet)? Are users' consumption diets heavily skewed towards a few topics of their interest, or do they also receive information on the broad variety of topics that are covered in mass media?

In the rest of this section, we begin by giving a brief overview of related work in Section 3.1.1. Next, to conduct our study, we needed a methodology to infer the topics of individual posts on Twitter. The short length of tweets makes it challenging to infer topics at the level of individual tweets. We propose a novel methodology to infer the topic of a social media post by leveraging the topical expertise of the Twitter users who have posted it, which we describe in Section 3.1.2. The topical vector determined by aggregating the topics of a set of social media posts that a user produces or consumes denotes the topical information diet corresponding to the set of posts. Finally, we end this section by presenting our findings on the production and consumption diets of users via mass media and online social media.

3.1.1 Related work

Two lines of research are related to our study of the topical diversity of information diets of social media users – (i) inferring topics of social media posts, and (ii) assigning social media users to topics for identifying topical authorities on social media. Next, we briefly discuss the prior work done in both these directions.

3.1.1.1 Inferring topics of social media posts

To our knowledge, all prior attempts to infer the topic of a tweet/hashtag/trending topic rely on the content itself – either applying NLP and ML techniques like topic modeling or mapping to external sources such as Wikipedia or Web search results – to infer the topics.

NLP and ML based approaches: A large section of prior work on topic inference for social media posts has applied NLP and ML techniques like topic modeling. Zhao *et al.* [263] compared the traditional media (NY Times) with Twitter using topic models, while, Hong *et al.* [105] performed an empirical study of topic modeling on Twitter itself. Yang *et al.* [252] developed a high precision topic modeling system for tweets using a supervised approach in real time. Lin *et al.* [138] used language modeling techniques for filtering posts on different topics in a continuous tweet stream. Whereas, Ramage *et al.* [194] used topic models to characterize users and content on Twitter on the dimensions of substance, status, style and social characteristics of posts. Several others [192, 61] have also performed topic modeling on the content posted by users to assign topics to users.

Mapping to external sources: On the other hand, some researchers [147, 28, 73] have taken the alternative approach of mapping social media posts to external sources such as Wikipedia or Web search results to infer their topics. Meij *et al.* utilized semantic linking, *i.e.,* identifying semantic concepts being talked about in a microblog post in an automated manner by linking it to related Wikipedia articles. Ferragina *et al.* [73] also leverage Wikipedia for inferring topics of hashtags by linking hashtags and Wikipedia entities using the topic annotator TagME. On the other hand, Bernstein *et al.* [28] utilized search engine as a distributed knowledge base and discovered topics of social media posts by making search queries using the filtered text of the posts.

Such methodologies that rely on the text of the social media posts are of limited utility in the case of social media like Twitter, primarily due to the short length of tweets, and the informal nature of the language used by most users [211, 238]. In contrast to these previous approaches which focus on the content, our methodology focuses on the characteristics of the *authors* of the content to infer its topic.

3.1.1.2 Assigning topics to social media users

Till now we described prior work that focussed on identifying topics of social media posts. We now shift our focus to studies which have attempted to assign topics to social media users, to either identify their topics of interest [150, 192, 251], or to identify topical authorities on social networks [238, 246, 174, 193, 61].

Inferring topics of interest of users: Quercia *et al.* [192] used topic modeling to assign topics to users and to infer users' topics of interest, while Michelson and Macsskassy [150] and Xu *et al.* [251] also inferred the topics of interests of users using author topic models. On the other hand, Bhattacharya *et al*. [29, 30] did not rely on the content that users post, instead they utilized crowdsourced topical expertise labels of the followees of a user to infer the user's topics of interest.

Identifying topical experts on social media: Canini *et al.* [193] used Twitter text search along with social links in the network to find topically relevant users, while Dimitrov *et al.* [61] automatically assigned users to topics using topic modeling and AlchemyAPI on the content the user produces. Weng *et al.* [246] built TwitterRank, which is similar to topic-sensitive PageRank and identifies influential users on different topics based on follow links and content similarities. Pal and Counts [174] compared and utilized network and content topic features to identify topical authorities on Twitter. Finally, Wagner *et al.* [238] used topic modeling on different types of data including the posts a user tweets or retweets, the profile bio and the lists to which the user belongs, for modeling the topical expertise of the user. They found that the tweet content is not very useful for inferring user's expertise. Instead, it is better to rely on other user-related information like the lists that they are included in. We also leverage a List-based methodology [211, 80] to retrieve topical expertise tags of users, for inferring the topic of a keyword in a tweet by leveraging the topical expertise of the different users posting the keyword on Twitter.

3.1.2 Methodology: Quantifying topical information diets

In this part of the thesis, we compute the *information diet* of a set of information items (*e.g.,* a set of tweets or hashtags), as the topical composition of the information items. We define the topical composition over a given set of topics as the fraction of information related to each topic. In this section, we present our methodology for quantifying the topical information diet for a set of tweets. The methodology for quantifying the topical information diets of social media users consists of two main steps: (i) inferring the topic of a keyword in a tweet (Section 3.1.2.1), and (ii) aggregating topics of keywords into the information diet of a set of tweets (Section 3.1.2.2).

3.1.2.1 Inferring topic of a keyword

We begin this section by describing our choice of keywords for inferring the topics of tweets, followed by a description of the topical hierarchy we use for this study. We then make use of an author-based crowdsourced methodology to infer topic of each keyword which relies on leveraging the topical characteristics of Twitter users who are authoring or posting the keyword. Finally, we present the evaluation of our topic inference methodology and show that our author-based crowdsourced method performs better than a state-of-the-art content based tool – AlchemyAPI.

Selecting keywords: We choose *hashtags* and *URLs* as the basic elements of information in a tweet and collectively refer to them as keywords. To justify our choice of keywords, we conducted an Amazon Mechanical Turk (AMT) [17] survey. In the survey, we showed 500 randomly selected tweets from Twitter's 1% random sample that did *not* contain any hashtags or URLs to AMT master workers [18]. A majority of the workers judged 96% of these tweets without keywords to be non-topical, *i.e.,* consisting of mostly conversational babble. Therefore our selected keywords – hashtags and URLs – contain important signal about the topicality of tweets, justifying our choice to consider them as keywords for inferring topics of tweets. However, our methodology can be easily extended to include other kinds of keywords, such as named entities.

Level of topical hierarchy	Number of nodes
Level-0	18
Level-1	262
Level-2	529
Level-3	231
Level-4	47
Total	1087

Table 3.1: Number of nodes at different levels of the topical hierarchy.

Selecting topical hierarchy: We combine two standard topical hierarchies – the Open Directory Project[1] and AlchemyAPI[2] to construct our topical hierarchy. The 18 top-level topical categories were selected by combining the top categories of the two hierarchies and comprise of arts-crafts, automotive, business-finance, career, education-books, entertainment, environment, fashion-style, food-drink, health-fitness, hobbies-tourism, paranormal, politics-law, religion, science, society, sports, and technology. The lower levels were derived from the lower levels of the AlchemyAPI hierarchy by showing them to two human annotators and asking them to independently map each node (with all its descendants, if any) in the hierarchy to one of the chosen 18 top-level topics and then coming to an agreement. Doing so we constructed a topical hierarchy with a total of $1,087$ nodes, with the number of nodes in each level indicated in Table 3.1. Additionally, Table 3.2 shows the topics in the top two levels of our topical hierarchy.

Mapping Twitter users to the topical hierarchy: To identify the topical expertise of Twitter users, we leveraged the List-based methodology developed in prior work [211, 80] to retrieve *expertise tags* for topical experts. Table 3.3 shows some sample topical experts from our experts' dataset along with a sample of their expertise List-tags. For instance, some of the tags inferred by this methodology for Lady Gaga are 'music', 'entertainment', 'singers', 'celebs' and 'artists'.

To map these experts to our topical hierarchy, we mapped each node in our topical hierarchy to one or more semantically similar List-tags. The semantically similar tags to a node in our topic hierarchy were identified using a semi-automatic process. A tag co-occurrence graph was constructed, where nodes were List-tags, and two nodes were linked if more than k (for this work we used $k = 3$) experts are annotated with both these tags. We started with a 'seed tag' (*i.e.,* a word directly appearing in our topic hierarchy), and then manually checked the neighboring tags of the seed tag in the tag

[1] See **www.dmoz.org**.

[2] See **https://www.ibm.com/watson/developercloud/alchemy-language/api/v1/#taxonomy**.

Level 1 topic	Level 2 topics
arts-crafts	art-and-technology, crafts, interior-decorating, visual-art-and-design
automotive	auto-parts, auto-repair, bicycles-and-accessories, boats-and-watercraft, buying-and-selling-cars, campers-and-rvs, cars, certified-pre-owned, commercial-vehicles, driving, electric-vehicles, minivan, motor-shows, motorcycles, off-road-vehicles, roadside-assistance, scooters-and-mopeds, trucks-and-suvs, vehicle-brands, vehicle-manufacturers, vehicle-rental
business-finance	advertising-and-marketing, aerospace-and-defense, automation, biomedical, business-news, business-operations, business-software, chemicals-industry, company, construction, dairy, energy, finance, home-and-garden, hospitality-industry, iron-and-steel-industry, logistics, manufacturing, metal-industry, mining-industry, paper-industry, pharmaceutical-industry, publishing, real-estate, record-company, shipping-industry, shopping, tanning, textile-industry, war-industry
careers	career-advice, career-planning, job-fairs, job-search, nursing, resume-writing-and-advice, telecommuting, us-military
education-books	books, books-and-literature, education
entertainment	adult-entertainment, celebrity-fan-and-gossip, comics-and-animation, dance, games, humor, movies, movies-and-tv, music, radio, shows-and-events, theatre
environment	agriculture-and-forestry, animals, environmental-safety, green-solutions, renewable-energy, weather-info
fashion-style	accessories, beauty, body-art, clothing, fashion-designers, fashion-industry, footwear, jewelry, luxury-fashion, mens-fashion, swimwear, underwear
food-drink	barbecues-and-grilling, beverages, cuisines, desserts-baking, dining-out, food, food-allergies, food-and-grocery-retailers, food-industry, gastronomy, health-lowfat-cooking, healthy-eating, kosher-food, vegan, vegetarian
health-fitness	addiction, aging, alternative-medicine, dental-care, disease, disorders, drugs, exercise, health-news, healthcare, incest-and-abuse-support, mens-health, nutrition, organ-donation, sexuality, sports-medicine, therapy, weight-loss, womens-health
hobbies-tourism	birdwatching, cigars, collecting, gardening-and-landscaping, getting-published, home-recording, inventors-patents, magic-and-illusion, needlework, reading, scrapbooking, tourism
paranormal	astrology, occult, parapsychology
politics-law	armed-forces, espionage-and-intelligence, government, immigration, law-commentary, law-enforcement, legal-issues, politics
religion-spiritualism	alternative-religions, atheism-and-agnosticism, buddhism, christianity, hinduism, islam, judaism
science	biology, chemistry, computer-science, ecology, engineering, geography, geology, mathematics, medicine, physics, science-news, social-science
society	charity, crime, dating, family-and-parenting, gay-lesbian, racism, rape, senior-living, social-institution, teens, unrest-and-war, welfare, work
sports	archery, auto-racing, badminton, baseball, basketball, bicycling, billiards, boat-racing, bobsled, bodybuilding, bowling, boxing, canoeing-and-kayaking, cheerleading, climbing, cricket, curling, diving, dog-sled, fencing, fishing, go-kart, golf, gymnastics, handball, hockey, horses, hunting-and-shooting, martial-arts, motorcycling, olympics, paintball, parachuting, polo, rodeo, rowing, rugby, running-and-jogging, sailing, scuba, skateboarding, skating, skiing, snowboarding, soccer, softball, sports-news, surfing, swimming, table-tennis, tennis, trekking, volleyball, wakeboarding, walking, water-polo, weightlifting, windsurfing, wrestling
technology	computer-certification, computer-crime, computer-reviews, computer-security, consumer-electronics, data-centers, electronic-components, enterprise-technology, hardware, internet-technology, mp3-and-midi, networking, operating-systems, programming, software, tech-news, technical-support, technological-innovation

Table 3.2: Topics in the top two levels of our topical hierarchy.

Sample experts	Sample expertise List-tags	Inferred topic
Lady Gaga (@ladygaga)	music, entertainment, singers, celebs, artists	Entertainment
Chuck Grassley (@ChuckGrassley)	senator, congress, government, republican	Politics
The Linux Foundation (@linuxfoundation)	linux, tech, software, computer, ubuntu	Technology

Table 3.3: Sample topical experts along with a sample of their expertise List-tags and the inferred topic (out of the 18 topical categories).

Topic categories	Some related terms
Arts-crafts	art, history, geography, theater, crafts, design
Automotive	vehicles, motorsports, bikes, cars
Business-finance	retail, real-estate, marketing, economics
Career	jobs, entrepreneurship, human-resource
Education-books	books, libraries, teachers, school
Entertainment	music, movies, tv, radio, comedy, adult
Environment	climate, energy, disasters, animals
Fashion-style	style, models
Food-drink	food, wine, beer, restaurants, vegan
Health-fitness	disease, mental-health, healthcare
Hobbies	photography, tourism, gardening
Paranormal	astrology, supernatural
Politics-law	politics, law, military, activism
Religion	christianity, islam, hinduism, spiritualism
Science	physics, chemistry, biology, mathematics
Society	charity, LGBT
Sports	football, baseball, basketball, cricket
Technology	mobile-devices, programming, web-systems

Table 3.4: The 18 topic categories to which keywords and tweets are mapped, and some terms related for each topic. The terms are matched with expertise-tags of Twitter users, to map expert users to different topics.

co-occurrence graph. We ignored unrelated tags and considered the related tags as being semantically similar to the corresponding seed tag. Some examples of semantically unrelated but frequently co-occurring tags are 'news', 'media' and 'celeb' which tend to co-occur with many tags and therefore we ignored them. Additionally, all experts mapped to lower-level topics were also considered to be mapped to the higher-level topics in the hierarchy. Following this procedure, we mapped $1,564,411$ experts to one or more topics in our topic hierarchy.

Author-based crowdsourced topic inference methodology: Much of the prior approaches for inferring the topic of a tweet or a keyword rely on the content itself and these tend to perform poorly on short social media posts which typically contain informal language [211, 238]. We propose an alternative *author-based crowdsourced topic inference technique* which relies on the topical expertise of users discussing the keyword. The basic intuition for our method is that if many users interested in a particular topic are discussing a specific keyword, then that keyword is most likely related to that topic. Our author-based methodology has two main advantages over other content-based techniques: (i) it is highly scalable and can be easily applied for millions of tweets, and (ii) it can cope with the informal and constantly-evolving vocabulary on social media and does not require constant re-training.

For performing this study, we only considered the 18 top-level topical categories from our topical hierarchy. These 18 topic-categories and their related terms are shown in Table 3.4. We quantify topical information diets of Twitter users by inferring the fraction of information from each of these 18 topics. In the future, more fine-grained topical information diets could be constructed using the full topic hierarchy.

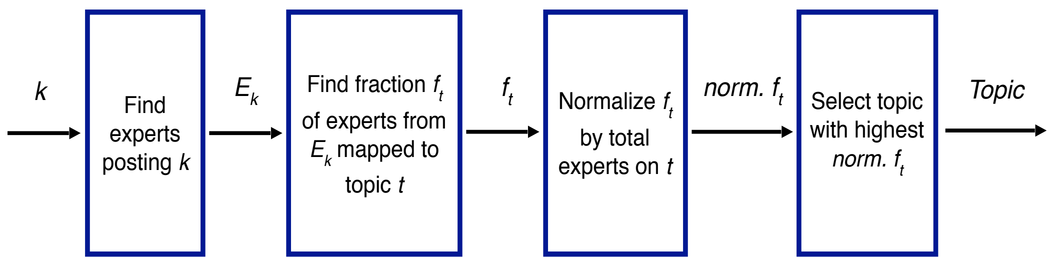


Figure 3.1: The different steps of our author-based crowdsourced methodology for inferring the topic of a keyword k.

Building upon the basic intuition, to infer the topic of a keyword k we first identify the set of expert users E_k from our topical experts dataset who have posted k. If less than 10 experts have posted k, we do not attempt to infer its topic. Next, for each of the 18 topics t (shown in Table 3.4), we compute the fraction of experts (f_t) out of E_k who are mapped to the topic t. We normalize this fraction f_t by the total number of experts on topic t in our dataset, to account for the varied number of experts mapped to different topics in our dataset. And finally, we select the topic with highest normalized f_t as the inferred topic for keyword k. Our author-based crowdsourced methodology for inferring the topics of a keyword is depicted in Figure 3.1.

Evaluating the topic inference methodology: We evaluated the performance of our topic inference methodology using two metrics: (i) coverage: fraction of keywords for which the methodology infers a topic, and (ii) accuracy: fraction of keywords for which the inferred topic is relevant. We also compared

Metric	Methodology	Hashtags	
		Popular	**Random**
Coverage	AlchemyAPI	22.5%	55.5%
	Proposed	98%	82.5%
Accuracy	AlchemyAPI	44.44%	51.35%
	Proposed	58.67%	49.69%

Table 3.5: Comparing the proposed author-based crowdsourced topic inference methodology with AlchemyAPI (which uses content-based NLP techniques) in terms of coverage and accuracy.

the performance of our proposed author-based methodology with a state-of-the-art content-based commercial service, AlchemyAPI, that makes use of NLP and deep learning to infer topics.

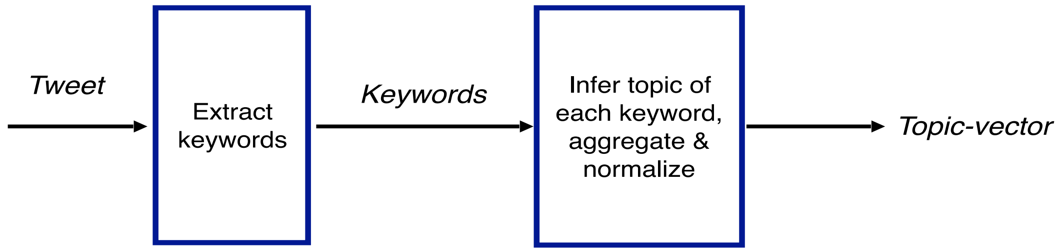
Our evaluation dataset, constructed from Twitter's 1% random sample for a week in December 2014, consists of (i) 200 popular hashtags which are tweeted the most number of times, and (ii) 200 randomly selected hashtags. The topics for these sets of hashtags were inferred using both our methodology and using AlchemyAPI. We passed 1000 randomly chosen tweets containing a hashtag to AlchemyAPI to infer the topic of the hashtag.

Table 3.5 shows the performance of our proposed topic inference methodology and AlchemyAPI. Regarding coverage, our methodology performs significantly better than AlchemyAPI, possibly because AlchemyAPI does not perform well for informal and shortened language used in most tweets. Also, our methodology has higher coverage for popular hashtags than for random hashtags, since we require the hashtag to be posted by at least ten experts to infer its topic.

To evaluate the accuracy, we measured the relevance of a topic for a keyword via an Amazon Mechanical Turk (AMT) survey where we showed five workers the hashtag, 20 random tweets containing the hashtag and the inferred topic, and asked the workers if they found the inferred topic to be relevant to the hashtag. Table 3.5 shows the majority opinion of the AMT workers, using which we discovered that our methodology is accurate for a significant fraction of hashtags (59% for popular hashtags, and 50% for random hashtags), and performs better than AlchemyAPI (44% for popular hashtags, and 51% for random hashtags). The performance is found to be similar for URLs.

Overall, our proposed author-based methodology performs better than a state-of-the-art NLP-based technique in inferring topics of hashtags. The proposed methodology is notably better at inferring the topics for popular hashtags, with not only higher coverage, but also higher accuracy for the inferred topics.

3.1.2.2 Constructing information diets from topics of keywords



Example:

Tweet: “This is the best #vegan #diet for good #health.”

Keywords: #vegan, #diet, #health

Topics of keywords: #vegan - Food, #diet - Health, #health - Health

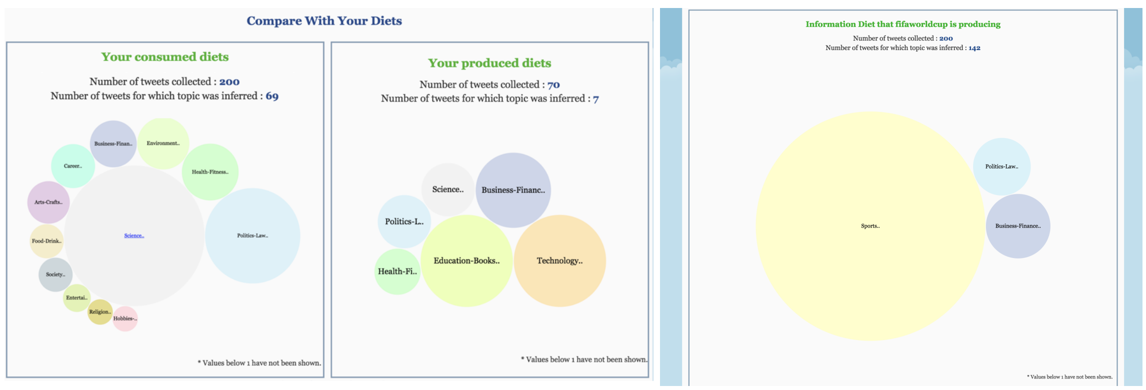
Topic-vector: Health : 0.67, Food : 0.33

Figure 3.2: Constructing information diets by extracting keywords from each tweet and inferring it’s topic and then aggregating and normalizing the contributions of all keywords, such that the weights in the final topic-vector sum up to one. A sample tweet is also depicted to exemplify the process of generating the topic-vector from the tweets.

To infer the information diet of a user, we first retrieve the full set of tweets produced (or consumed) by the user. Then, to construct the diet corresponding to this set of tweets, we first extract the keywords from each tweet and infer each keyword’s topic. Using these topics of the keywords, we construct a topic-vector for the given set of tweets, where the weight of a topic is the total contribution of all keywords inferred to be on that topic. Here, we normalize the contribution of each keyword within a tweet by the number of keywords in the tweet to ensure that each tweet contributes a total weight of 1 to the topic-vector. The information diet of the set of tweets is then given by this topic-vector. The process of constructing information diets from the tweets is shown in Figure 3.2, which also shows a sample tweet, the keywords extracted from it, the inferred topics for each keyword and the final aggregated and normalized topic-vector for the tweet.

3.1.2.3 Public deployment of information diets

We have publicly deployed a Twitter application to make users more aware of the information diets they are producing and consuming on Twitter social media, at **http://twitter-app.mpi-sws.org/information-diets/**. Using this system, the users can explore the diets they are consuming on Twitter, as well as filter the tweets they are consuming by the topics and selectively read them. Additionally, they can also search for other Twitter users to examine the diets they are producing. Figure 3.3 depicts the screenshots of our Twitter application showing the diets produced and consumed by the logged in user as well as the diet produced by the searched user '@fifaworldcup'. More details about the functionalities of our application and a pointer to a demo video can be found in Appendix A.1.



(a) **Diets consumed and produced by the logged-in user** (b) **Diet produced by '@fifaworldcup'**

Figure 3.3: Screenshots of our Twitter application for making users aware of their information diets showing the (a) diets consumed and produced by the logged-in user, and (b) diet produced by searched account '@fifaworldcup'.

3.1.2.4 Limitations

We briefly discuss some limitations of our approach for quantifying the topical information diets of Twitter users. Our methodology relies on the keywords in the tweets to be tweeted by multiple topical experts to infer the keywords' topics, and this method can sometimes lead to an incorrectly inferred topic, and here we outline some such problem scenarios.

First, if many experts from many different topics tweet a keyword, then we are more likely to infer it's topic incorrectly. This scenario may occur for keywords about issues of general interest (*e.g.,* Christmas) or about an event of global importance (*e.g.,* a terrorist attack), both of which may lead to many different topical experts to tweet about it (possibly about various aspects of it). For instance, in our dataset collected in December 2014, Christmas related keywords like #merrychristmas #christmaseve, #happyholidays were being posted universally. Therefore, though the AMT workers labeled these as 'religion-spiritualism', we ended up inferring 'entertainment' and 'food-drinks' *etc.* for them, depending on which topical experts were relatively more active in posting about them. In the future, this situation could be remedied by either imposing a per-topic threshold on the number of experts or using a measure of dispersion to identify the cases where comparable numbers of experts on too many topics are posting about a keyword of general interest.

Second, the topics of keywords may not remain constant over time [260]. To mitigate the impact of the evolution of topics of keywords over time and to get best topic inference results, it is important to use expert tweets from a similar period as the tweets for which we want to infer the topics. Especially in polarized scenarios, prior work [89] has shown that Twitter users may even actively "hijack" hashtags to further their agendas.

And finally, since we infer the topics of only those keywords which have been tweeted by at least ten topical experts, we have a lower coverage and accuracy for non-popular keywords. However, the later sections show that the popular information forms a significant fraction of users' diets, and therefore our approach is likely to be able to estimate the information diets of users reasonably accurately.

3.1.3 Production: How are the diets produced on social media and mass media different?

Traditionally, in mass media, editors of news-organizations are expected to follow specific guidelines to ensure that the news-stream has a balanced coverage across various topics of interest of the subscribers. In contrast, every user-account in social media can serve as an information producer, and there are no definite guidelines regarding the content that any account posts. One of the primary goals of this thesis is to analyze the effects of these differences, thus in this section, we compare and contrast the information diets produced on online social media with the information diets produced over mass media.

3.1.3.1 News organizations: Social media vs. mass media

We begin the comparison of information diets produced over mass media and online social media by focusing on the diets that well-known news organizations are producing over the two types of media.

Mass media diets of news organizations: To measure the mass media diets, we focused on three popular news organizations – New York Times, Washington Post, and The Economist – and collected their broadcast print editions for three days in December 2014. We obtained human annotated topical

Topic	NYTimes	Washington Post	Economist
Arts-Crafts	4.56%	0.0%	1.85%
Automotive	1.34%	0.0%	0.37%
Business-Finance	7.51%	8.65%	28.04%
Career	0.8%	0.48%	0.74%
Education-Books	1.88%	5.29%	3.32%
Entertainment	12.33%	13.94%	1.48%
Environment	3.49%	0.96%	7.01%
Fashion-Style	0.0%	1.44%	0.0%
Food-Drink	4.83%	6.25%	2.21%
Health-Fitness	6.17%	5.29%	2.95%
Hobbies-Tourism	1.34%	0.0%	0.37%
Paranormal	0.27%	0.0%	0.0%
Politics-Law	29.49%	37.5%	35.06%
Religion	2.14%	0.96%	2.95%
Science	1.34%	0.96%	2.58%
Society	3.75%	6.73%	3.32%
Sports	15.01%	9.62%	1.11%
Technology	3.75%	1.92%	6.64%

Table 3.6: Mass media information diets of three news organizations, where the topics of the news-articles were judged by AMT workers (top topics highlighted).

labels (out of our 18 topical categories shown in Table 3.4) for each of their articles via an AMT survey. Each news article was shown to five distinct workers, and the majority verdict was considered as the topic for the news article.

Using these topical labels, we constructed the mass media diets for these news organizations. Table 3.6 shows the mass media information diets of the three news organizations. We find that all the news organizations tend to focus (*i.e.,* post majority of their news articles) on a few popular topics – politics, entertainment, and sports for NYTimes and Washington Post, and mainly politics and business-finance for The Economist. However, despite their bias towards these few popular topics, the mass media diets also have a spread over the remaining less popular topics – the 12 least popular topics contribute 25% of the diet for NYTimes and 17% for both Washington Post and Economist.

In the following sections, we use these mass media diets as a baseline for comparing information diets produced on social media.

Social media diets of news organizations: We begin by addressing the question – *are there differences between the information diets published by news organizations over mass media and social media?*. To answer this question, we collected the tweets posted by the Twitter accounts of our selected three news

Social media account	Topic of specialization	Contribution of topic	
		Social media	Mass media
NYTSports	Sports	66.6%	15.0%
nytimesbusiness	Business	66.1%	7.5%
nytimesbooks	Edu-Books	59.1%	1.9%
EconUS	Business	74.4%	28.0%
EconWhichMBA	Education	37.6%	3.3%
	Business	32.1%	28.0%
PostSports	Sports	88.5%	9.6%
PostHealthSci	Science	34.5%	0.96%
	Health	25.1%	5.3%
WaPoFood	Food	60.3%	6.3%

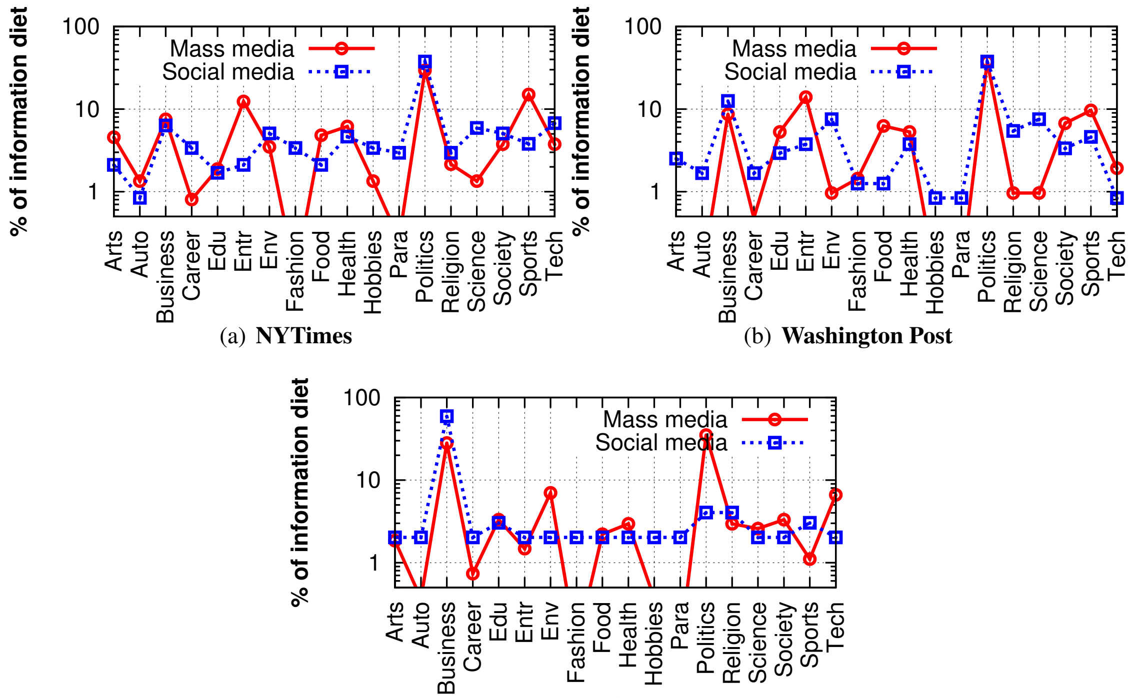
Table 3.7: Examples of topic-specific Twitter accounts of news organisations, along with the contribution of their topics of specialization in their produced diet.

organizations for the same three days in December 2014 and computed the diets they are producing over social media.[3]

Interestingly, we observe that each of the three news organizations have multiple accounts on Twitter. These include one primary account (@nytimes, @washingtonpost, and @economist) and several *topic-specific accounts* (*e.g.,* @NYTSports, @EconSciTech, PostHealthSci) each of which specializes in posting news stories on a particular topic. Table 3.7 shows some of the topic-specific accounts of the three news organizations, along with the fraction of their produced diet that is on their topic of specialization. It is evident that the topic-specific accounts produce a much larger fraction of their diet on their specific topics of specialization, as compared to the mass media diet of the same news organization.

While the topic-specific accounts of the news organizations have thousands to hundreds of thousands of followers, a much higher number of users subscribe to the primary accounts. For instance, the primary account @nytimes has 40.6M followers, while the topic-specific accounts @NYTSports and @nytimesbusiness have 93K and 794K followers respectively. Since most social media users consume the diet produced by the primary account, we compare the social media diet produced by the primary account with the mass media diet of the same news organization.

[3]The results presented in this section are for the same three days in December 2014, over which both the mass media diets and social media diets were analyzed. However, we observed that the information diets remain relatively unchanged over longer time-durations too.



(a) **NYTimes** (b) **Washington Post**

(c) **Economist**

Figure 3.4: Comparing the information diet posted by news organizations in mass media (news articles in print editions) and social media (tweets posted by their primary Twitter accounts) for the same days in December 2014. (Topics with contribution less than 0.5% not shown.)

Figure 3.4 compares the information diets produced by the three news organizations over mass media, with those produced by their primary accounts over Twitter social media. We find two main differences between the mass media and social media diets of the same news organization. First, the primary accounts of the news organizations in social media tend to publish less content (as compared to the corresponding mass media diets) on those topics for which there exist topic-specific accounts. For instance, for both NYTimes and Washington Post, topics such as sports and food are covered much lesser in the social media diets than in the corresponding mass media diets. Additionally, both the primary and the topic-specific social media accounts of the news organizations tend to be more specialized in their production by focusing on fewer topics, as compared to their mass media diets. For example, while the mass media diet of Economist focuses on both business and politics, the social media diet of @economist focuses solely on business and publishes far lesser content on politics.

Therefore, in summary, we find that there is an *unbundling of content production on social media* by the news organizations, with multiple accounts per news organization and each in turn specializing on a particular topic as compared to their mass media editions. This unbundling would enable users in

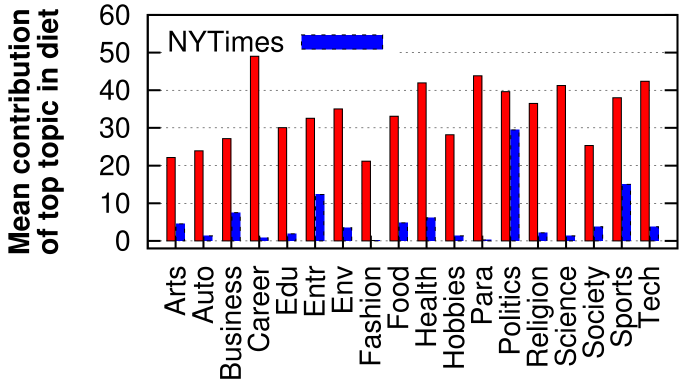


Figure 3.5: Mean contribution of the top topic (on which a user posts the highest fraction of their diet) in the produced diets of popular users, grouped according to their top topic of production.

social media to get focused information on their topics of interest by subscribing to the topic-specific accounts. However, the users who subscribe to only the primary account of the news organizations might not be aware that they are receiving a different information diet as compared to that of the mass media editions.

3.1.3.2 Popular social media accounts vs. mass media

Next, we study whether our observations about the specialized production of the social media accounts of news organizations generalizes to other popular user-accounts on Twitter. Prior research has shown that a substantial fraction of the information being consumed by users on social media sites like Twitter is produced by a small fraction of popular users [250]. Hence, we next study the information diets of the content posted by some popular accounts on Twitter. There are several ways to identify popular or influential accounts on Twitter, such as by the number of followers, or by the number of times an account is retweeted. In this study, we consider *verified users* as examples of popular user-accounts on Twitter. Out of all the verified users on Twitter who declared their language as English, were not protected accounts or news organizations, we randomly selected a set of 500 verified users. We collected the tweets posted by them during December 2014 and computed the information diets posted by these users using the methodology presented earlier.

For studying the specialization in the produced diet of each user, we define the *top topic* for the user as the topic on which they post the largest fraction of their diet. For the group of users having a common top topic, we compute the mean percentage contribution of their produced diet that is on their

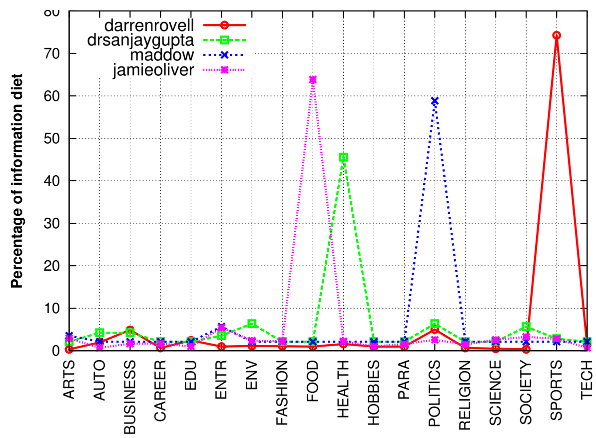


Figure 3.6: Information diets produced by four selected verified users – @darrenrovell (ESPN Sports Business Reporter, ABC News Business Contributor.), @drsanjaygupta (Staff Neurosurgeon, Emory Clinic; CNN Chief Medical Correspondent), @maddow (American television host and political commentator), @jamieoliver (British celebrity chef and restaurateur). All of them produce focused diets on their topics of expertise.

top topic. Figure 3.5 shows this mean percentage contribution for the group of users specializing on each topic. As a baseline, we also show the contribution of each topic in the NYTimes mass media diet (which was stated in Table 3.6). We find that the popular users, on average, post a significant fraction of their diet (between 20% and 50%) on just their top topic. Further, users having different top topics are focused to different degrees – for instance, popular users having career, health, paranormal, science, and technology as their top topic post more than 40% of their diet on their top topic. For instance, Figure 3.6 shows a sample of verified users with different fields of expertise and their produced diets, and we can again observe that they produce a large fraction of their diets on just the topic of their expertise. Anyone who subscribes to these popular sources of information on social media will get a much higher fraction of content on the corresponding topic, than what is obtained from a typical mass media source (as shown by the NYTimes baseline in Figure 3.5).

Additionally, we looked at the distribution of the 500 randomly selected verified users across their top topics. Figure 3.7 shows the distribution of these users according to their top topic. Most of the users have their top topic as one of the three topics – entertainment, sports, and politics. However, there are small fractions of popular users focusing their diets on all the other topics as well. These observations agree with recent findings [29] that although Twitter is primarily thought to be associated with few popular topics such as entertainment, sports, and politics, there are popular accounts who are experts on a wide variety of topics.

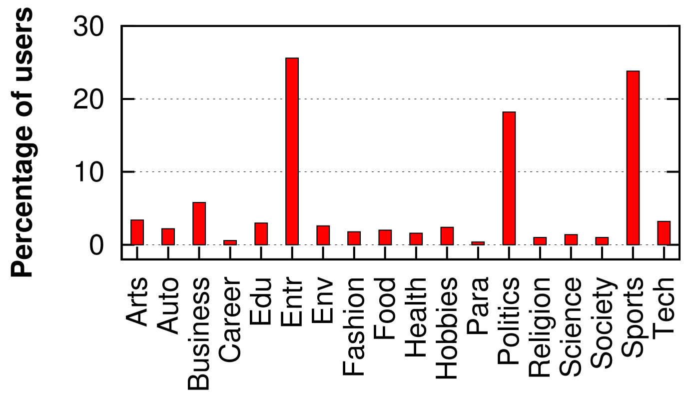


Figure 3.7: Distribution of the 500 randomly selected verified users, according to the topic on which they produce the maximum fraction of their diet.

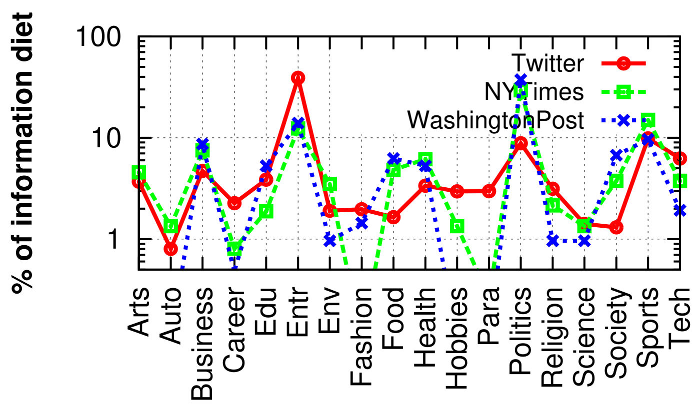


Figure 3.8: Comparing the information diet of the Twitter 1% random-sample with the mass media diet of news organizations (NYTimes and Washington Post).

These observations imply that, similar to mass media, there are sources of information on a wide variety of topics on Twitter social media. However, since every source produces a diet that is specialized on just a few topics, the consumers of information on social media need to be careful in deciding whom they subscribe to, especially if they desire to get a topically balanced information diet.

3.1.3.3 Random sampling of social media vs. mass media posts

Till now, we have shown that the individual sources of information in social media (popular user accounts as well as accounts of news organizations) produce diets that are very focused on specific topics. Now we shift the focus to the overall information being produced over the two media. We use the Twitter 1% random sample (for December 2014) to represent the overall information being produced on social media, and compare the information diet of the Twitter random sample with the mass media diets of NYTimes and Washington Post in Figure 3.8.

We observe that the diets from both social media and mass media are skewed, but towards different topics. Though both diets have entertainment, politics, sports, and business amongst the top topics, the Twitter social media diet is more heavily biased towards entertainment (39%), while the mass media diets focus more on politics (30%). Further, some topics are over-represented in the social media diet as compared to mass media diet, such as technology, hobbies-tourism, paranormal, and career. On the other hand, topics such as food, health, and society are covered more in mass media than in social media, which is probably because these topics are of general interest to many people in the offline world. Whereas, topics such as entertainment and technology are more dynamic, with new information being regularly generated, leading to them being covered more in a real-time information dissemination medium like Twitter.

3.1.4 Consumption: Are the social media users consuming balanced diets?

Unlike in mass media, where everyone consumes the same broadcast information, every user on social media shapes their own personalized channel of consumption by subscribing to other users. Having observed the unbundling of content creation on social media with each source producing a specialized diet, we next turn our attention to the diets being consumed by the social media users.

For this analysis, we selected 500 users randomly from the Twitter user-id space (*i.e.,* the user-ids were randomly selected from the range 1 through the id assigned to a newly created account), with the constraint that the selected users follow at least 20 other users (to ensure that the selected users have a meaningful consumption behavior to study). We then computed the consumed information diet for each user, considering the tweets that a user received from her followings (*i.e.,* via word-of-mouth) during December 2014.[4]

Similar to the previous section, we define the top topic for a user as the topic on which they consume the largest fraction of their diet. For the group of users having a common top topic of consumption, Figure 3.9 plots the mean contribution of the top topic in the consumption diet of these users.[5] As a

[4]We consider all tweets received by a user to compute their consumption diet in the absence of data about what they read.

[5]In our set of 500 randomly selected users, we did not find any user whose top topic of consumption was society, hence we will not consider this topic further in this section.

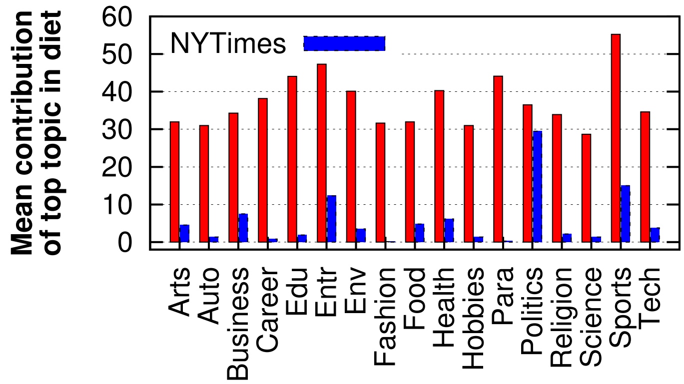


Figure 3.9: Mean contribution of the top topic in the consumption diets (on which a user consumes the highest fraction of their diet) of random users grouped according to their top topic of consumption.

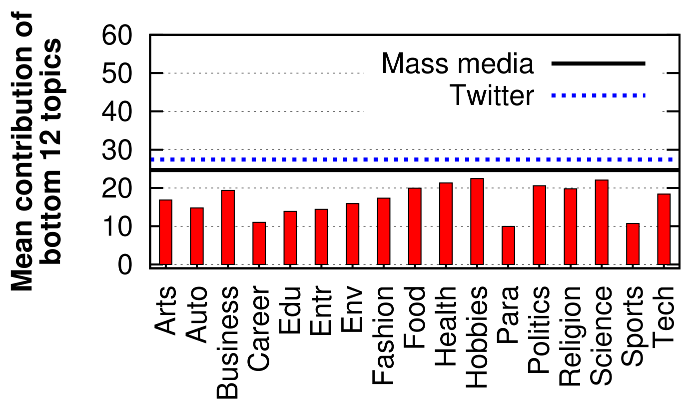


Figure 3.10: Mean contribution of the bottom 12 least dominant topics in the consumption diets of random users grouped according to their top topic of consumption.

baseline, the figure also shows the contribution of each topic in the NYTimes mass media diet. Across almost all topics, the consumers are very focused on their top topic, and on average, consume 30% or more of their diet on that topic. Moreover, when we compute the contribution of up to top two topics, we observe that 80% of the users consume more than half of their diet on only these one or two topics. These observations imply that users in social media consume a much larger fraction of their information diets on their primary topic(s) of interest, as compared to what they would consume on the same topics from a typical mass media source (as shown by the NYTimes mass media baseline).

Additionally, Figure 3.10 depicts the mean contribution of the *bottom 12 topics* on which the users consume the least information, for the same groups of users. We find that the 'tail topics' account for an inordinately low fraction of their consumed diet. Across all topics, the mean tail topics contribution

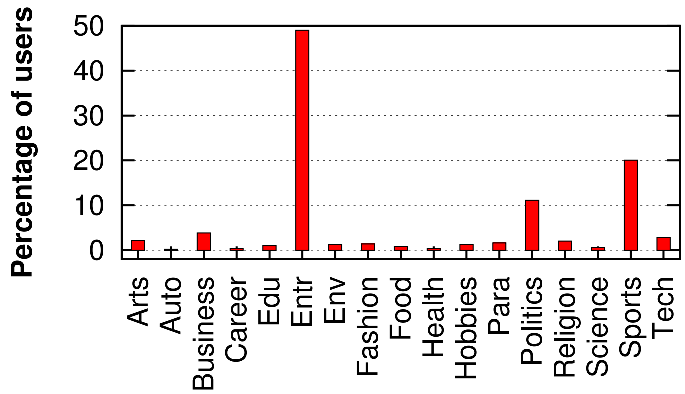


Figure 3.11: Distribution of the 500 randomly selected users, according to the topic on which they consume the maximum fraction of their diet.

for users focusing on a particular topic is even lower than the contribution of the bottom 12 topics in the NYTimes mass media diet (24%) and the Twitter random sample diet (27%).

Finally, Figure 3.11 plots the distribution of the 500 randomly selected users according to their top topic of consumption. We find that the users' consumptions are very unevenly spread amongst the different topics – as much as half the user population consumes most information on the topic entertainment, while a sizeable fraction focuses on sports and politics. When we compare this distribution to the production distribution of popular users in Figure 3.7, we observe that consumption behaviors are even more skewed across topics than production.

Thus we observe that users are extremely selective in the information they consume via social media, with a considerable bias towards one or two topics of their interest; moreover, this bias comes at the cost of the tail topics. In future, as users rely more and more on social media like Twitter to consume information, their diets may get progressively more skewed towards the one or two topics of their interest. Users who wish to have a more balanced consumption on social media need to be careful about the sources to which they subscribe. Alternatively, the biases in the consumption diets of users can potentially be mitigated by the information supplied to them by recommender systems deployed on the social media sites. In the next chapter, we will investigate the role of recommender systems in shaping the diets of social media users.

3.1.5 Summary: Topical diversity in information diets of social media users

Our study of the topical diversity in the information diets of social media users revealed many interesting findings. We found that mass media sources cover a wide range of topics from politics and business to entertainment and health. But on social media, the individual sources of information are very focused and publish information dominated by a few topics. Therefore, it is up to the social media users to select sources to obtain a balanced diet for themselves. Furthermore, we find that for most users, a significant fraction of their consumed diet comprises of as few as one or two topics, and they hear very little about other niche topics like health and environment (unless they are interested in these topics).

Our work and findings have a number of significant implications. With the increasing popularity of social media, it is vital to raise awareness about the balance or imbalance in the information diets produced and consumed on social media. Our findings raise the need for better information curators (human editors or automated recommendation systems) on social media that provide a more balanced information diet. Finally, our work is an early attempt, and much future work remains to be done both on understanding the impact of the diets on consumers in shaping their opinions and designing mechanisms for helping users to have a more balanced diet.

3.2 Geographical source diversity of information diets

In this part of the thesis, we focus on studying the information diets of social media users along the dimension of geographic source diversity. Here we consider a country as a geographical unit, and instead of studying individual user's diets, we investigate the aggregated diet that all the users within a country are consuming as a whole. In particular, we are interested in answering the following questions:

1. **Production:** What is the geographic diversity of the sources of information from which different countries are consuming? How are the information sources – Twitter users and elites – spread across the world?

2. **Consumption:** How diverse are the consumption channels for different countries? Is there geographic source diversity in the diets of different countries? What is the relationship between a country and the greatest contributing countries to its consumed diet?

We begin by giving a brief overview of the related work in Section 3.2.1, and then continue to describe our methodology for inferring the geolocation of users on Twitter social media in Section 3.2.2. Finally, we present the analysis of geographic source diversity in the diets of countries, where we start by analyzing the geographic diversity of sources on Twitter, followed by an investigation of the geography of consumption channels, and end with the exploration of geographic source diversity in the diets of different countries.

3.2.1 Related work

Prior work related to our study arises from two main areas: (i) inferring or predicting geolocations of social media users, and (ii) analyzing the impact of geography and language on online user interactions. Next, we outline some of the related work in these two areas.

3.2.1.1 Inferring geolocations of social media users

Various techniques have been explored to infer geolocations of social media users. Hecht *et al.* [94] used map APIs to resolve location field data provided by the users as part of their profile information and found that many users do not provide real location information. Java *et al.* [109] used Yahoo! Geocoding API [6] to resolve the location string entered by users, while Krishnamurthy *et al.* [123] used the UTC offset field in the users' profiles to infer their location.

Others have tried to predict the location of users who do not provide their profile information, either based on the location of the users' neighbors in the social graph [200, 21], or based on the content of their tweets [48, 143], or a combination of the two [56]. Lieberman and Lin [137] leveraged Wikipedia edit histories to determine the location of Wikipedia users. Since they leveraged features specific to Wikipedia, their method is not very generalizable to other platforms. Popescu and Grefenstette [186] utilized the place names that users tag photos with on Flickr to infer the home location of Flickr users.

In this work, we rely only on the profile information provided by the users themselves, since we found it to be sufficient for inferring the country level location information for a considerable fraction of all Twitter users in our dataset.

[6] See **http://developer.yahoo.com/maps/**.

3.2.1.2 Analyzing the impact of geography and language on online user interactions

There is a growing interest amongst researchers to understand how offline boundaries (*e.g.,* geographic, linguistic, national, and cultural boundaries) impact users' interactions in the online world. Some recent studies have analyzed the geographic distribution of Twitter users, albeit on small datasets consisting of tens of thousands of users. Java *et al.* [109] and Krishnamurthy *et al.* [123] examined and discovered differences between the properties and growth of the networks of Twitter users in different geographic regions (like North America, Europe, South America and Asia-Pacific) and continents. More recently, Takhteyev *et al*. [222] found that geographic distances, national boundaries, and languages hold considerable influence on the formation of social ties on Twitter. Hong *et al.* [104] studied the differences in usage patterns between different language communities on Twitter. Mocanu *et al.* [154] studied the language geography using Twitter social media, and reported on the language usage of different countries, as well as distribution of languages in multilingual regions, while Kim *et al.* [118] studied the multilingual societies on Twitter and the role that bilingual users play as "bridges" between different societies. Multiple studies [93, 202, 201] have also investigated different Wikipedia language editions to study the linguistic and cultural patterns of usage and editing.

Similar to these prior studies, our current work shows that both linguistic similarity and geographical proximity play a significant role in shaping the users' online interactions and diets. Compared to these previous studies, our work presents a considerably more detailed examination of how geolocations of users impact their participation, connectivity and information diets, using a significantly larger dataset containing tens of millions of users.

3.2.2 Methodology: Quantifying geographical information diets

In this section, we first describe the Twitter dataset we used in this study, followed by the methodology that we used to infer the geographical locations of users. We end by briefly outlining the methodology for constructing information diets of countries from the source geolocations.

3.2.2.1 Twitter dataset

For studying the geographical source diversity in the information diets consumed by different countries, we use the Twitter dataset described in [45]. The dataset includes the profile information of 51.9 million

	Bing & Yahoo	Yahoo & time zone	Bing & time zone	At least 2 match
Overlap	10.58 M	12.24 M	10.19 M	12.86 M
Match	9.78 M (92.4%)	10.85 M (88.7%)	8.99 M (88.2%)	12.22 M (94.5%)

Table 3.8: Match between the different sources for geolocation resolution.

users, and their 1.9 billion follow links, based on the snapshot of the network taken in September 2009. The dataset also contains the 1.7 billion public tweets posted by these users from the launch of Twitter in March 2006 till September 2009.

3.2.2.2 Inferring users' geolocations

In this study, we focus on inferring location information for Twitter users at the granularity of countries. For inferring a user's country of residence, we make use of two fields in their profiles: (i) the location field (free-text string entered by the user), and (ii) the timezone field (selection made by the user from a drop-down menu, consisting of location name and a UTC offset).

Out of the total 51.9 million users, $13,148,002$ (25.3%) users filled in the location field. For them, we use public map APIs provided by Yahoo Maps[7] and Bing Maps[8] to convert the free-text string entered by the users into countries. We could do the conversion for $10,709,638$ (81.5%) users using Bing Maps, and $12,908,671$ (98%) users using Yahoo Maps. Furthermore, out of the 51.9 millions users, $19,365,683$ (37.3%) users provided timezone information, which we converted to corresponding countries.

Prior work has suggested that location inference using individual map APIs can be error prone [94]. Therefore, we compared the results obtained using the two map APIs and the timezone, to minimize inference errors. Table 3.8 shows the number of users that were common between the sets of users whose location information was successfully resolved using each of these three sources. We also show the fraction of these overlapping users for whom the inferred locations matched. We find a high agreement in the resolved country name between any two of the three sources.

[7] See **http://developer.yahoo.com/geo/placefinder/**.

[8] See **http://www.microsoft.com/maps/developers/web.aspx**.

To minimize the inference errors from the three sources (the two map APIs and timezone translation), we only consider the set of users for whom two of the three sources match in their resolved location. Doing so, we obtain 12.2 million users which account for 23.5% of all users in our dataset. These users are distributed across 231 countries and account for 73.65% of all tweets posted and 37.6% of all social links in the network.

3.2.2.3 Evaluating geolocation inference methodology

For inferring users' geolocations, we are relying on the users themselves to provide correct location information. However, prior work [94] has reported that for 19.5% of the users they could not correctly infer the geolocation because either the users had entered non-geographic information in the free-text location string or the map APIs had not returned the correct result.

Therefore, for evaluating our methodology, we take a sample of 1000 randomly selected users from our final set of 12 million users and manually examine the timezone and location string entered by them and judge whether they were correctly resolved to the corresponding country. In 94.7% of cases, the country resolution was judged to be correct, while out of the remaining, 4.4% users had entered non-geographic information in their location string, while for 0.9% of the users the map APIs had resolved the country incorrectly.

3.2.2.4 Constructing information diets from source geolocations

To construct the information diet being consumed by a country C, we consider the tweets being consumed by each user in the country C from the users they follow.[9] For each tweet being consumed by a user in the country C, we determine the geographic location of the source. Aggregating across these tweets, we obtain a vector of countries where the weight of each country denotes the normalized contribution of that country's sources (via the follow links between the two countries) in the consumed diet of C. The information diet consumed by the country C is given by this country vector.

[9] As before, we consider all tweets received by a user to compute their consumption diet, in the absence of data about what they read.

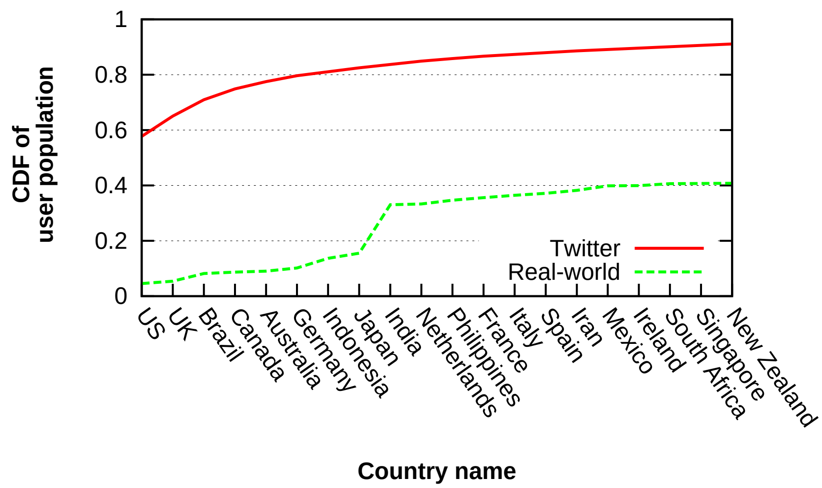


Figure 3.12: Distribution of Twitter population and world population of 20 countries with the most Twitter users.

3.2.2.5 Limitations

Since users from different countries may not have the same probability of sharing their location information, it is possible that our set of users and hence the information diets of different countries may be biased towards certain countries whose users tend to share their location on social media more. Another source of potential bias is the fact that our dataset is from 2009, and therefore may not capture the current distribution of Twitter users. And finally, for inferring users' locations, we are relying on the information that users themselves are providing, and biases may also creep in due to this self-reportage.

3.2.3 Production: Are the sources geographically diverse?

Before discussing the geographic source diversity of diets that different countries consume on Twitter, we first analyze the geographic diversity in the production of information on Twitter, *i.e.,* geo-distribution of sources or users on Twitter. In particular, the question we answer with our analysis is "how are Twitter users and elites spread across the world?".

3.2.3.1 Geo-distribution of Twitter users

The 12 million Twitter users in our dataset, for whom we successfully inferred location information, are spread across 231 countries worldwide. The number of Twitter users varies considerably across the different countries, with only a small number (13) of countries with 100,000 or more users, while a large number (167) of countries have 10,000 or fewer users. Not surprisingly, the top few countries account for a vast majority of the total Twitter population.

In Figure 3.12, we show the skew in Twitter population towards a few countries, by plotting the cumulative distribution function (CDF) of the Twitter users from the 20 countries with the most Twitter users. The US, the country with the highest number of users, by itself accounts for 57.7% of the total Twitter population in our dataset. The top 10 countries alone account for 84.9% of the whole Twitter population, while the bottom 80% of countries only account for 2.3%.

Interestingly, the top countries account for a significantly higher fraction of the Twitter population than the share of the world population living in those countries [2] (also shown in Figure 3.12). The difference between the two curves exemplifies the geography-based digital divide in today's world, where users outside of a small number of developed and developing countries have limited reach to online services like Twitter [173].

So far our analysis of geolocations of Twitter users has been limited to a snapshot of the population in 2009. As Twitter adoption grows worldwide, one would expect the adoption rates to change over time. We analyzed the temporal evolution of Twitter user population by studying several snapshots of the network during the period from 2006 to 2009. While the number of users in each country increased considerably during this period, our observations about the skew in Twitter service adoption towards a small number of countries held true at all times.

3.2.3.2 Geo-distribution of elite Twitter users

Not all users on Twitter are equal. Studies have shown that a small number of Twitter users – *elites* – account for a disproportionately large number of followers and tweets consumed on Twitter. For instance, Wu *e al.* [250] have shown that roughly 50% of the URLs consumed are generated by just 20K elite users (*i.e.,* 0.05% of all users). Such influential users in the network can be detected using ranking methods such as PageRank or FollowerRank [131].

We now focus our attention on the distribution of elite Twitter users (identified as users with highest PageRank) across different countries, as depicted in Figure 3.13. We observe that the distribution of elite users across the countries is even more skewed than the distribution of Twitter users themselves. For example, if we consider the top 0.1% of users with highest PageRank in our dataset, then 80.7% of them are in the US, which is much higher than its 57.7% share of the total Twitter population. The

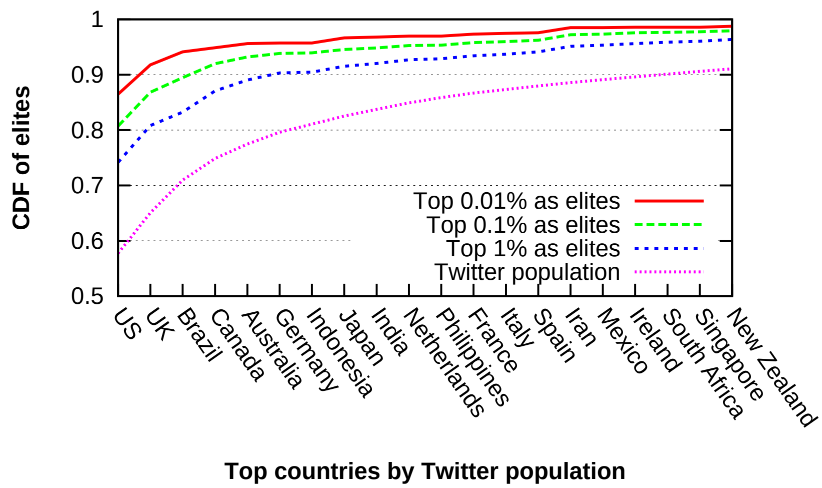


Figure 3.13: Distribution of number of elites in the 20 countries with the most users.

ten countries with the most users account for more than 95% of the top 0.1% elites, even though they represent only 85% of the user population. Therefore, our results indicate that the digital divide is even more massive amongst the elite users. They also suggest that for building location-specific search or recommendation services, global ranking algorithms might not be sufficient as they would ignore local elites, *i.e.,* we would also need a local ranking scheme.

3.2.4 Consumption: Is there geographic source diversity in the diets that different countries consume?

For investigating the consumption of different countries, we only consider the 100 countries with the most users for our analysis, since the remaining countries have too few users (less than 1000) in our dataset to extract meaningful and representative information.

3.2.4.1 Geography of consumption channels

Before delving into the geographic source diversity in the consumed diets of different countries, we examine the geography of the consumption channels for the different countries, *i.e.,* the social (followee) links between the users of different countries.

Transnational vs. intra-national links: On Twitter, 35.15% of all social links are transnational, *i.e.,* they connect a follower and a followee that are located in different countries. The percentage increases to 37% when we exclude the US, which accounts for a majority of users and links in the Twitter network.

Country	% of Trans-national Followings	% of Intra-national Followings	% of Total users
India	82.28%	17.72%	1.21%
Canada	79.84%	20.16%	3.91%
Australia	78.57%	21.43%	2.62%
Indonesia	73.19%	26.81%	1.46%
UK	69.79%	30.21%	7.33%
Netherlands	62.42%	37.58%	1.16%
Germany	62.26%	37.74%	2.12%
Brazil	32.9%	67.1%	5.9%
Japan	26.41%	73.59%	1.45%
US	18.44%	81.56%	57.74%

Table 3.9: How much do countries rely on other countries for the information they consume?: The fraction of trans- and intra-national following links for the 10 countries with the most users (ranked by their fraction of transnational followings).

Thus, even as a majority of social links stay within national boundaries, a considerable fraction (more than a third) of all links cross national boundaries, highlighting the global nature of consumption channels in the Twitter network.

However, the fraction of transnational links varies considerably from country to country, with different countries relying to different extents on other countries for the information they consume. For the 10 countries with the most number of users, Table 3.9 shows this reliance as the fraction of their transnational and intra-national following links along with their share of total users. There are two striking takeaways from this table. First, even amongst the top-10 countries, the fraction of trans-national links varies from as high as 82% in some countries to as low as 18% in others, suggesting that users in some countries seek information from around the world, while those in others look for information primarily from their compatriots. In the former category, we have countries like India, Australia, Canada, Indonesia, and the UK, with more than two-thirds of their following links going to users in other nations. At the other end, users in the US, Japan, and Brazil have more than two-thirds of their links remaining within their national boundaries. Netherlands and Germany lie in the middle with a more even division between national and transnational links. Thus, users in some countries have much more global consumption channels than others.

Second, comparing the fraction of intra-national links for countries with their share of the Twitter population, we observe that there is a significant bias towards following other users from the same

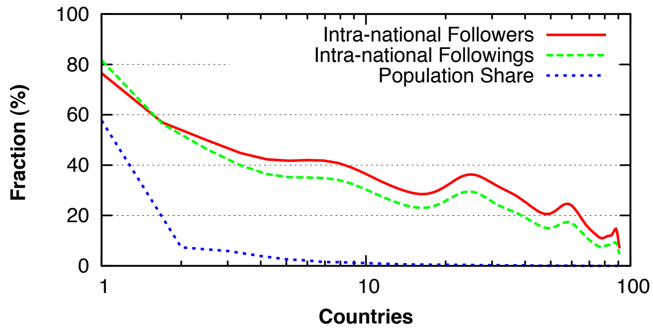


Figure 3.14: Fraction of intra-national followers and followings in comparison to the Twitter user population share in different countries, ranked by their Twitter user population.

country to get information. For example, 37.74% of all users followed by German users are from within Germany itself, even though German users account for only 2.12% of the total Twitter population, which suggests that German users prefer to follow and consume information from other German users almost 18 times more than users elsewhere. Figure 3.14 plots the fraction of intra-national followers and followings for the different countries in our dataset along with their share of user populations. The figure shows a clear bias towards intra-national links for users in all the countries. The ratio of the percentage of intra-national links to the percentage of user populations is very high across the different countries; average ratio across all countries for following links is 1085, and for follower links, it is 756.5. Thus, even as users connect to others globally, they also exhibit a significant preference for connecting to local users for consuming information.

Figure 3.14 also shows that for most of the countries the percentage of intra-national followers is slightly but consistently higher than intra-national followings, with the US being an exception. The higher percentage of intra-national followers suggests that there is less global demand for information from users in countries outside the US than there is demand for global information from users within those countries. This imbalance could be potentially explained by the relatively large fraction of elite users within the US as shown in Section 3.2.3. Users in other countries follow these elite users in the US to consume information, but the countries themselves contain few elites, leading to lower demand for consumption channels (*i.e.,* follower links) from outside of them.

Country	Closest 5 Followers	Closest 5 Followings
Chile	Argentina, Bolivia, Ecuador, Peru, Uruguay	Argentina, Bolivia, Ecuador, Spain, Uruguay
Egypt	Lebanon, Morocco, Saudi Arabia, Tunisia, UAE	Kuwait, Lebanon, Morocco, Qatar, Saudi Arabia
Japan	China, Hong Kong, South Korea, Taiwan, Vietnam	China, Hong Kong, Jamaica, South Korea, Taiwan
Russia	Belarus, Greece, Latvia, Lithuania, Ukraine	Belarus, Estonia, Greece, Latvia, Ukraine
Spain	Argentina, Bolivia, Ecuador, El Salvador, Uruguay	Argentina, Bolivia, Ecuador,Mexico, Uruguay
US	Australia, Canada, Nepal, New Zealand, Pakistan	Australia, Canada, New Zealand, Singapore, UK

Table 3.10: Closest 5 follower (who seek information from them) and following (from whom they seek information) countries for a few example countries around the world.

Impact of geography & language on consumption channels: We now focus on the consumption channels between different pairs of countries to investigate whether users from a country preferentially seek information from other countries that are geographically or linguistically close to this country.

To conduct our analysis, for each country, we ranked all other countries based on the density of consumption channels between them – *i.e.,* how *closely* their users followed (or were followed by) users in the other countries. We computed the closeness of a country A with another country B based on the number of channels (both followers and followings separately) that go between the countries, normalized by the number of users in country B.

Table 3.10 shows the top-5 closest follower and following countries for a few countries around the world. We make two observations: first, while the top-5 closest follower and following countries are not the same, there is considerable overlap between the lists. In fact, when we compared the lists of top-10 closest countries according to the follower and following links there was, on average, an overlap of 75.8%. Second, for some countries, such as Japan, the closest countries correspond to geographical neighbors in East Asia, while for others, such as Spain, the closest countries are geographically distant countries in South America that share the same language. Thus, both language and geography appear to play a role in determining the consumption channels that users from different countries choose for themselves.

Type of neighbors	Closest 5 followers	Closest 5 followings
Linguistic	37.58 %	38.46 %
Geographic	55.16 %	55.16 %
Continent	74.73%	70.99 %
Linguistic or geographic	72.53 %	73.41 %
Linguistic or continent	90.11 %	87.25 %

Table 3.11: Percentage of closest follower and following country pairs that share a geographic boundary, or a common language, or lie within the same continent.

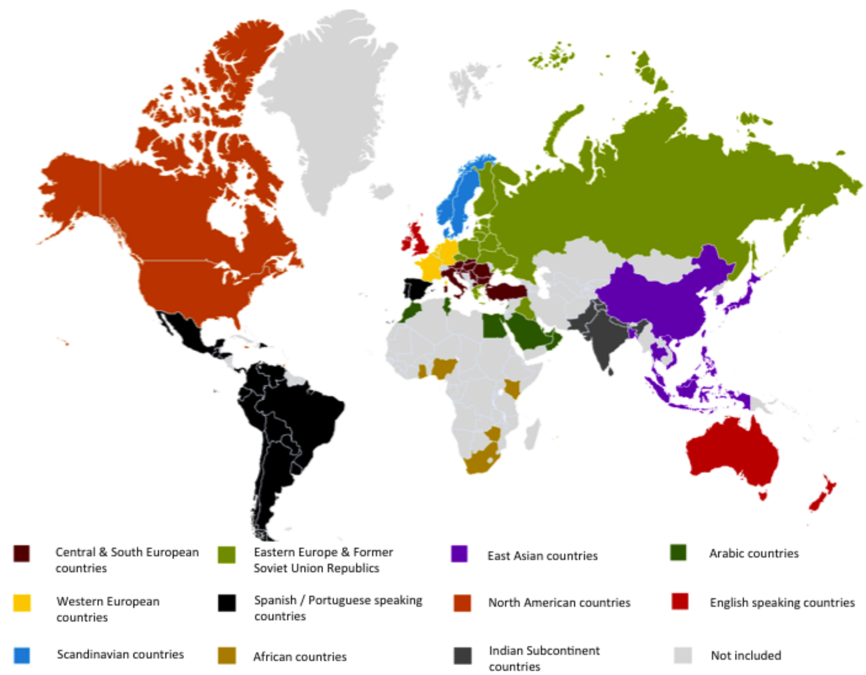


Figure 3.15: Groups of countries whose users are closely connected with one another via consumption channels.

We further investigated the impact of geography and language by computing the percentage of top-5 closest pairs of countries that are geographical neighbors (share a border or lie within the same continent) or linguistic neighbors (share a common language). Table 3.11 shows the results for pairs of top-5 closest follower and following countries. The percentages for both closest follower and following countries are similar. They show that a vast majority of closest countries are geographical or linguistic neighbors: 55% of closest pairs of countries share a common border, while 38% share a common language. In fact, 73% of countries share either a boundary or language, indicating that both language and geography influence the consumption channels between countries.

Lastly, we use the closeness rankings to create a consumption graph between countries, where each country is connected to its closest five follower or following countries. We then applied the Louvain method for community detection [32] to detect closely interconnected groups of countries within the consumption graph. Figure 3.15 shows the country groupings resulting from the graph of closest five following countries on a world map, though we got similar results when we used closest five follower countries instead.

Figure 3.15 shows that the 91 countries in our dataset fall into eleven distinct groups of countries. These groups correspond strikingly with well-recognized geographic, linguistic, political, and cultural groupings of countries in the offline world. For example, the East Asian countries such as China, Vietnam, Thailand, South Korea, and Japan form a grouping distinct from countries in the Indian sub-continent, such as India, Pakistan, Sri Lanka, and Nepal. Similarly, Arabic speaking countries in the Middle East and North Africa, such as Egypt, Tunisia, Morocco, Saudi Arabia, UAE, and Qatar form one group. Interestingly, the western European countries of Spain and Portugal are grouped with Spanish- and Portuguese-speaking countries in South and Latin America, such as Argentina, Brazil, and Mexico. Similarly, the Scandinavian countries of Sweden, Norway, and Denmark form their own group distinct from other western European countries such as France, Germany, Belgium, and the Netherlands. While eastern European countries like Poland and the Czech Republic are grouped with countries that were formerly republics of Soviet Union, the central and southern European countries like Austria, Hungary, Greece, and Romania are grouped with Turkey.

The existence of these eleven distinct groupings of countries corresponding to well known national, political, linguistic, and cultural boundaries underscores the importance and influence of these offline factors on the formation of consumption channels in the online world.

3.2.4.2 Geographic source diversity in consumed diets of countries

Having studied the geography of consumption channels, we next focus our attention on the geographic source diversity in the consumed diets of countries. For this analysis, we use the 41 million tweets that were posted in a week at the end of May 2009 to construct the consumed information diets of the different countries.

Across the different countries, we find that on an average, the same country accounts for about 62.46% of the consumed diet, *i.e.,* roughly two-thirds of the tweets consumed by a country are produced

Country	Contribution of other countries in the consumed diet
US	19.78 %
UK	74.02 %
Brazil	42.8 %
Canada	85.82 %
Australia	83.57 %
Germany	71.2 %
Indonesia	89.32 %
Japan	16.09 %
India	83.11 %
Netherlands	62.03 %

Table 3.12: Contribution of other countries in the consumed diets of the 10 countries with the most users.

locally in the same country. However, a non-trivial fraction (37.54%) are produced internationally in other countries, indicating that there is some geographic source diversity in the consumed diets.

When we consider the top 10 countries with the most number of users in our dataset, we find that there is a considerable variation in the contribution of other countries in their consumed diets, as depicted in Table 3.12. On one end of the spectrum, there are countries like Japan (16.09%) and the US (19.78%) who consume very little from outside of their own countries, and on the other end are countries like Indonesia (89.32%), Canada (85.82%), Australia (83.57%) and India (83.11%) which rely heavily on other countries for their information consumption. Therefore, some countries demonstrate higher geographic source diversity in their consumption diets than the others.

3.2.5 Summary: Geographical source diversity of information diets

While investigating the geographic distribution of sources on Twitter social media, we observed that a few countries accounted for a vast majority of all users, and an even more significant fraction of elite users. Furthermore, our analysis revealed that though the country itself accounts for a substantial fraction of its consumed diet, on average a third of the diet is contributed by other countries. Therefore, we observe that on an average most countries' consumption channels, as well as, consumed diets consist of a fair amount of geographic source diversity, with the top contributors being either geographically close or linguistically similar countries – emphasizing the role these offline boundaries play in shaping the online information diets.

3.3 Conclusion

In this part of the thesis, we examined the impact of the paradigm shift from traditional mass media to online social media by studying the word of mouth diets being produced and consumed. We first discussed the topical diversity in the word of mouth diets of social media users. We observed that not only do social media users (including news organizations) produce diets focussed on just one or two topics, but the diets that users are consuming are even more skewed. Next, we explored the geographic source diversity in the diets of social media users, and again we observed that a considerable proportion of the consumed information is contributed by the country itself, or a few geographically close or linguistically similar countries.

The focused nature of produced diets on social media has lead to an unbundling of content production where each source is a very specialized producer. However, we observe that users do not do a very good job of following multiple different sources to stitch together a topically or geographically balanced diets for themselves. Therefore, in the next chapter, we explore whether the recommendations provided by social media platform providers mitigate or exacerbate the imbalances in the users' consumed diets.

CHAPTER 4

Recommendations: Impact on diversity in consumption

Given the deluge of user-generated content on social media platforms, personalized recommendation systems are today being deployed on all popular social networking sites such as Twitter, Facebook, and YouTube, to enable users to find interesting users and content [10, 7, 88] easily. Though the exact recommendation algorithms (which are not publicly known) deployed in various OSNs may be very different, all of them largely depend upon the *social neighbourhood* of target users for finding interesting items to recommend to them [87, 10, 7, 88]; hence these are also known as *social recommendation* algorithms.

In the previous chapter of the thesis (Section 3.1), we observed that for most users just one or two topics of their interest constitute a majority of their consumed diets on Twitter. In this chapter, we study the impact of recommendations (given by Twitter) on the information that users are exposed to by answering the question, "*do the recommendation systems exacerbate or mitigate the topical imbalances in the consumed diets of social media users?*".

The primary challenge for studying the impact of Twitter's recommendations on the information diets of users is that these recommendations are personalized for each user and can not be publicly crawled as these are only visible to the users. To overcome this challenge for data collection, we created test accounts on Twitter which mimic randomly selected real users on Twitter by following the same set of users as the real users. Using this methodology, we can approximate the recommendations for these real users by collecting the social recommendations provided to our test accounts. We show that the recommendations given to users on Twitter are predominantly from their 2-hop neighborhoods and the more popular the tweet is within the 2-hop neighborhood, the higher the chance of it being recommended to the user. Extensive analysis of the different factors that contribute to the recommendations on Twitter is out of scope of this study – our goal is not to reverse-engineer the recommendation algorithm, but to estimate its impact on the users by evaluating what information they serve to the users on top of the information that users are getting directly from their followings.

We constructed the recommended diets of the test accounts by inferring the topics for the tweets recommended to them, and also computed the consumed diets for them using the tweets they receive via their follow-links. We compare the consumed and recommended diets of social media users to determine whether the recommendations mitigate or exacerbate the topical imbalances in users' consumed diets.

We find that social recommendations, *i.e.,* recommendations about information popular in a user's social network neighborhood [7, 88], often do not match the user's consumed diets. The differences between recommended and consumed diets are likely due to differences in the interests of users and the interests of their network neighbors. As a result, social recommendations introduce topical diversity to a user's diet and can help balance its topical composition by mitigating the imbalances.

Our research contributions in this work can be summarized as follows:

1. We developed a methodology for collecting and characterizing personalized social recommendations by creating test accounts on Twitter which mimic real users by following the same set of users like them. Using this method, we could construct both the consumed diets as well as the recommended diets of social media users (Section 4.2).

2. We compared the topical diversity in the consumed diets of the test accounts with their recommended diets to evaluate the impact of recommendations on the diversity in consumption. We found that social recommendations on Twitter mitigate the imbalances in the users' consumed diets by exposing them to a more diverse set of topics.

We envision that our work will create awareness among social media users about imbalances in their information diets and how the algorithmic systems like recommendations are impacting them. Additionally, we also anticipate that our findings have implications for the designers of future algorithmic systems on social media for discovering and recommending information.

4.1 Background & related work

There has been a lot of research over the years towards developing personalized recommender systems. Some popular approaches for personalized recommendation include *collaborative filtering* [101, 198] and *content-based approaches* [179]. Especially, in collaborative filtering, a user's preferences are compared to those of all other users to identify other users who have similar preferences to the given user, and the items that are liked by those users are recommended. The approach of social recommendations can be said to be a special case of collaborative filtering, where the set of other users is limited to the network neighborhood (*i.e.,* friends or friends of friends) of the given user.

In this section, we begin by giving a short background on Twitter's personalized recommendations. Having provided the context, we briefly describe the different evaluation metrics that have been used to evaluate recommendation systems. We end by discussing the impact of personalized recommendations on the diversity of news and information that users get and how this impact can be measured.

4.1.1 Background: Twitter's personalized recommendations

The 'Discover tab' feature [231, 133] in Twitter gave personalized user [87] and tweet (content) recommendations to every user.[1] Though the exact algorithms used by Twitter are not known publicly, the recommendations are known to be based on the accounts whom the target user follows and who these immediate neighbors follow in turn [231]. Twitter attempts to "identify your connections and rank them according to how strong and important those connections are to you", and then recommends the content posted by the strong or important connections of the target user [231]. However, the exact methodology for determining the strength or importance of the connections is not specified by Twitter. Prior attempts [46] at developing a personalized system for recommending interesting tweets to a user found that such content can be effectively found in the 2-hop social neighborhood of the user (*i.e.,* the tweets posted by the followings of the followings of the user). Since these methods rely on the social neighborhood of target users (*i.e.,* the users the target user follows, and the users they follow in turn), they are often referred to as *social recommendations* [224].

[1] Since Twitter's Discover tab feature has been discontinued; we refer the readers to a Mashable article [229] which delineates its different functionalities, to learn more about it.

4.1.2 Evaluation metrics for recommender systems

A large spectrum of criteria and metrics have been proposed and employed to evaluate recommendations systems in the past and significant research effort has focused on how to do this evaluation in an effective manner [102, 210]. Traditionally, the recommendation systems community has regarded *relevance* as the most important metric. Relevance measures whether the recommendations given to users are relevant to their topics of interest. However, more recent studies [139] have postulated that several other metrics are equally important for evaluating the performance of recommendation systems, such as *coverage* [78] , *popularity* [218], *novelty* [236, 43], *serendipity* [121, 34, 78], and *diversity* [265, 266] and we next describe them briefly.

Coverage [78] is usually measured as the degree to which recommendations cover the set of all available items. Novelty [236, 43] in recommendation systems refers to introducing the target user to items which might be different from their existing preferences and interests, but which could still be of interest to them [43]. The related concept of serendipity [121, 34, 78] has no one agreed upon definition, though it is related to the novel recommendations such that the users are positively surprised by the inclusion of the unexpected items. While diversity [265, 266] and serendipity are related concepts, diversity is typically measured as the differences in the content being recommended. Kunaver and Porl [129] have written a survey about diversity in algorithmic recommendations, which covers all three aspects of it – defining and evaluating diversity, studying the effect of diversity on recommendations, and integrating diversity in algorithm design.

Since in this thesis we quantify the impact of recommendations on the topical diversity of the information diets of social media users, before going further, we briefly discuss diversity in the context of recommendation systems.

4.1.3 Impact of recommendations on diversity

Diversity in the topics and issues being discussed in a society is considered to be one of the most critical conditions for a public discourse in modern societies and a core public value in media law and policy [155, 66, 99]. Council of Europe (2007) has stated that "media pluralism and diversity of media content are essential for the functioning of a democratic society" and the right to freedom of expression "will be fully satisfied only if each person is given the possibility to form his or her own opinions from diverse sources of information" as guaranteed by Article 10 of the Convention for the Protection of Human Rights and Fundamental Freedoms [12] [2]. Given the importance of diversity in information diets of users, researchers have encouraged governments [97] and public service media [216] to play a role in integrating diversity into the recommendation systems by design to help people to choose a more diverse diet for themselves.

Traditionally, recommendation systems literature has typically focused on diversity as a design element for the algorithm. Prior studies [267, 247, 120] have demonstrated that user satisfaction can be increased by bringing diversity into the recommendation sets.

As shown in the previous chapter (Section 3.1.4) users often choose to focus their attention selectively on only a few sources that reflect their preferences and topics of interest. This phenomenon is also called selective exposure [208, 219, 111, 75]. Additionally, with personalized recommendations becoming more popular and playing the role of gatekeepers of information and news in the society, some have feared that these personalized recommendations may reinforce the users' preexisting preferences by showing them more of the same perspective as their own and limiting their exposure to other perspectives and ideas, and thereby trap them in filter bubbles [177].

Social scientists and more recently a broader spectrum of researchers concerned about the impact of recommendation algorithms on the society are now investigating diversity in the output of personalized recommendation systems [216]. Eskens *et al.* [66] examined how news personalization affects the right to receive information of news consumers, along five perspectives – political debate, truth-finding, social cohesion, avoidance of censorship and self-development, and what policy choices need to be made to counter the ill-effects. Moeller *et al.* [155] examined the impact of personalized news use on issue agenda diversity by using data from a representative set of the Dutch population and found that personalized news does not lead to a smaller core of issues that are discussed. They also found that younger and more educated news users make use of news personalization more and are interested in niche topics at the fringes of common issues agenda. Similarly, empirical examinations of the filter bubble phenomenon [177] have shown that though real, the effect is often rather small in magnitude [23, 75]. Recommendations can, in fact, diversify the set of recommended articles as compared to recommendations by editors, or popularity based recommendations [171, 108, 156, 262].

[2] Article 10, Paragraph 1: "Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authorities and regardless of frontiers."

There are two main perspectives from which diversity can be studied in the context of recommendations: (i) diversity of supply can be measured as the source or content diversity of the information being published, and (ii) exposure diversity is the diversity that users are choosing to expose themselves to through active or passive choices. The diversity that is being published is not the same as the diversity that is being consumed by people. We believe, it is not sufficient to study diversity of supply (reported in Section 3.1.3, which examines the topical diversity in information produced on social media), but it is also important to study the exposure diversity, *i.e.,* what are users actually consuming [169, 99, 39, 170]. In the previous chapter (Section 3.1.4), we took a step in this direction by measuring the topical diversity in the consumed diets of social media users, while in this chapter, we continue in this direction by examining the topical diversity in the recommended diets of users.

Some studies have looked into measuring the diversity of the recommendations. However, recommendation algorithms are today evaluated in isolation, whereas they should be evaluated by taking into consideration their interplay with the word of mouth propagation. Hence, we propose to evaluate the diversity of recommendations as of the *additional* benefit they provide a target user over and above what they consume directly through word of mouth. In our study, we find that social recommendations on Twitter increase the diversity in the information diets of users by exposing them to a more heterogeneous set of topics as compared to their consumed diets.

4.2 Data collection & methodology

One of the main challenges for studying recommendations on most social media platforms, including Twitter, is that the recommendations provided to an individual user are visible only to them and cannot be crawled publicly. Since the recommendations provided on many social media platforms, including Twitter, are *social recommendations* which rely on the accounts the user follows and the accounts whom they follow to identify what to recommend to the users [231]. To overcome this challenge, we adopt the methodology of creating *test accounts* on Twitter which mimic real users, *i.e.,* the test accounts follow the same users that the mimicked real users follow and thus have the same network neighborhood as them.

4.2.1 Experiment design

Our basic methodology is to create test accounts which mimic real Twitter users by following the same set of people as the real users. However, there are several issues which one needs to be careful about while adopting this methodology; we discuss these issues in this section.

Mitigating the effect of other factors affecting personalized recommendations: Twitter's personalized recommendations to a user can potentially be impacted by a number of personalization factors other than the social network neighborhood, such as the user's profile features (*e.g.,* location, language, declared interests), their IP address for accessing Twitter, their activity (*e.g.,* their tweets or retweets, or new users they may follow), or the recent websites they visited [231, 233]. For instance, a user may be recommended 'promoted' tweets (tweets purchased by advertisers to be promoted), based on who the user follows, how they interact with a tweet and what they retweet, amongst other factors [232].

We design our experimental setup to minimize the effect of other factors mentioned earlier. Therefore, for each test account, we specify the same profile attributes (*i.e.,* language – English, location, and timezone – Germany), and we do *not* perform any activity (*e.g.,* posting tweets) from our test accounts. Furthermore, we collect the recommendations for each test account from within the same IP subnet in a German city.

Estimating random noise in the recommendations: There can be several sources of random noise in the recommendations. For instance, a part of Twitter's recommendations may involve randomly selecting from a pool of potential candidate items, or Twitter may perform A/B testing [176] where they may show slightly different results to different users to judge which is liked more by them.

To estimate the level of such noise in Twitter recommendations, we created two *identical* test accounts (for each real user mimicked) around the same time, which have the same profile features and follow the same set of users. We then collected the set of recommendations given to the identical accounts, at the same instants of time. We gathered multiple such snapshots of tweet recommendations over a week in December 2014, at 30 minute time intervals, and checked the overlap between the sets. Fig. 4.1 shows the percentage of tweets that are recommended to the two identical accounts at the same

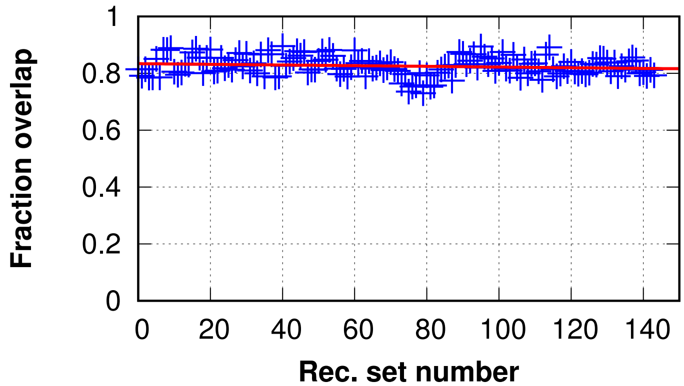


Figure 4.1: Percentage of common items in the set of tweets recommended to two exactly identical accounts at the same instant of time.

point in time. We observe an overlap of about 80% in the tweets recommended to the identical test accounts (across snapshots during the whole period), while the rest is random noise.

We computed the results reported in subsequent sections using several pairs of identical accounts and found them to be very similar for any two identical accounts. This high similarity indicates that even though there is some randomness/non-determinism in the set of tweets that get recommended to a particular account, the general characteristics of the recommended set of tweets remain the same for two identical accounts.

4.2.2 Data collected to construct recommended diets

Our methodology consists of creating a set of test accounts on Twitter and then investigating the recommendations given to these test accounts to study the impact of recommendations on the information diets of users. To ensure that the test accounts have realistic properties, we randomly selected a set of *real* Twitter users and created the test accounts to mimic these users.

Specifically, we randomly selected a set of 15 real Twitter users, such that their number of followings show a wide variety, ranging from few tens to close to a thousand. We next created 15 test accounts on Twitter, each of which cloned one of these real Twitter users by following the same set of people as the real user. We refer to these test accounts as $u1$, $u2$, ..., $u15$. We observed that topics of interests of these 15 selected accounts computed using the top topics in their consumed diets (based on the tweets they receive directly from the users they are following) display a large variety, including both generally popular topics like entertainment, sports, and politics, as well as more specialised topics

Test account	#users at 1 hop	#users within 2 hops	Top topics in consumed diet	KL div between consumed and recommended diets
u1	20	180,721	Sports, Politics, Entertainment	0.352
u2	34	12,260	Entertainment, Politics, Fashion	0.265
u3	71	136,441	Entertainment, Arts, Business	0.264
u4	84	58,834	Automotive, Environment, Sports	0.831
u5	90	34,128	Sports, Entertainment, Politics	0.236
u6	118	143,772	Entertainment, Education, Sports	0.165
u7	159	751,346	Politics, Business, Sports	0.111
u8	169	576,030	Sports, Entertainment, Business	0.331
u9	204	1,053,750	Business, Politics, Education	0.285
u10	395	638,305	Entertainment, Sports, Food	0.349
u11	520	479,300	Politics, Sports, Entertainment	0.121
u12	577	1,707,228	Sports, Entertainment, Politics	0.144
u13	604	3,063,196	Science, Hobbies, Education	0.423
u14	705	5,579,809	Politics, Food, Entertainment	0.466
u15	977	3,547,878	Politics, Business, Science	0.043

Table 4.1: Characteristics of the 15 test accounts created, which mimic the followings of 15 real Twitter users.

such as automotive, environment, and science. Table 4.1 gives the basic characteristics of these 15 test accounts.

The recommendations given in Twitter are dynamic and are updated in real-time [88]. Hence, for each of our test accounts, we gathered a snapshot of Twitter recommendations every 30 minutes, for a week in December 2014. On an average, each account received 708 recommended tweets in each gathered snapshot. Since these are too many for any user to view practically, we considered only the top 10 recommended tweets per snapshot to construct their recommended diets. [3] We also collected the tweets received by each test account from all their followings, during the same time in December 2014, to be able to construct their consumed diets too.

Constructing recommended diets To construct the recommended diet for a test account, we first collect all the tweets recommended to it. Then to construct the diet corresponding to this set of tweets using the topic inference methodology described in the previous chapter in Section 3.1.2, we first extract the keywords from each tweet and infer the topic for each keyword. Aggregating these keyword topics we construct the normalized topic vector, where the weight of a topic is the total normalized

[3]We verified that the insights presented later in the chapter hold even if we consider all recommended tweets (instead of the top 10).

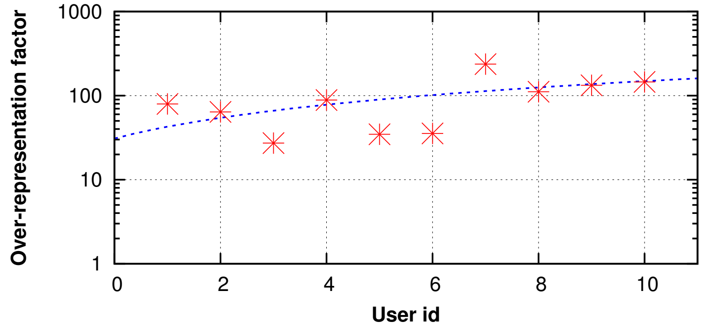


Figure 4.2: Over-representation of tweets that are tweeted/retweeted by at least ten users within u's 2-hop neighborhood, in the recommendations provided by Twitter to u.

contribution of all keywords inferred to be on that topic. The recommended diet of the test account is given by this topic vector.

Before comparing the consumed and recommended diets of our test accounts, we briefly evaluate our approach of collecting the recommendations for real users by creating test accounts that mimic their social neighborhoods, by investigating the role played by network neighborhood in determining the recommendations that an account gets. We consider the 2-hop network neighborhood of a given user u, as their social neighborhood, *i.e.,* the users u follows, and the users they follow in turn. Specifically, we investigate the following questions: (i) are most of the tweets recommended to u tweeted within u's social neighborhood?, and (ii) which particular tweets tweeted within the social neighborhood of u are more likely to be recommended to u?

Majority of recommendations from within 2-hop neighborhood: More than 98% of the tweets recommended to a target user u are drawn from the tweets that have been tweeted/retweeted by users within u's 2-hop neighborhood.

Selection of tweets to recommend from within the 2-hop neighborhood: Tweets that are retweeted by *multiple* users in u's 2-hop neighborhood are much more likely to be recommended to u. For instance, tweets that are retweeted by at least ten distinct users within u's 2-hop neighborhood are 20 - 200 times more likely to be recommended to u, as compared to the likelihood of selecting such a tweet at random (shown in Figure 4.2).

Thus we see that Twitter relies heavily on the social neighborhood of a user to provide personalized recommendations to them. In Section 4.2.1, we saw that Twitter's recommendation algorithm involves some amount of random noise. This randomness likely originates from the non-determinism in selecting the exact set of tweets to recommend out of the candidate set of tweets, *i.e.,*, the tweets posted or retweeted in the 2-hop neighborhood of the target user.

4.2.3 Limitations & ethical considerations

Our test accounts are passive accounts which do *not* perform any activity such as tweeting or favoriting *etc.* They only gather the recommendations given to them by Twitter. Even though the creation of such test accounts results in some users gaining an extra follower, we believe that this has a negligible effect on an extensive social network like Twitter.

In Section 4.2.2, we have shown that the recommendations on Twitter are social recommendations with a majority of recommendations being from within the 2-hop neighborhood of the target user. In such a scenario, our methodology of creating test accounts which mimic the social neighborhood of real users works well. However, if Twitter were to change its recommendation strategy to also include other factors like user's posting/retweeting behavior or if we wanted to study recommendations on another platform which takes into account more features beyond the social neighborhood, our methodology would not work in its current form. For such systems, we could have two options for studying the impact of recommendations on the information diets of users: (i) We could mimic not only the social neighborhood but also the activity of users. Though this method would perturb the system under observation a lot more and this may not be desirable. (ii) We could use donated data from real users to study the recommendations, for instance, by getting users to install a browser extension which would collect their recommendations on our behalf. Some recent studies have used a similar methodology for studying personalization in search results [199, 190, 191, 239]. This method would also overcome our data collection limitations and help scale up the size of study without perturbing the system significantly.

Finally, in the current study, we have focussed on studying the topical diversity of the recommended diets. In the future, it would be interesting to analyze diversity along multiple dimensions such as political leaning, tone, geographic sources *etc.*

4.3 Recommended diets vs. consumed diets

For each of our test accounts, we construct and compare three information diets:

Topics	Range of contribution to recommended diet (%)
Arts-Crafts	0.67 – 6.03
Automotive	0.59 – 10.83
Business-Finance	2.01 – 18.01
Career	0.47 – 3.46
Education-Books	0.71 – 5.47
Entertainment	5.14 – 40.36
Environment	1.27 – 6.11
Fashion-Style	0.19 – 2.64
Food-Drink	0.49 – 4.32
Health-Fitness	0.79 – 5.45
Hobbies-Tourism	0.84 – 4.31
Paranormal	0.60 – 4.39
Politics-Law	9.03 – 33.34
Religion	1.76 – 6.81
Science	3.57 – 13.05
Society	0.91 – 2.46
Sports	6.14 – 46.97
Technology	1.13 – 9.79

Table 4.2: Range of contributions of different topics in the recommended diets given to the test accounts.

i. *Consumed diet*: constructed from the tweets the test account receives directly from the users it is following,

ii. *Recommended diet*: constructed from the top tweets recommended to the test account, and

iii. *Combined diet*: assuming that a user pays equal attention to the consumed and the recommended diets, this is constructed by considering the average contribution of each topic from the test account's consumed and recommended diets.

4.3.1 Are the recommendations personalized for each user?

We begin by investigating whether different users get different recommended diets. To answer this question, we have generated Table 4.2, where we show the contribution of the various topics in the recommended diets given to the 15 test accounts. Examining the range of contributions of the topics in the recommended diets given to different test accounts, we find that there is a wide range of the contributions of different topics – politics varies between 9.03% and 33.34%, science between 3.57%

and 13.05%, and entertainment between 5.14% and 40.36%. This variation makes it evident that different accounts are being recommended different diets, with differing contributions of topics. [4]

4.3.2 To what extent do the recommendations given to a user match their consumed diet?

Table 4.1 depicts different characteristics of our 15 test accounts, including the number of users at 1-hop (*i.e.,* the number of followings), the number of users within 2- hops, the top 3 topics from the consumed diet (to capture the topics of interest of the users) and the KL-divergence between the consumed and recommended diets. We use the standard measure of KL-divergence between the consumed and recommended diets to quantify how well the recommended diet matches the consumed diet of a user. The smaller the value of KL-divergence, the closer the two diets are. We observe that the KL-divergence values for the 15 test accounts vary in the range of 0.043 to 0.831, with five accounts having KL-divergence values below 0.2, and three having values above 0.4. This variation in the KL-divergence values suggests that the recommendations match the consumed diets to different extents for different users.

Figure 4.3 shows the topical compositions of the consumed and recommended diets for two test accounts – (i) $u15$ which has the minimum KL-divergence, and (ii) $u4$ which has the maximum KL-divergence of the recommended diets from their consumed diets. It can be seen that the recommended diet of $u15$ mostly matches the consumed diet, while for $u4$ there is a greater mismatch between the two diets. For instance, though $u4$ consumes a lot of information on the topics automobile and environment, its recommended diet has a much lower fraction of these topics. On the other hand, the recommended diet for $u4$ has higher fractions of politics, religion, and science, topics which are not that significant in its consumed diet.

These observations suggest that the recommended diet that a user will get does not always match their consumed diet. We also observe cases where two accounts are consuming approximately the same amount of information on a particular topic, but they receive very different amounts of information on this topic in their recommended diets. These differences may be driven by the fact that different users

[4] Note that the ranges shown in Table 4.2 are for our 15 test accounts, and actual ranges may be even larger for other Twitter user-accounts.

have different social neighborhoods, and the social recommendations given by Twitter are derived from what information is popular in the social neighborhood of the user [88].

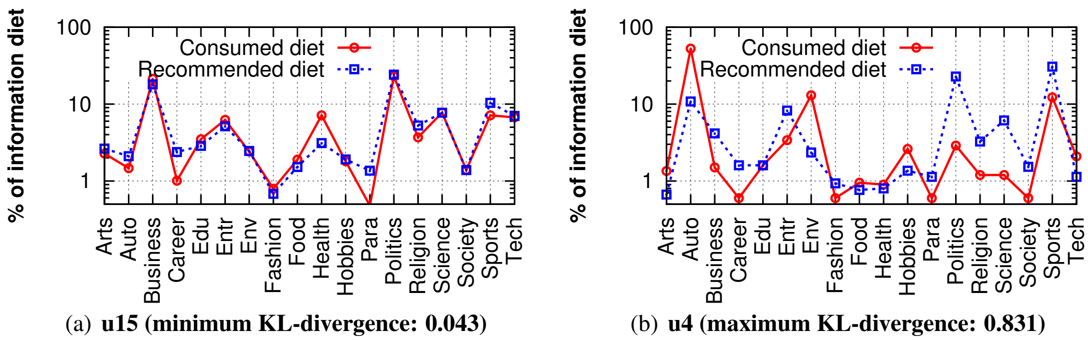
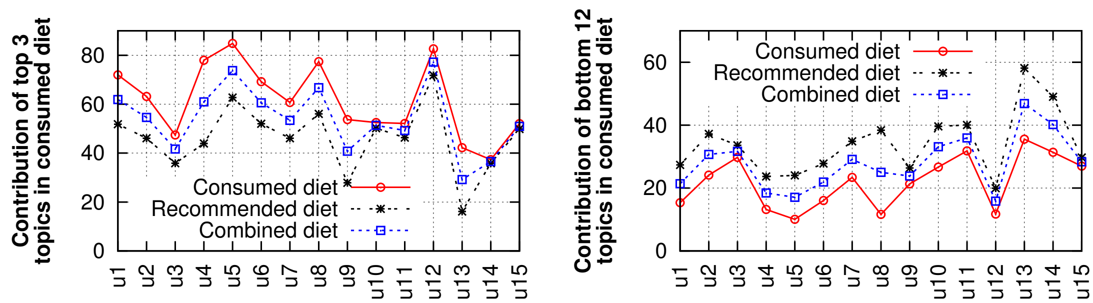


Figure 4.3: Comparing the consumed diet and recommended diet of two test accounts – (i) the one with the minimum KL divergence, and (ii) the one with the maximum KL divergence of the recommended diet from the consumed diet.

The effect of the social neighborhood can also be observed from Table 4.2 where it is seen that popular topics like entertainment, politics, and sports are being recommended to everyone irrespective of whether they are interested in these topics. Every account is getting recommended at least 5%, 9% and 6% in entertainment, politics, and sports respectively, which is significantly higher than for other topics. On an average, every test account receives up to 17%, 19% and 17% on entertainment, politics, and sports respectively. As observed in the previous chapter (Section 3.1.3.2), there are a large number of users tweeting about these topics of general interest (see Figure 3.7), and hence everyone's neighborhood is likely to contain significant discussions on these topics, which get included in the social recommendations.

4.3.3 Do the recommendations mitigate or exacerbate the imbalances in the users' consumed diets?

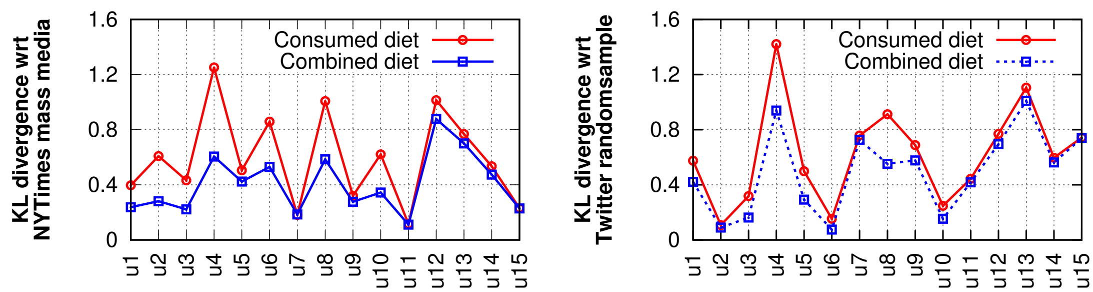
Finally, to address the question whether the recommendations mitigate or exacerbate the biases in the users' consumed diets, we consider the top 3 topics in the consumed diet of an account (*i.e.,* the 3 topics on which the account consumes most information from its followings), and measure the contribution of these 3 topics in the consumed and recommended diets of the user. These are plotted for the 15 test accounts in Fig. 4.4(a). Similarly, the Fig. 4.4(b) shows the contribution of the bottom 12 topics in the consumed diet of an account, in the consumed and recommended diets.



(a) **Top 3 topics** (b) **Bottom 12 topics**

Figure 4.4: The contribution of the (i) top 3 consumed topics and (ii) bottom 12 consumed topics in the consumed, recommended and combined diets for the test accounts.

Interestingly, we observe that the top 3 consumed topics account for a significantly smaller share in the recommended diets of the users, as compared to the consumed diets. Again, the contribution of the bottom 12 topics is higher for the recommended diets, as compared to the consumed diets of the users. Thus, the recommendations tend to even out the imbalances in the consumed diets of the users, by including information from the lower ranking topics in users' consumed diets. Hence, social recommendations are reducing the gap between the information that different users are exposed to by mitigating the biases in the user's diets.



(a) **NYTimes mass media diet** (b) **Twitter's 1% random sample diet**

Figure 4.5: KL divergence of the consumed and combined diets of test accounts from the baseline diets of (i) NYTimes mass media diet, and (ii) Twitter's 1% random sample diet.

To quantify this mitigation, we computed the KL-divergence between a user's consumed and combined diets, from the baselines of: (i) NYTimes mass media diet (introduced in Section 3.1.3.1), and (ii)Twitter's 1% random sample diet (introduced in Section 3.1.3.3), depicted in Figures 4.5(a) and 4.5(b) respectively. We can note that for each test account, the divergence from either of the baselines is lesser for the combined diet than for the consumed diet, showing that the social recommendations are having an equalizing effect across the users (and driving the combined diets towards the baselines).

Thus, we find that social recommendations mitigate the imbalances in the users' consumed diets, bringing in more heterogeneity into what the users are being exposed to.

4.4 Conclusion

In this chapter, we presented a methodology for investigating the topical diets that social media users are consuming via the algorithmic channel of recommendations. We used test accounts which mimic real users to collect the social recommendations on Twitter, which typically rely on a user's neighborhood to recommend content to them. We observed that the social recommendations somewhat mitigate the topical imbalances in the users' consumed diets by adding some topical diversity.

We envisage that this work will not only create awareness among social media users about potential imbalances in their information diets, but will also have implications for the designers of future information discovery, curation, and recommendation systems for social media. For instance, we found that social recommender systems are bringing in more heterogeneity into what the users are being exposed to. While this is good for broadening the horizons for the users, topic-specific recommendations might be necessary to provide information focused on the users' interests. Studying the information diets provided by different types of recommender systems, and their impact on the information that a user is exposed to is a promising direction to pursue in the future.

CHAPTER 5

Search: Quantifying political bias in social media and Web search

With the increasing amounts of user-generated information available online, another set of automated retrieval algorithms which form a vital input channel to the users' information diets are the search systems. We all rely on search for a wide variety of goals in our day-to-day lives – ranging from finding specific website or content (navigational queries) to learning more broadly about entities, people, topics or events (informational queries) [245]. For instance, during election season, people are known to make repeated queries about political candidates and events (*e.g.,* "democratic debate", "Donald Trump", "climate change") on the Web, as well as on social media sites like Facebook and Twitter [83, 226] to learn more about their queried terms.

While the goal of informational search queries is to provide users with greater knowledge about a topic, this knowledge is not necessarily always impartial. When a query is issued to a search system, a set of relevant items for the query are first extracted from the whole corpus of data items (*e.g.,* web links or social media posts). This set of relevant items for the query are in turn fed to the ranking system which returns a ranked list of search results to the user who made the query. For polarizing topics like politics, many of these returned results can be biased towards one political perspective or the other, therefore by ranking items from one perspective higher than the other the ranking system could (possibly inadvertently) return a list of politically biased search results to the user. This bias in the search results could be introduced because of biased data that forms the input to the ranking system, or because of the ranking system itself.

The potential biases that search systems can introduce and users' unquestionable trust in search results have lead to growing concerns about search systems' impact on the behavior of users, especially in scenarios where they may potentially misinform or mislead the users. Prior field studies have shown that not only do the users place greater trust in highly ranked search results [175], but the opinions of undecided voters can be manipulated by biasing the search results about political candidates [63]. In such polarizing scenarios, where multiple different perspectives about the searched topic exist (*e.g.,* political candidates or events), the bias in the top search results can influence the user's opinion and shape public opinion towards (or away from) certain competing perspectives. However, such biases of search systems are challenging to detect and quantify, since multiple sources of bias exist (*e.g.,* input data and ranking system) whose effects are hard to disentangle.

As the last part of the thesis, we explore the impact of search systems – the third channel of consumption on social media – on the information diets of the users. Since the topic of information being consumed is implicitly provided by the user when they make a search query, in this part of the thesis we focus on the information diets within a particular topic – US politics.

We investigate the biases introduced by the search systems into the information diets of users by proposing a novel generalizable search bias quantification framework. This framework not only captures the bias in the search results output by a search system but is also capable of decoupling this output bias into different components to identify the sources of bias – the input data or the ranking system. For our chosen context of 2016 US Presidential primaries, we first apply our search bias quantification framework to political searches on social media (Twitter) to quantify and investigate its sources of political bias, and then we use our framework to quantify and compare relative bias for political searches on social media search (Twitter) and Web search (Google).

To apply our framework to study the sources of bias in political searches on Twitter social media, we first needed a methodology to measure the political bias of an individual search result, *i.e.,* a tweet. We operationalized the political bias of a tweet as its source bias, *i.e.,* the political bias of the author of the tweet, and developed a highly scalable and accurate author based crowdsourced methodology for inferring the political bias of a Twitter user. We then utilized these inferred biases of tweets to quantify the sources of bias for Twitter search. Not only could we observe the search results output by the ranking system of Twitter search, but we were also able to gather the tweets containing a query which form the input to the ranking system. Armed with this data, we were able to disentangle the bias of different sources of bias for Twitter search, using our bias quantification framework. In our analyses of Twitter search results, we show that the bias in the search results does not only originate from the ranking system, but the bias of the input data (that is input to the ranking system) is also a significant contributor to the overall search bias. Moreover, we observe that the top Twitter search results display varying degrees of political bias that depends on several aspects, such as the topic (event/person) being searched for, the exact phrasing of the query (even for semantically similar queries), and also the time at which the query is issued.

After quantifying the bias in social media search, we proceed to use our quantification framework to compare the *relative bias* for political searches on two popular search systems - Twitter social media search and Google Web search. Our motivation for performing this comparison is to make the biases of different channels more visible and accessible to the users. Traditional media channels like Fox News or CNN have often been scrutinized by academics [197, 20, 41, 79, 86, 27, 161] as well as media watchdog groups (like FAIR (`fair.org`) and AIM (`aim.org`)) for fairness, accuracy and balance in the news they report. Additionally, tools have also been developed to mitigate or expose the media bias [188, 146, 178, 161, 3] to users. However, the relative biases of newer digital algorithmic channels like search systems are not as well studied and documented as yet, and thus users may not be taking their relative biases into account while selecting the channel to get their information from. In fact, many users believe that these algorithmically curated channels (as opposed to human editorial curation) are powerful, infallible and thus unbiased [67, 217], which is far from being true. This lack of awareness can results in "blind faith" in search systems [175], and impairs the users from making an informed choice of which search channel to use. With this study, we aim to highlight the differences in the political bias of these two popular search systems – Twitter social media search and Google Web search – and make their relative bias more visible.

Our comparison of relative bias of the two search systems reveals that the bias for political candidates is much more favorable to the candidates on Web search than on social media search. This difference is mainly due to multiple neutral or supportive (candidate-controlled) web-links (for instance candidate's homepage or their social media profile links) that get included in the top results on Web search. We also observed that the bias in Web search results is less dynamic over time as compared to bias in social media search. Our findings show that search systems exhibit not only political bias in their search results but also different search systems exhibit different biases. It is important to highlight these differences in political bias of varying search systems, since the users currently may not be taking these biases into account when choosing one search system over the other to get information from.

Our research contributions in this work can be summarized as follows:

1. We propose a *novel generalizable search bias quantification framework* to measure not only the bias in the search output but also to discern the contribution of different sources – input data or ranking system (Section 5.2).

2. We apply the framework to *investigate the sources of bias* in political searches on Twitter social media search where we show that both input data and ranking system contribute to the final output bias seen by the users. We also observe that the bias varies with the topic being searched for, the exact phrasing of the query and the time at which the query is made. (Section 5.3).

3. We also utilize our framework to *compare the relative bias* for political queries on two popular search systems: Twitter social media search and Google Web search. As compared to social media search, we find that the political bias on Web search is a lot less dynamic, more favorable for the candidate queries, and has a higher fraction of top search results containing links to candidate-controlled sources, such as links to their website or social media profiles. (Section 5.4)

Our work is aimed towards making social media users aware of the potential political biases of social media search and how it compare with the bias in web search and encouraging the development of novel information retrieval systems and mechanisms for presenting search results which could represent multiple competing perspectives on the same event or person.

5.1 Related work

In this section, we discuss the prior research done in three related areas: (i) measuring political bias in media, (ii) characterizing bias on social media, and (iii) characterizing bias in Web search systems.

5.1.1 Measuring political bias

The first step in quantifying the bias for political searches on social media and web search is measuring the political bias of an individual result (*i.e.,* a tweet or a web-link). There have been several attempts to infer the political bias of news media, news articles, and blogs. However, there has been limited work on inferring the bias of content of short social media posts like tweets. Next, we briefly outline the research done in measuring the political bias in news media as well as social media.

5.1.1.1 Measuring political bias in news media

News media organizations have traditionally acted as the *gatekeepers of information* in the society, by regulating the news that is consumed by the people in the society [213]. Given the huge impact that the media organizations can have on the societal evolution, media studies scholars have long studied and monitored their activities. Their interest is also partially driven by their concern that ideologically partisan and unregulated media can adversely affect many societal structures including political outcomes [86, 49]. As a result, significant research effort in the past has focussed on investigating the political bias in news media [86, 79, 55, 41, 255, 253, 49, 23, 134, 264, 197, 153].

Content-based approach: The first set of studies [86, 79, 55, 41, 255, 253, 49] directly scrutinized the content produced by the news organizations to identify their political biases. Groseclose *et al.* [86] matched the media outlets and the US Congress members who co-cited political think tanks and then used the ADA scores of the Congress members to assign a political bias score to the media outlets.[1] While Gentzkow and Shapiro [79] examined whether the language used by a media outlet was more similar to Republican or Democrat members of the Congress to infer the 'media slant' for the outlets. Finally, some studies [55, 41] focussed on the coverage of important events and societal issues by the media outlets to infer their political biases. Covert and Wasburn [55] measured the media bias by manually analyzing the content published by media outlets to cover various social issues. Whereas, Budak *et al.* [41] used a combination of crowdsourcing and machine-learning methods to study the selection and framing of political issues by different news organizations. Yigit-Sert *et al.* [255] leverage user comments and the content of the online news articles to automatically identify sources which are biased towards the same direction for a topic.

Audience-based approach: On the other hand, the second approach for inferring the political bias of media outlets examines the biases of audiences of news organizations to infer the political biases of the

[1] ADA scores are assigned to the US Congress members by the political watchdog group "Americans for Democratic Action" (**www.adaction.org**).

news outlets. This approach is based on the idea that the ideological biases of the users get reflected in their news consuming and sharing practices [208]. Bakshy *et al.* [23] estimated the content alignment scores for news articles being shared on Facebook by averaging the leanings of Facebook users sharing it (where users' leanings have been self-reported), and then averaged these content alignment scores over the articles published by a media outlet to determine the alignment scores for the news outlet. Le *et al.* [134] presented a method to measure ideological slant of individual news articles by monitoring their propagation on Twitter, where they analyzed the connectivity of the users tweeting an article to label them as Republican or Democrat leaning. Weber *et al.* [242] developed a tool for examining the political polarization of hashtags on Twitter, where they inferred the leaning of a hashtag using the users' leanings who are posting it. Zhou *et al.* [264] inferred the leanings of news articles and users on Digg using a semi-supervised learning scheme that propagates liberal and conservative labels, with the assumption that liberal users are more likely to vote for liberal articles and users and conservative users for conservative ones. While the Balance study [161] assigned political bias scores to many of the popular news websites based on the political leanings of the websites, blogs and Digg users that link to or vote for the news website. More recently, Ribeiro *et al.* [197] leveraged the advertiser interfaces of social media sites (which offer detailed insights into the audiences of the news outlets), to infer the ideological leaning of the outlet by examining to what extent the liberals and conservatives are over- or under-represented in the outlet's audience. Not just researchers, but think tanks like Pew Research [153] have also examined the media bias and political polarization by conducting extensive readership surveys of the audiences.

5.1.1.2 Measuring political bias of social media users

There has been limited work on inferring the bias of content of short social media posts like tweets. Instead, researchers have heavily focussed on inferring the bias of the users posting on these social media platforms.

Content based methods: The first set of studies infer the political leaning of social media users by modeling how users with different leanings use language [189, 144, 70, 248, 195, 37].

Makazhanov and Rafiei [144] developed a language model for each party using the tweets of the party's candidates and then inferred a user's political preference by matching the user's tweets with the language models of the parties. Fang *et al.* [70] inferred the voting intentions of users by the content of their tweets, by employing a topics-based Naive Bayesian classifier. Wong *et al.* [248] utilized tweeting and retweeting behavior of users to infer the political leanings of Twitter accounts of prominent users and media sources. Rao and Yarowsky [195] used user language on social media (in the form of sociolinguistic features) to infer multiple user-attributes including political orientation. Boutet *et al.* [37] used the tweets referring to a political party or retweets of users with a known political leaning to infer the political leaning of users on Twitter.

Network based methods: The second category of studies [81, 54, 52, 33, 249, 81, 24, 25] infer the political leanings of social media users by leveraging network information, with the networks being considered ranging from following social networks to retweeting and endorsing networks. Golbeck and Hansen [81] computed the political preferences of Twitter users by starting with the seed set of Congressional liberal/conservative ADA scores and propagating them on the follower network. Barberá [24, 25] also developed a Bayesian Spatial following model which places Twitter users along a universal ideological scale by leveraging who these users follow. Conover *et al.* [52] applied label propagation to a retweet graph for user classification and found the approach to outperform tweet content based machine learning methods. They [54] showed that the network of political retweets has a highly partisan structure with limited connectivity between users with different leanings. Bond and Messing [33] inferred the political leanings of Facebook users by observing the endorsements of Facebook Pages of known politicians, while Wong et al. [249] measure the endorsements in terms of retweeting behavior of users to infer their political leanings.

Network + content based methods: The last set of studies [180, 181, 261, 132, 142] combine network and content features for identifying the political leaning of social media users. Pennacchiotti and Popescu [180, 181] leveraged user behavior, network features and linguistic features based on users' Twitter feed to classify users along political leaning, ethnicity, and affinity to particular businesses. Al Zamal *et al.* [261] leveraged network homophily by utilizing both user features as well as features extracted from the user's friends' (network neighborhood) profiles and posts, for the task of inferring latent attributes of users including political leaning. Lahoti *et al.* [132] utilized the network structure and the content consumed by a user within a non-negative matrix factorization problem with shared latent factors to map the users on a liberal-conservative ideology space. While BiasWatch [142] is a system to discover and track bias themes and users from opposing sides of a topic in a semi-supervised manner. They first identify a seed set of biased partisans, and then build user similarity networks around these users (where the edges capture the similarity between users via content- and retweeting link-based features) and propagating the bias along this similarity network.

Most of these prior studies assume that the leaning of the user is explicit in their language, social connections or endorsements. However, this may not always be true. We propose a methodology for inferring the bias of a Twitter user by leveraging their interests, which are correlated with their political affiliation. We build upon prior studies that have shown that people's political affiliations are correlated with their personality attributes and responses to different stimuli [42, 187, 212]. Shi *et al.* [212] have used co-following graphs to measure the alignment between political and social issues (using common set of followers) in the US society and found that the partisan divide extends to cultural and lifestyle preferences (as also observed by prior studies: [74, 13, 58, 31, 59, 60, 130]). Another set of studies [77, 240] examined how out of context accounts (*e.g.,* non-political accounts) that a user follows can be used to map the user to a contextual (*e.g.,* political) setting. We show that our proposed interest-based method can be used to infer the political leaning of users with varying levels of political activities. In the past, Cohen and Ruths [51] used supervised methods to classify users into different groups of political activities and showed that it is hard to infer the political leaning of "normal" users. Following their advice, later in this chapter, we present an evaluation of our methodology over three different datasets of test users on Twitter. Finally, we utilize our proposed methodology to quantify the bias for political searches on Twitter social media.

5.1.2 Characterizing bias on social media

With more and more users relying on social media platforms like Twitter and Facebook to receive news [136] and information about on-going events and public figures [226], there has been a debate about the impact these platforms are having on the news that users are consuming. While some have envisioned increased democratization with users from different political ideologies engaging with each other [209], others warned that use of social media platforms could encourage selective exposure by reinforcing users' existing biases [140]. Further inspection of cross-ideological exposure [103] revealed that political discourse on social media is highly partisan [54, 15] and users are unlikely to get exposed to cross-cutting content via their social neighborhood. These results have been reinforced by studies showing that not only are social media users more willing to communicate with other like-minded users [140, 215]; they are also unable to engage in meaningful discussions with users with different beliefs than their own [254].

Further, this selective exposure effect is often amplified by personalized recommendation systems provided by social media platforms that recommend to users more items matching their own preferences. While this approach may work well for recommending items such as movies or music, there are concerns that such systems might limit exposure to ideologically cross-cutting content leading to the formation of 'filter bubbles' or 'echo chambers' [23, 38, 75, 177]. Researchers have long worried that people are predominantly interacting with other like-minded individuals, and are not being exposed to alternate views, which can lead to a worrying increase in *societal polarization* [220, 205].

To combat this polarization in news consumption some systems intended to promote diversity have been proposed. These systems deliberately expose users to different points of view with the hope of nudging them to read opposing points of views also [163, 178, 16, 168, 167, 116]. Unfortunately, such systems have had limited success in practice. While some diversity-seeking users may be satisfied, many users either ignore or reject disagreeable points of view [165]. In fact, in some instances, exposure to opposing views can increase the political polarization by causing the readers to retreat to an even more entrenched view of their prior beliefs [141, 151, 160, 22]. As a solution, Babaei *et al.* [20] recently proposed a system called "Purple feed" which shows the users news posts which are likely to have a high consensus in the reactions of both Republican and Democrat leaning readers.

While these studies give evidence of polarized content generation and sharing on social media platforms, it is unclear how this data impacts automated retrieval systems like search systems. In this work, we propose a search bias quantification framework which not only quantifies the bias in the output ranked list shown to the users, but it also discerns to what extent is this bias due to the ranking system of the search system, or the biased data generated on the platform that is input data to the ranking system.

5.1.3 Characterizing bias in Web search systems

In recent years, Web search engines and their potential biases have received a lot of scrutiny [225, 234, 76, 237, 159, 158, 19, 199, 149]. This scrutiny has typically stemmed from the concern that dominant search engines like Google might favor certain websites over others when ranking relevant search results. For example, Edelman argued that Google manipulates its search results to rank its services (such as Google Health links) higher than other competing services [62]. In another example, Vaughan *et al.* examined the geographical bias in Web search and observed that sites from some countries like the US are covered more than sites from other countries [237].

Several studies have focused on the *political bias* of Web search queries and search results during recent years. Researchers [36, 241] developed an online searchable database of politically charged queries, where the political leaning of the queries was generated using anonymized search engine queries which resulted in a click on US political blogs. In another line of research, researchers conducted field studies to examine the influence of political bias seen in search results on users' voting decisions. For instance, Epstein and Robertson found that by manipulating the political bias in top search results they could impact the voting preferences of undecided voters by 20% or more [63] and they termed this phenomenon as search engine manipulation effect. As a continuation of this line of research, Epstein *et al.* have shown that modifying the design of search engines to include alerts about the bias in the search results shown to the users can mitigate the aforementioned search engine manipulation effect significantly [64]. Also, Metaxas and Pruksachatkun conducted another recent audit of several search engines during the 2016 US Congressional Elections [149].

Motivated by these prior studies and their findings, in this work, we propose a generalizable search bias quantification framework and apply it to investigate the sources of bias in social media search, as well as, apply it to compare relative biases of social media and Web search.

Personalization in Web search: A complementary line of work has focussed on the personalization effects and studied the differences in the results seen by different users for the same query due to personalization. Various factors including geolocation of users have been found to lead to personalization of search results [90, 119]. In another study [122] it was shown that during disruptive events such as shootings, the users tend to changes their information-seeking behavior and use the search engines to seek information that they agree with. Robertson *et al.* [199] designed a Chrome extension for surveying participants and collecting search results that they would have seen for a set of political queries to audit Google search rankings and suggestions. They observed significant differences in the composition and degree of personalization of search results based on query type and the time of querying. Another set of researchers [190, 191, 239] also followed a similar methodology of designing a browser plugin using which users can donate their search results for a predetermined set of queries to study personalization of search results for political queries.

In contrast, we study the bias in consistent, non-personalized search results for political queries shown to all users on social media search and Web search by adopting measures to mitigate the personalization effects as described later in the chapter. We find that biases exist even for such non-personalized results. We do acknowledge that in reality most searches made by users are personalized. Therefore our results may not be representative of most searches done in the wild. However, we believe that the personalization is most likely to exacerbate the biases we observe and report in this work and this forms a potential direction of future research.

Black-box approach to auditing algorithms: Recently, the rise of algorithmic platforms' influence on users' online experience has motivated many studies [57, 221, 91, 203, 69] to audit these platforms and understand their biases. While some of these algorithmic systems' functionalities are open to the public, making the auditing process easier, most of them are not. The walls of intellectual proprietary, high complexity of these algorithms and the perils of gaming a system via malicious users put these algorithms in a black box, making it almost infeasible to have access to an algorithm's specifications from outside, like in our study. While we know about a few general factors that a search engine takes into account in curating the search results (such as relevancy, popularity, and recency), there are hundreds of other features that are hidden in a black-box, preventing us as researchers from being able to pinpoint the exact feature(s) of the algorithm which might be leading to the bias being introduced in the search results.

Therefore, building on previous studies that have adopted the "black-box" view for an algorithmic system while auditing it [68, 135, 91, 90, 47], we characterized the bias of the ranking algorithm in Twitter's search platform and Google Web search platform, without knowing their internal functioning. Our proposed auditing framework can help users, system designers and researchers to become aware of possible biases of a search process, while they might not be aware of the details of the process itself. For users, this awareness can result in more intelligent use of a system, knowing that their search results can be far from neutral in some cases. For system designers, such auditing platforms can be used to investigate the algorithm's specifications, particularly when the bias has been introduced by the algorithm and not the system input. And finally, researchers and watchdog organizations can actively utilize such auditing platforms to measure bias and compare it among different search platforms, making the research community and the system designers aware of potentially misleading biases.

5.2 Search bias quantification framework

In this thesis, we quantify the bias for political searches on Twitter social media search and Google web search in the context of the US political scenario, which has two primary political parties: the Democratic party and the Republican party. In this section, we propose a bias quantification framework which captures the bias introduced at different stages of a search process.

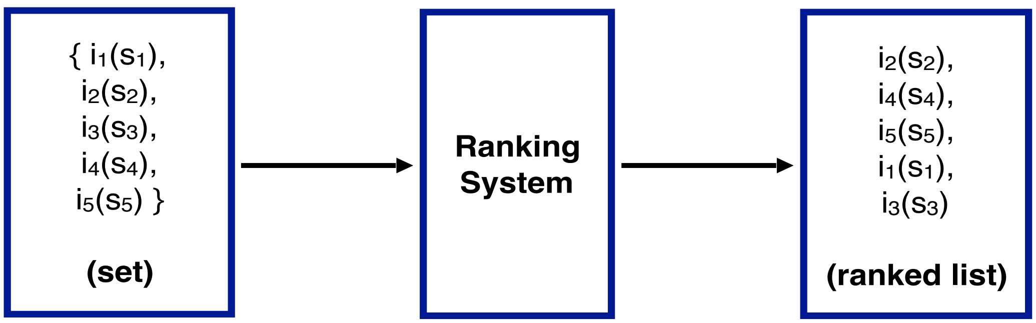


Figure 5.1: Overview of our search bias quantification framework. For a given query q, a set of data items relevant to the query is first selected. Each individual data item (e.g., i_1, i_2) has an associated bias score (e.g., s_1, s_2). These set of relevant items is input to the ranking system which produces a ranked list of the items. Our framework includes metrics for measuring the bias in the set of relevant items input to the ranking system (input bias), and the bias in the ranked list output by the ranking system (output bias).

Figure 5.1 gives a high-level overview of the different stages of information retrieval via an algorithmic search system. The search system retrieves information from a corpus of data, where each individual data item (e.g., i_1, i_2) has an associated bias score (e.g., s_1, s_2). In the later sections (Section 5.3 and 5.4), we describe methodologies for computing the bias score for political searches on Twitter social media and Google web search platforms. When a user makes a query q, a *set of data items* relevant to the query is first selected out of the whole corpus. Then, this set of retrieved relevant items forms the input data to the ranking system which produces a *ranked list of the relevant items*, which is shown as the search output to the users. The framework can also be generalized to modern-day

Rank r	Bias till rank r	Value
1	$B(q,1)$	s_2
2	$B(q,2)$	$\frac{1}{2}(s_2+s_4)$
3	$B(q,3)$	$\frac{1}{3}(s_2+s_4+s_5)$
4	$B(q,4)$	$\frac{1}{4}(s_2+s_4+s_5+s_1)$
5	$B(q,5)$	$\frac{1}{5}(s_2+s_4+s_5+s_1+s_3)$
Output bias at rank 5		$\frac{1}{5}[s_2(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5})+$ $s_4(\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5})+$ $s_5(\frac{1}{3}+\frac{1}{4}+\frac{1}{5})+$ $s_1(\frac{1}{4}+\frac{1}{5})+$ $s_3(\frac{1}{5})]$

Table 5.1: Explaining the bias metrics with reference to Figure 5.1.

IR systems which perform retrieval and ranking together, such as systems using topic modeling. We comment on this issue in the Section 5.6.

Within our framework, we define three different components of the bias of a search system, each of which is quantified in terms of the biases of the individual data items: (i) *input bias*: the bias in the set of retrieved items relevant to the query that are filtered out of the whole corpus. This set of retrieved items serve as the input data to the ranking system, (ii) *ranking bias*: the bias introduced by the ranking system, and (iii) *output bias*: the cumulative bias in the ranked list output by the search system and shown to the users. In the rest of the section, we discuss the metrics we proposed to quantify these different components of bias of a search system.

5.2.1 Bias score of an individual data item

We are interested in quantifying the search bias for political queries in the context of US politics. Since there are two primary political parties in the US, each data item (*e.g.,* a tweet or a web-link) can be positively biased (*i.e.,* supportive), negatively biased (*i.e.,* opposing), or neutral towards each of the parties. Therefore the bias score of each item must capture the extent to which it is biased with respect to the two parties.

To apply our bias quantification framework in the context of political searches on a search platform, we need a methodology for inferring the bias scores for each data item (indicated by s_i in Figure 5.1). Later in the paper, we present methodologies for measuring the bias scores of individual items for our chosen scenario of political searches on Twitter social media (Section 5.3) and Google Web search platforms (Section 5.4).

Next, we use these bias scores of individual data items to define the metrics for the input bias, output bias, and ranking bias.

5.2.2 Input Bias

Once a user issues a query, the search system retrieves a set of items from the whole corpus that are relevant to the query and provides them as an input to the ranking system. Since this input data captures the bias introduced due to the filtering of the relevant items from the data corpus according to the issued query, we measure the input bias for a query as the aggregate bias of all the items relevant to the query in this input data set. In other words, input bias gives a measure of the bias a user would observe if they were shown *random* items relevant to the query, instead of the output list ranked by the ranking system.

Specifically, the Input Bias $IB(q)$ for query q is the average bias of the n data items that are relevant to q

$$IB(q) = \frac{\sum_{i=1}^{n} s_i}{n} \tag{5.1}$$

where the summation is over all the bias scores (s_i) of the n data items found relevant to q. For instance, for the query q shown in Figure 5.1, the input bias is $IB(q) = \frac{1}{5}(s_1 + s_2 + s_3 + s_4 + s_5)$.

5.2.3 Output Bias

The output bias of a search system is the cumulative bias in the final ranked list of search results presented to the user who issued the search query. Prior studies have shown that not only are the users more likely to browse the top search results [145], but they also tend to put more trust in them [175]. Therefore, we propose an output bias metric inspired from the well-known metric – mean average precision [145] – which gives more importance to higher ranked search results.

For a given search query q, we first define the bias till a particular rank r in the ranked results (*i.e.,* the aggregate bias of the top r results). The bias $B(q, r)$ till rank r of the output ranked list is defined as

$$B(q, r) = \frac{\sum_{i=1}^{r} s_i}{r} \tag{5.2}$$

where the summation is over the top r items in the ranked list.

As an example, the first five rows in Table 5.1 depict the bias till ranks 1, 2,..., 5, for the sample search scenario shown in Figure 5.1.

The Output Bias $OB(q, r)$ for the query q at rank r is then defined by extending the above definition as follows,

$$OB(q, r) = \frac{\sum_{i=1}^{r} B(q, i)}{r} \tag{5.3}$$

The last row of Table 5.1 depicts $OB(q, r)$ at rank $r = 5$ with respect to Figure 5.1. In this formulation, the bias score s_2 of the top-ranked item i_2 is given the highest weight, followed by the bias score s_4 of the second-ranked item i_4, and so on, following the intuition that the bias in the higher ranked items is likely to influence the user more than bias in lower ranked items.[2]

5.2.4 Ranking Bias

If the internal details of the deployed ranking system were known, then the ranking bias could be measured by auditing the exact features being used for ranking. However, for most of the real-world commercially deployed search engines, the internal details of the ranking system are not known publicly. Therefore, building on previous studies that have adopted a "black-box" view for an algorithmic system while auditing it [68, 135, 91, 90, 47], we treat the ranking system as a black-box, such that we only observe its inputs and outputs. In such a scenario, the ranking bias captures the *additional* bias introduced by the ranking system, over the bias that was already present in the input set of relevant items.

Therefore, we define the Ranking Bias $RB(q, r)$ for the query q as simply the difference between the output bias and the input bias for q (as given by Equations 5.1 and 5.3).

$$RB(q, r) = OB(q, r) - IB(q) \tag{5.4}$$

[2]Similar to how missing relevance judgements are handled in the Information Retrieval literature [256], in case there exists an item for which the bias score cannot be computed, we just ignore the item and compute the rankings.

5.2.5 Time-averaged Bias

To capture the overall trend in the bias, we collect multiple snapshots of search results, compute the different bias metrics for each snapshot, and then compute the time-averaged values of the aforementioned metrics. For instance we compute the time-averaged output search bias $TOB(q, r)$ as the average of the $OB(q, r)$ (given by Equation 5.3) values measured at various instants of time. Similarly, we define $TIB(q)$ and $TRB(q, r)$ as the time-averaged input bias and time-averaged ranking bias for query q respectively.

5.3 Investigating sources of bias for political searches on social media

Having described our search bias quantification framework, we next apply it to political searches on Twitter social media for queries related to 2016 US presidential primaries. With this study, we highlight an important application scenario of our framework, where not only can we observe the search system's output results, but we also can observe the set of relevant items that form the input to the ranking system.

We begin by describing our selected queries and data set for Twitter search (Section 5.3.1), followed by the methodology we used for measuring the political bias of an individual Twitter search result (Section 5.3.2), and then we finally present our findings about how biased are the search results for political topics on Twitter and where does this bias in the search results comes from (Section 5.3.3).

5.3.1 Collecting Twitter search data

Here, we describe the queries we considered and the data gathered from Twitter for conducting the analyses.

5.3.1.1 Selecting search queries

In an ideal scenario, for studying the bias in political searches on Twitter, we would use the actual search queries that people are making on the platform for following news and information related to 2016 US presidential primaries. However, we did not have access to this proprietary data about the queries issued on Twitter. In the absence of the actual search queries issued on Twitter, we followed the methodology used in [122] of first identifying a seed set of queries and then expanding them to identify a larger set of potential queries. Our seed set consists of the queries *democratic debate*, and *republican debate*, and their shortened versions (*dem debate* and *rep debate*) popular on Twitter because of their short lengths.

We wanted our expanded set of queries to satisfy two properties: (i) they should be popular and be used by many users, and (ii) they should not be biased towards any particular party, candidate or organization in their formulation, *i.e.,* the leaning of the user issuing the query should not be obvious from the query.

To satisfy the first property of selecting popular queries, we focused on hashtags for expanding the query set. This choice was bolstered by the knowledge that hashtags are used extensively on Twitter to tag and follow discussions about politics [52]. Additionally, every time a user clicks on a hashtag, a Twitter search page with the hashtag as the query opens up, making hashtags effectively act as recommended queries on Twitter. To identify such popular hashtags, we collected the Twitter search results for our four seed queries during the November 2015 Republican and Democratic debates. We then identified top 10 most frequently occurring hashtags for each of the debate's dataset which contained the term "debate" in them (to ensure they are about the primary debates), resulting in a total of 15 distinct hashtags (*#debate, #demdebate, #democraticdebate, #republicandebate, #gopdebate, #debatewithbernie, #hillarycantdebate, #debatewithbe, #nprdebate, #cnndebate, #cnbcgopdebate, #fbngopdebate, #foxbusinessdebate, #gopdebatequestions, #gopdebatemoderators*).[3]

Due to our second desirable property, we wanted to retain only the unbiased queries from the above 15 hashtags, to avoid over-estimating the bias in the search results. Doing so, we removed queries which were biased towards (or against) a candidate (*#debatewithbernie, #hillarycantdebate, and #debatewithbe*)[4], an organization (*#nprdebate, #cnndebate, #cnbcgopdebate, #fbngopdebate, and #foxbusinessdebate*), or a party (*#gopdebatequestions, and #gopdebatemoderators*). Therefore, we were left with the expanded set of 8 queries which are popular and for whom it was hard to guess the

[3]We did not include *#debate* in our selected query dataset because it was too generic and many tweets containing it were about topics unrelated to 2016 US Presidential Primaries.

[4]For example, we observed that the hashtag #debatewithberine was biased towards Bernie Sanders (and the Democratic party), with #FeelTheBern, #BernieSaidItFirst and #Bernie2016 being the hashtags which co-occurred with #debatewithbernie the most.

political leaning of the user issuing the query – *democratic debate, dem debate, #democraticdebate, #demdebate, republican debate, rep debate, #republicandebate* and *#gopdebate*.

In addition, we also included the names of the 17 presidential candidates, resulting in a total of 25 queries, which we used to measure the bias for political searches on Twitter. Table 5.8 shows the exact phrasings of the 25 queries from our dataset.

5.3.1.2 Data collection from Twitter

For applying our bias quantification framework to Twitter search, we needed to collect data about the output search results given out by Twitter's ranking algorithm, as well as the set of tweets which were relevant to our selected queries that form the input to the ranking system. For performing our bias analysis, we collected the search data for a one week period in which both a Democratic debate (December 19, 2015) and a Republican debate (December 15, 2015) took place – 14 - 21 December 2015.

Even though Twitter provides multiple different filters for their search functionality, we collected the search snapshots for our set of selected queries for the default filter of "top" search results [9]. The "top" search results are the output of Twitter's proprietary ranking system, which performs ranking based on a multitude of factors, including the number of users engaging with a tweet [8]. During the one week period, search snapshots were collected at 10-minute intervals for each query. Each snapshot consists of the top 20 results on the first page of search results, and we used these snapshots to compute the output bias for the queries. Across all queries, we collected a total of $28,800$ snapshots which consisted of $34,904$ distinct tweets made by $17,624$ distinct users.

Finally, we used Twitter's streaming API to collect the tweets containing our selected queries during this one week period, and this set of tweets formed the input to the ranking system and were used to compute the input bias for the queries.[5] Across all queries, we collected more than 8.2 million tweets posted by 1.88 million distinct users.

Collecting non-personalized search results : In this work, we focus on quantifying the bias in consistent, non-personalized search results shown to every user, therefore to mitigate the personalization

[5]We observed that 74.8% of tweets included in the search results were also included in the data that we collected via the streaming API. In comparison, prior work that compared [157] data collected using Twitter's Streaming API with Twitter's Firehose (full Twitter stream), found that on average, the Streaming API contained 43.5% of data available on the Firehose on any given day.

effects we made all the search queries from the same IP subnet (in Germany), and without logging in to Twitter.

5.3.2 Measuring political bias of an individual search result

To apply our bias quantification framework to Twitter search for queries related to US presidential primaries, we need a methodology for inferring the political bias of an individual result – a tweet. The short length of tweets (140 characters) makes it very challenging to infer the bias of a tweet from its content (*i.e.,* to measure its *content bias*). Instead in this work, we operationalize the bias of a tweet as its *source bias*, *i.e.,* we approximate the bias of a tweet with the political bias of the author of the tweet.

In the rest of this section, we begin by presenting our methodology for inferring source bias of a tweet and then present our evaluation results. Finally, we end with a short analysis of how well source bias and content bias of a tweet match each other in practice for political searches on Twitter.

5.3.2.1 Source bias - Inferring political bias of Twitter users

Prior studies have shown that people's political affiliations are correlated with their personality attributes and responses to different stimuli [42, 187**?**]. Based on this knowledge, we propose a methodology for inferring political leaning of Twitter users by leveraging their interests. Therefore, our methodology for inferring the political bias of a Twitter user u, is based on the following three steps:

1. **Generating representative sets of Democratic and Republican users:** We use the crowd-sourced methodology described in [80, 211], which infers the topical attributes of a user v by mining the Twitter Lists that the other users have included v in. By relying on what others are reporting about a user, rather than what the users are identifying themselves as, we avoid the self-reportage problem, as well as avoid biasing the sets towards the group of users who have self-reported. Following this methodology, we identified a seed set of 865 users labelled as "Democrats" and 1348 users labelled as "Republicans". These seed sets include known politicians (*e.g.,* Steny Hoyer, Matt Blunt), political organizations (*e.g.,* DCCC, Homer Lkprt Tea-party) as well as regular users.

2. **Inferring topical interests of a user:** To infer the interests of a user u we rely on the methodology in [29, 30], which for a user u, returns a list of topics of interest of u along with the number

of users whom u follows who have been labeled with this topic using [80, 211]. Therefore, our method leverages the network neighborhood of u to infer the interests and hence the political leaning of u. For instance, if a user u follows three users tagged with 'politics' and four users tagged with 'entertainment', then the returned list would be {politics: 3, entertainment: 4}. We convert this <topic, #users> list into a weighted tf_idf vector for user u (where the idf-s are computed considering the interest lists of all the users in our dataset) and refer to it as the *interest-vector* I_u of the user u.

We are not able to infer the topical interests of a user when either their accounts are protected, and we can not gather the users they are following or because they follow too few other users (less than 10). But in prior work, it has been shown that such cases are few, and this methodology infers the interests of a significant fraction of active users on Twitter [29, 30].

3. **Matching user's interests to interests of Democrats and Republicans:** We first compute the representative interest vectors for Democrats (I_D) and Republicans (I_R) by aggregating the interest vectors of users in each set and normalizing such that I_D and I_R vectors sum up to 1 each. These aggregate vectors not only capture the differences in the political interests of Democrats and Republicans (*e.g.,* [progressive, democrats, obama, dems, liberals] & [patriots, conservative, tcot, right, gop] are the top terms in I_D and I_R respectively), but also the differences in their non-political interests (*e.g.,* I_R has higher weight for sports-related terms, while I_D has higher weight for technology and entertainment related terms). Therefore, even in the case of users who don't follow any politicians on Twitter or the ones who follow politicians from both parties, these representative vectors can be used to infer their likely political bias.

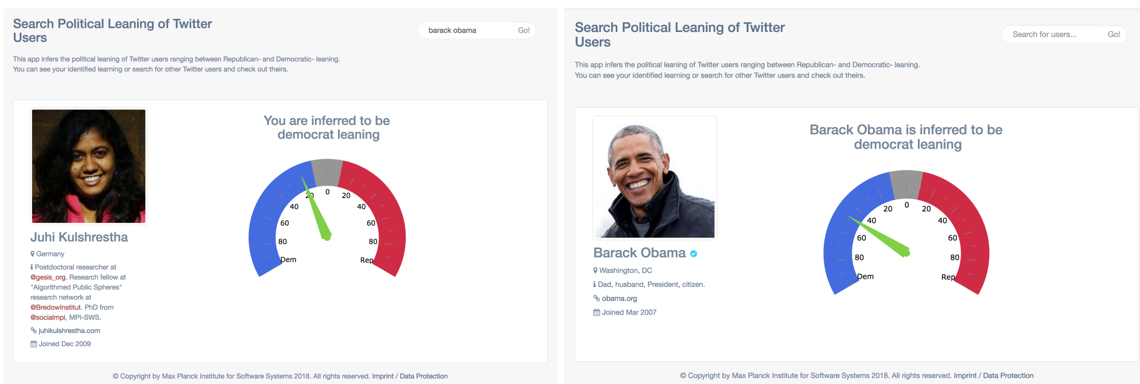
   Finally, the *bias score* of user u with interest vector I_u is given by the difference in the cosine similarities of I_u with I_D and I_R,

$$Bias(u) = cos\_sim(I_u, I_D) - cos\_sim(I_u, I_R). \tag{5.5}$$

   We max-min normalize the scores such that the bias score of a user lies in the range $[-1.0, 1.0]$, with a score closer to $+1.0$ indicating more Democratic bias, while a score closer to -1.0 indicating more Republican bias.

5.3.2.2 Public deployment of the source bias inference methodology

We have publicly deployed the aforementioned source bias inference methodology in the form of a Twitter application [184], at **http://twitter-app.mpi-sws.org/search-political-bias-of-users/**. One can log in to the application using their Twitter credentials, and see their inferred political affiliation. Figure 5.2(a) depicts a screenshot of our application showing my inferred political leaning. One can also search for other Twitter users to check out their inferred political leaning. Figure 5.2(b) depicts a screenshot of our application showing our inferred political leaning for the former US president 'Barack Obama'.



(a) **Me (the logged-in user)** (b) **Barack Obama**

Figure 5.2: Screenshots of our Twitter application for inferring political leaning of users showing the inferred political leaning for me and 'Barack Obama'.

Details about the functionalities of our application and a pointer to a demo video can be found in Appendix A.2.

5.3.2.3 Evaluation of political bias inference methodology

To validate whether our bias inference method works well for a whole spectrum of politically interested users, we perform the evaluation over three test sets of Twitter users – (i) politically interested common users, selected randomly from the set of users who have retweeted the two parties' accounts on Twitter, (ii) the current US senators, and (iii) self-identified common users (with fewer than 1000 followers), who have identified their political ideology in their account bios. We use two metrics for evaluating the methodology: (i) *coverage* – for what fraction of users can the methodology infer the political bias, and (ii) *accuracy* – for what fraction of users is the inference correct.

We begin by using the set of politically interested common users to evaluate our inferred bias scores, followed by a description of how we discretize our bias score into three distinct categories – Republican, neutral and Democratic, and we end by presenting our methodology's performance in inferring the political bias of senators and self-identified common users.

Evaluation for politically interested common users

Identifying politically interested common users: Following the methodology in [140], we collected up to 100 retweeters of each of the latest 3,200 tweets posted by the accounts of the two political parties – @TheDemocrats and @GOP. We removed the retweeters which retweeted the accounts of both the parties, obtaining 98,955 distinct retweeters of @TheDemocrats, and 71,270 distinct retweeters of @GOP. From each of these two sets of retweeters, we randomly selected 100 retweeters, giving us a total of 200 politically interested common users.

Ground truth bias of politically interested test users: We collected the ground truth bias annotations for these 200 politically interested users by conducting an AMT survey where human workers were shown a link to user's Twitter profile. We only used Master workers from the US who have had at least 500 HITS approved, with an approval rating of 95%. We paid the workers 4$ for judging the political leaning of 45 Twitter users. The workers were asked to infer the user's political leaning as either pro-Democratic, pro-Republican or neutral based on the user's profile and tweets. For each user, we aggregated the judgements of 50 workers, adding $+1$ for each pro-Democratic, -1 for pro-Republican and 0 for each neutral judgement and normalizing by the total number of judgements to get an AMT bias score in the range $[-1.0, 1.0]$, where a more positive score indicates a stronger Democratic bias, while a more negative score indicates a stronger Republican bias.

Evaluating our inferred score: With our methodology, we were able to infer the bias of all 200 users (*i.e.,* coverage is 100%). To quantify the accuracy of the methodology, we checked whether our inferred bias scores correlate well with the AMT bias scores. To verify this, we binned our inferred bias score into three bins: Bin 1 $[-1.0, -0.5]$, Bin 2 $(-0.5, 0.5)$, and Bin 3 $[0.5, 1.0]$ and computed the average AMT bias scores for each bin. We observe a strongly Republican leaning score (-0.86) for Bin 1, while a strongly Democratic leaning score (0.93) for Bin 3. We observe a similar trend if we bin according to

AMT bias score	**Inferred Rep**	**Inferred Neutral**	**Inferred Dem**
AMT Bin 1	84.05%	13.04%	2.89%
AMT Bin 2	18.18%	45.45%	36.36%
AMT Bin 3	3.89%	12.98%	83.11%

Table 5.2: Confusion matrix of the match between AMT bias scores and Inferred bias scores.

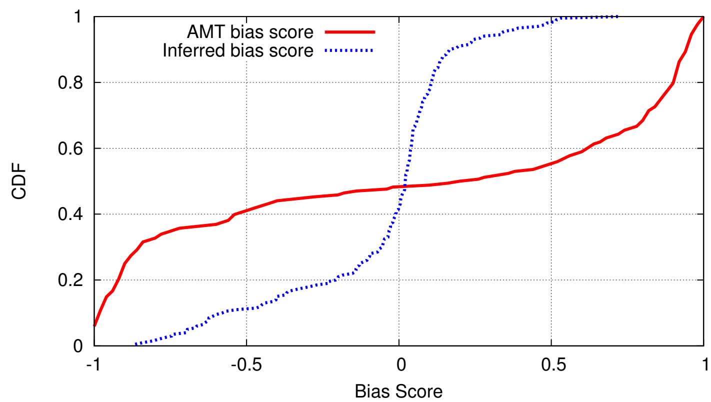


Figure 5.3: CDF of AMT bias scores and Inferred bias scores for politically interested common users.

the AMT bias scores and compute the average inferred score for each bin (-0.32 for Bin 1 and 0.14 for Bin 3), demonstrating a good correlation between the two bias scores.

Discretizing the bias score into categories: While the inferred bias scores are highly correlated with the AMT bias scores, we observe that the distribution (CDF) of the two scores in the interval $[-1.0, 1.0]$ are different, as shown in Figure 5.3. Due to this difference in the distributions of the two scores, we decided to discretize our inferred bias score, and categorize users as - neutral, Democratic or republic leaning.

In order to do the discretization, we needed to identify a suitable threshold x on our inferred score, such that users with scores in the range $(-x, x)$ are categorized as neutral, while the ones with scores x and above are identified as Democratic leaning, while $-x$ and below are identified as Republican leaning. We experimented with x = 0.01, 0.03, 0.05, 0.08 and 0.1, and for each of these values computed a confusion matrix of the match between the AMT bias score and our inferred bias score. We selected $x = 0.03$ to be the threshold as it maximizes the sum of the diagonal of the confusion matrix, as shown in Table 5.2. In the rest of this section, we will only label the users as Republican or Democratic leaning when their bias scores lie outside of the neutral zone $(-0.03, 0.03)$. We make this conservative choice to not overestimate the bias in the search results.

Political Bias	Coverage	Accuracy
Current US Senators		
Democratic (n=45)	97.78%	86.36%
Republican (n=54)	98.15%	98.11%
Average	97.96%	92.23%
Self-identified common users		
Democratic (n=426)	92.01%	88.52%
Republican (n=675)	90.22%	82.95%
Average	91.12%	85.73%

Table 5.3: Coverage and accuracy of the political bias inference methodology for (i) current US senators, and (ii) common users who have declared their political ideology in their Twitter account profiles.

Evaluation for US senators

Table 5.3 outlines the performance of our methodology for the 100 current US senators (45 Democrats, 54 Republicans, 1 Independent), showing that our methodology has very high coverage. Closer inspection of the two senators, for whom we could not infer the bias, disclosed that one them does not follow any other users on Twitter while the other follows only one, making it impossible for us to infer their interests and consequently their bias. Our methodology also performs well in terms of accuracy by correctly identifying the bias for 86.4% of Democratic senators and 98.1% of Republican senators, out of the ones for whom we could infer the bias.

Evaluation for self-identified common users

We collected our final set of self-identified common users using the service Followerwonk and gathering users located in the US, with less than 1000 followers, and whose Twitter account biographies contained keywords matching Democrats ('democrat', 'liberal', 'progressive') or Republicans ('republican', 'conservative', 'libertarian', 'tea party'). We manually inspected each user, and pruned out any users whose bios did not reflect their political ideology, for instance, users with erroneous bios like "*I am a #conservative #Christian who is neither a #Democrat nor a #Republican, but an #Independent voter*" and "*We hate Politicians - Democrats, Republicans, all of them.*" were removed. Following this procedure, we collected a total of 426 self-identified Democratic users, and 675 self-identified Republicans.

Gold standard Tweet bias	Source bias		
	Republican	**Neutral**	**Democratic**
Republican [-1.0, 0.5]	**70.44%**	9.36%	20.2%
Neutral (-0.5, 0.5)	27.61%	16.96%	55.43%
Democratic [0.5, 1.0]	11.71%	10.24%	**78.05%**

Table 5.4: Confusion matrix for source bias classification – gold standard tweet bias (based on AMT workers' judgement) vs. source bias.

Table 5.3 also depicts the performance of our methodology for these self-identified users. The average coverage is again high (91.1%), with the users for whom we could not infer the bias either having protected accounts or following too few users such that it was impossible for us to infer their interests and therefore their political bias. Our proposed method also has a high accuracy of 85.7% on average across all these self-identified common users for whom we could infer the bias.

Further inspection of interest vectors of the users for whom we correctly inferred the political leaning reveals that the interest vectors of Democratic users not only contain political terms like 'liberal', 'progressive', and 'dem', but also other terms including 'gay', 'lgbt', 'science', and 'tech', while the interest vectors of Republican users contain terms like 'tea', 'gop', and 'palin' along with other related terms like 'patriots', 'military', and 'vets'.

5.3.2.4 Match between source bias & tweet bias

In this section, we focus on answering the question, "*how closely do source bias and bias of a tweet reflect each other?*".

Measuring tweet bias: For each of our selected queries, we gathered two search snapshots from our chosen period in December 2015, one during the Republican debate and one during the Democratic debate. Across all these snapshots, we gathered a total of 881 distinct tweets, and we use these to evaluate the extent to which the tweet bias matches the inferred source bias. We use AMT workers to measure the tweet bias by showing each tweet (but not the user who posted it) to 10 AMT workers and asking them to label the tweet as either pro-Democratic, pro-Republican or neutral. Then following the methodology in Section 5.3.2.3, we computed a tweet bias score for each tweet by aggregating the judgments of the 10 AMT workers. Using these scores, we generated the gold standard labels for the bias of the tweets, by dividing the range of AMT tweet bias scores into 3 intervals and labelling

Gold standard Tweet bias	Content bias (SVM)		
	Republican	**Neutral**	**Democratic**
Republican [-1.0, 0.5]	**39.11%**	40.22%	20.67%
Neutral (-0.5, 0.5)	16.67%	76.19%	7.14%
Democratic [0.5, 1.0]	24.35%	50.00%	**25.65%**

Table 5.5: Confusion matrix for content bias classification (Support Vector Machine (SVM) classifier) – gold standard tweet bias (based on AMT workers' judgement) vs. content bias.

Gold standard Tweet bias	Content bias (GBDT)		
	Republican	**Neutral**	**Democratic**
Republican [-1.0, 0.5]	**79.88%**	11.17%	8.94%
Neutral (-0.5, 0.5)	57.14%	35.72%	7.14%
Democratic [0.5, 1.0]	64.10%	8.97%	**26.93%**

Table 5.6: Confusion matrix for content bias classification (Gradient boosted decision tree (GBDT) classifier) – gold standard tweet bias (based on AMT workers' judgement) vs. content bias.

tweets in interval $[-1.0, 0.5]$ as Republican, in interval $(-0.5, 0.5)$ as neutral, and in interval $[0.5, 1.0]$ as Democratic-leaning.

How closely do source bias and tweet bias match each other?: To investigate the match between source bias and tweet bias, Table 5.4 presents the confusion matrix for our source bias inference methodology. We observe that when the content is biased on either side, the match between source and AMT gold standard tweet bias is high (70% or more) indicating that strongly biased content is produced mostly by users with the same bias.

How does our source based scheme compare to content-based scheme for inferring the bias of a tweet?: To evaluate how well does a content-based scheme work for inferring bias of social media posts, especially in comparison with our source based methodology, we represented the tweets by a bag-of-words model (*i.e.,* using every distinct unigram as a feature) and applied two well-known classifiers – Support Vector Machine (SVM) and Gradient Boosted Decision Tree (GBDT). The unigram features were generated from the tweet text by applying preprocessing steps of case-folding, stemming, stop word removal and removal of URLs. We used 5-fold cross validation for all the classification experiments.

Tables 5.5 and 5.6 depict the confusion matrices for the SVM and GBDT content based classifiers respectively. Comparing with Table 5.4, we observe that our source based method performs better than the content based scheme. While the accuracy for SVM classifier is quite low, the GBDT classifier seems to classify most tweets as Republican. However, we want a classifier where errors for the different classes are balanced, so that one class is not grossly over-estimated, and from this perspective also our source-based classification performs better.

5.3.3 Characterizing the bias for political searches on Twitter social media

Having described our bias inference methodology, as well as the search data that we collected for political searches on Twitter social media, we next focus on analyzing the collected data to characterize the bias for political searches on Twitter. We begin by investigating the contributions of the two sources of bias – input data and ranking system – to the final output bias seen by the users. Then we examine the interplay between the input data and the ranking system that produces the output bias seen by the users. We end with an analysis of the variation of bias over time.

5.3.3.1 Where does the bias come from – Input data or ranking system?

It is not always the ranking system, input data matters: We show the three biases (output, input and ranking bias) for all our selected queries in Table 5.7. When we compute the average biases for the four sets of queries – Democratic and Republican candidate and debates – we find that the average input biases for all four sets are Democratic-leaning (*i.e.,* larger than 0). Although the average input bias for Republican candidates and debates is less Democratic-leaning than Democratic ones, the full tweet stream containing all these query terms (without any interference from the ranking system) on an average contains a more Democratic slant. We observe that the input bias proves to be a prominent contributor to the final output bias seen by the users. For instance, the output bias for *Bernie Sanders* is very Democratic (0.71), with only a small amount of the bias being contributed by Twitter's ranking system (0.16); the majority of bias originates from the input data (0.55), indicating that most of the users that discuss *Bernie Sanders* on Twitter have a Democratic leaning. The effect of input data on the output bias highlights the importance of also taking into account the input data while auditing algorithms, to discern how much of the bias is due to the data and how much is contributed by the algorithmic system. This insight is particularly crucial in this digital era where many algorithms are trained using vast amounts of data [26].

We also measured the *bias of overall Twitter corpus* in two ways: (i) *User population bias*: measured as the average bias of 1000 Twitter users selected randomly from the Twitter user-id space (*i.e.,* the user-ids were randomly selected from the range of 1 through the id assigned to a newly created

Query	Output Bias (TOB)	Input Bias (TIB)	Ranking Bias (TRB)
Queries Related to Democratic Candidates			
Hillary Clinton	0.21	0.03	0.18
Bernie Sanders	0.71	0.55	0.16
Martin O'Malley	0.64	0.57	0.07
Average	0.52	0.38	0.14
Queries Related to Republican Candidates			
Donald Trump	0.29	0.19	0.10
Ted Cruz	-0.48	-0.11	-0.37
Marco Rubio	-0.41	-0.12	-0.29
Ben Carson	0.46	0.20	0.26
Chris Christie	-0.14	0.27	-0.41
Jeb Bush	-0.31	0.09	-0.40
Rand Paul	-0.37	-0.18	-0.19
Carly Fiorina	0.16	0.38	-0.22
John Kasich	-0.09	-0.13	0.04
Mike Huckabee	0.30	0.12	0.18
Rick Santorum	-0.04	0.18	-0.22
Lindsey Graham	-0.45	0.07	-0.52
George Pataki	-0.17	0.09	-0.26
Jim Gilmore	-0.35	-0.11	-0.24
Average	-0.11	0.07	-0.18
Queries related to Democratic debate			
democratic debate	0.43	0.38	0.05
dem debate	0.52	0.29	0.23
#democraticdebate	0.28	0.19	0.07
#demdebate	0.57	0.56	0.01
Average	0.45	0.35	0.10
Queries related to Republican debate			
republican debate	0.53	0.27	0.26
rep debate	0.31	0.40	-0.09
#republicandebate	0.39	0.34	0.05
#gopdebate	0.04	0.10	-0.06
Average	0.32	0.28	0.04

Table 5.7: Time-averaged bias in Twitter search "top" results, for selected queries (related to political candidates and debates) – output bias TOB, input bias TIB, and ranking bias TRB. Here a bias value closer to $+1.0$ indicates Democratic bias and a value closer to -1.0 indicates Republican bias.

account in December 2015), and (ii) *Full tweet stream bias*: measured as the average source bias of 1000 tweets selected randomly from Twitter's 1% random sample for December 2015. We found the user population bias to be 0.25 and a full tweet stream bias to be 0.3 indicating that not only is the population of Twitter Democratic-leaning, but the active users (whose tweets have been included in Twitter's 1% random sample) are even more Democratic-leaning. These findings are in-line with prior studies [182] which have shown that Twitter has a high fraction of Democratic-leaning users.

Although Twitter has a Democratic-leaning corpus bias, the input bias (TIB) of the different queries varies across the spectrum (as shown in Table 5.7). This variation in bias likely occurs because each query acts as a filter to extract a subset of Twitter users whose tweets are relevant to that query, and the sets of users filtered out by different queries have differing biases. Therefore, even with the corpus bias of Twitter being the same, each query determines the input data set and hence the input bias, which in turn affects the final output bias observed by the user for that query.

The Power of the Ranking System: Although input data does contribute to the final output bias, the ranking system also exerts power over the final bias by shifting the bias or even changing its polarity, as demonstrated by the ranking biases shown in Table 5.7. Even though we observed that the input biases for both the Democratic and Republican candidates on an average were Democratic-leaning, we notice that on an average the ranking system adds a Democratic-leaning ranking bias for the Democratic candidates making the output more Democratic-leaning ($TOB = 0.52$), while it adds a Republican-leaning ranking bias for Republican candidates making the output more Republican-leaning ($TOB = -0.11$). This change of polarity from a Democratic-leaning input bias to a Republican-leaning output bias is particularly noticeable for some Republican candidates like *Chris Christie*, *Jeb Bush* and *Lindsey Graham*. These shifts in the bias caused by the ranking system (that can also result in a polarity change), exhibit the ranking system's power in altering the inherent bias of the input data.

The ranking of posts in social media search systems is a complex process with the platform providers trying to provide the most relevant posts within the highest ranked items. They use a number of factors to measure the relevance of posts for ranking search results, including the keywords it contains, the popularity of the post in terms of users' engagements with it (*e.g.,* number of retweets, favorites or replies) [8, 5], as well as the recency of the post [6]. Our goal in this thesis is not to reverse engineer Twitter's ranking system. However, we take a step towards gaining insight into the ranking system

Query	TRB of Ranking Strategies		
	Twitter's Ranking	Most Retweeted First	Most Favorited First
Queries Related to Democratic Candidates			
Hillary Clinton	0.18	0.33	0.25
Bernie Sanders	0.16	0.22	0.16
Martin O'Malley	0.07	0.001	0.1
Queries Related to Republican Candidates			
Donald Trump	0.10	0.06	0.09
Ted Cruz	-0.37	-0.49	-0.35
Marco Rubio	-0.29	-0.36	-0.27
Ben Carson	0.26	0.23	0.25
Chris Christie	-0.41	-0.40	-0.34
Jeb Bush	-0.40	-0.46	-0.34
Rand Paul	-0.19	-0.25	-0.17
Carly Fiorina	-0.22	-0.17	-0.18
John Kasich	0.04	0.04	0.11
Mike Huckabee	0.18	0.11	0.19
Rick Santorum	-0.22	-0.34	-0.16
Lindsey Graham	-0.52	-0.45	-0.56
George Pataki	-0.26	-0.22	-0.23
Jim Gilmore	-0.24	-0.22	-0.21
Queries related to Democratic debate			
democratic debate	0.05	0.21	0.12
dem debate	0.23	0.22	0.22
#democraticdebate	0.07	0.08	0.14
#demdebate	0.01	-0.01	0.01
Queries related to Republican debate			
republican debate	0.26	0.274	0.268
rep debate	-0.09	-0.09	-0.09
#republicandebate	0.05	0.08	0.17
#gopdebate	-0.06	-0.06	-0.02

Table 5.8: Time-averaged ranking bias for different ranking strategies: (i) Twitter's ranking (Twitter search "top" results), (ii) Most retweeted tweet first ranking, and (iii) Most favorited tweet first ranking. Here a bias value closer to $+1.0$ indicates Democratic bias and a value closer to -1.0 indicates Republican bias.

of Twitter by examining the impact of the popularity of the posts on the search rankings. For doing so, we take the posts included in Twitter's top search results and rerank them based on the popularity of the post (*i.e.,* the number of retweets and the number of favorites). We then compared the bias of these simulated rankings with the ranking bias of Twitter's ranking system (shown in Table 5.8). For most of our queries, the ranking biases of the three strategies are quite similar to each other, indicating that popularity of the post can explain much of the observed bias in Twitter's ranking. However, in case of some queries (*e.g., Martin O'Malley*, *John Kasich*, *democratic debate* and *#republicandebate*), the difference in the ranking bias values between Twitter's ranking and the popularity based rankings indicates that there are probably other factors also that contribute to the overall bias of the search results. Note that this analysis is just a first step towards understanding the influence of the different factors on the overall bias of search results and we defer a more in-depth analysis for the future.

5.3.3.2 Collective contribution of the input data and the ranking system

Having observed that both the input data and the ranking system contribute prominently to shape the final output bias seen by the users, we next explore the dynamics between these two sources of bias. Here, we discuss two cases in which the interplay between the input and ranking biases lead to an output bias which can noticeably affect a user's search experience.

The case of popular candidates: Comparing the output biases for the candidate queries in Table 5.7, we found that the search results for the more popular candidates have a higher bias towards the opposing perspective.[6] For example, the top search results for the most popular Democratic candidate – *Hillary Clinton* – contained lesser Democratic-leaning results than other Democratic candidates, while the results for the most popular Republican candidate – *Donald Trump* – contained fewer Republican-leaning results as compared to other Republican candidates. In Figure 5.4, we plotted the output bias for the Republican candidates ranked by their popularity. The negative slope of the line of best fit the figure seems to suggest that the more popular a candidate is, the more is the opposing perspective in their top search results (however we are limited in the number of data points to be able to make any statistical inferences).

[6]The popularity of a candidate is estimated from the polling data obtained from [185] for December 2015.

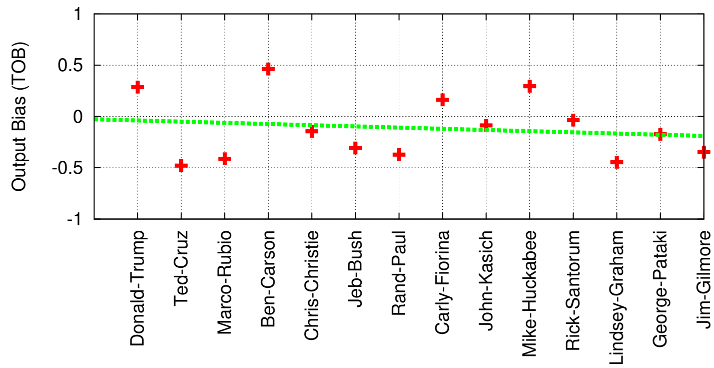


Figure 5.4: The time-averaged output bias TOB in Twitter "top" search results for the Republican candidates – candidates are listed left to right from highest to lowest popularity.

This situation may be undesirable for popular candidates, especially if users from the opposite perspective are more likely to speak negatively about the candidate and indeed this is what we find. Table 5.9 shows tweets randomly sampled from the set of tweets included in the top search results for a candidate, which were posted by users with an opposing polarity as compared to the candidate and they all either criticize or ridicule the candidates. Such negative tweets could alter the opinions of undecided voters [63] and thus the situation is less than ideal for the popular candidates.

When we examine the input biases for *Hillary Clinton* and *Donald Trump*, we observe that they too lean towards opposite leaning indicating that opposite leaning users are more likely to talk about the popular candidates as opposed to less popular candidates. However, we observe that the ranking system altered input bias for the two most popular candidates in different manners – while the ranking system improved the situation for *Hillary Clinton* by adding a Democratic-leaning ranking bias and directing the search results towards her own party's perspective, it does the opposite for *Donald Trump* by adding Democratic-leaning ranking bias and thus increasing the opposite leaning bias for him. These opposing interplay between the input data and ranking system (though possibly inadvertent) can have serious implications for the candidates, especially the one for whom the ranking system made the tweets of opposite leaning users more visible in the final output search results.

Different Phrasings of Similar Queries: While looking for information about the same topic, different users may use different phrasings of the query. For instance, for searching for the event Republican debate, users can use different queries like *republican debate*, *rep debate*, *#republicandebate* or *#gopdebate*. If users from different leanings preferentially use different keywords, phrases or hashtags to refer

Randomly selected tweets from search results for *Hillary Clinton*, which are posted by a Republican-leaning user	Randomly selected tweets from search results for *Donald Trump*, which are posted by a Democratic-leaning user
WT: Watchdog wants federal ethics probe of Clinton, possible improprieties http://bit.ly/1NvlrPA	Williamsburg, #Brooklyn Dec 15 #trump2016 #MussoliniGrumpycat #MakeAmericaHateAgain #DonaldTrump @realDonaldTrump pic.twitter.com/Hj6DC7M7V1
The Clintons both Bill and Hillary have a very long history of framing others while they commit the Crimes. History has destroyed the proof	Scotland defeats Trump on clean energy. Hopefully hell have a lot of time for golfing soon [url]
@CarlyFiorina: @realDonaldTrump is a big Christmas gift wrapped up under the tree for @HillaryClinton. [url]	Dirty little secret: Donald Trump is not a good debater.
@CNN @HillaryClinton @BernieSanders hell no shes a murderer pic.twitter.com/zGQwR7dLZj	http://MLive.com - Where Donald Trumps Michigan campaign donations come from http://ow.ly/39hCWt
I dont care if youre a Democrat or Republican, how can you trust a word Hillary Clinton says and how can you consider voting for her??	Enjoy the sweet music of Donald Trump in Carol of the Trumps [url]

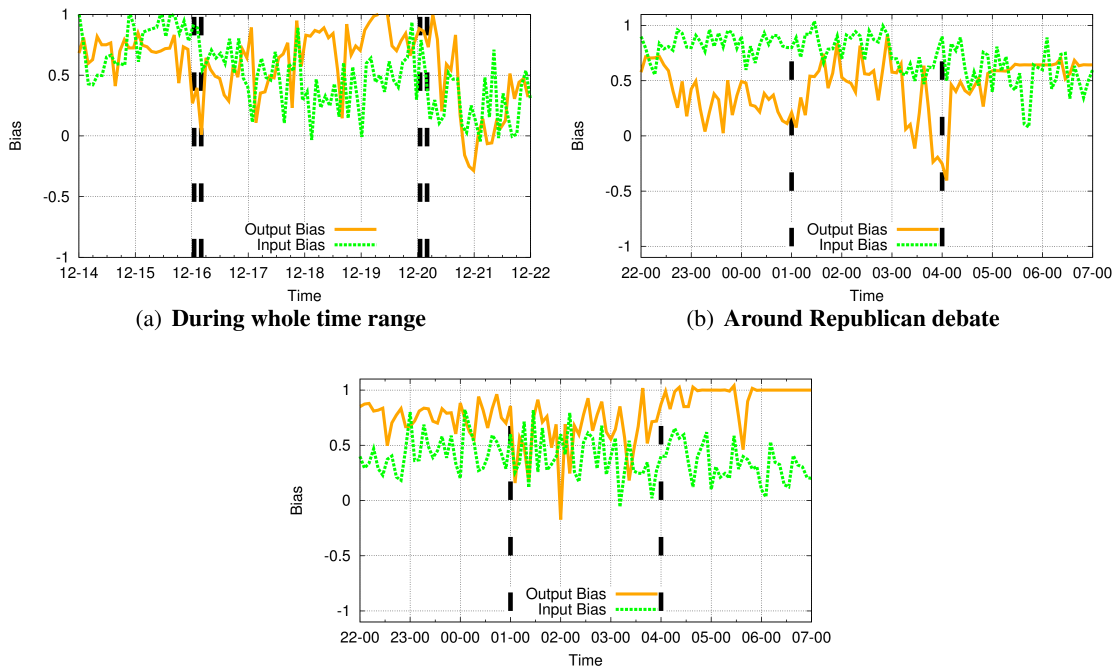
Table 5.9: Randomly selected tweets from the search results for the queries *Hillary Clinton* and *Donald Trump*, which are posted by a user with an opposite bias as compared to the candidate.

to the same event in their tweets, then this might lead to differing biases for these differently phrased queries about the same event. To investigate whether different phrasings of the query about the same event lead to different biases, we compare the bias values for the queries related to Democratic and Republican debates, shown in Table 5.7. The first thing we observe is that the output biases for similar queries are noticeably different. For instance, the output bias of *republican debate* ($TOB = 0.53$) has a lot more Democratic-leaning bias than the query *rep debate* ($TOB = 0.31$), while the bias in search results for *#demdebate* ($TOB = 0.57$) are much more Democratic-leaning than bias for the query *#democraticdebate* ($TOB = 0.28$).

When we examine the input and ranking biases for our similarly phrased queries from Table 5.7, we observe that for most of them, the input bias is the more prominent contributor to the final output bias. However, in some cases even when the input biases are similar, as in the case of queries *rep debate* ($TIB = 0.40$) and *republican debate* ($TIB = 0.27$), the ranking system shifts their biases in opposite directions, by adding a Democratic-leaning ranking bias for *republican debate* ($TRB = 0.26$), while a Republican-leaning ranking bias for *rep debate* ($TRB = -0.09$). This example illustrates the power the ranking system exerts on the input data, which can lead to search results for similar queries with similar input biases having different output biases. These observations about different biases for similar queries raise questions about the impact that features like autocomplete queries and suggested queries can have on the bias that the users see, and what mechanisms can be designed to make the users aware of these effects. These are open research questions that can be pursued in the future and in Section 5.7 we briefly discuss some solutions for signaling the bias in the search results to the users.

5.3.3.3 Variation of bias over time

Finally, we explore whether the bias in the search results for a particular query varies with the time at which the query is issued. As described earlier, we collected the Twitter top search results for our selected queries at 10-minute intervals during the period December 14–21, 2015, which included both a Republican debate (December 15) and a Democratic debate (December 19).



(a) **During whole time range**

(b) **Around Republican debate**

(c) **Around Democratic debate**

Figure 5.5: Temporal variation of output and input bias for the query *dem debate* – (a) variation across the full duration over which we collected data (December 14–22, 2015), (b) variation during a 9-hour window around the Republican debate on December 15, 2015, (c) variation during a 9-hour window around the Democratic debate on December 19, 2015.

To illustrate how the bias in the search results for a query varies with time, Figure 5.5 shows the variation in the output and input biases for the query *dem debate* during the entire one week period (Figure 5.5(a)), during a 9-hour interval around the Republican debate (Figure 5.5(b)), and during a 9-hour interval around the Democratic debate ((Figure 5.5(c)). We observe noticeable variation in the bias over time. The variation is lower for the input bias because we compute input bias over cumulative sets of tweets and hence it is less affected by instantaneous events. However, the variation in output bias is much higher, especially during and immediately after the debate events. (Fig. 5.5(b) and Fig. 5.5(c)). In Table 5.10, we present the statistical analysis for the temporal variations in the output bias for the query *dem debate* (corresponding to Figure 5.5). The first row of Table 5.10 shows the comparison between the output bias values for the search snapshots for the query *dem debate* in the 3 hour period before the start of the Republican debate and the 3 hour period after the end of the Republican debate. For comparison, we computed the significance of difference by performing paired t-test and determining the p-value for 95% confidence interval, and also computed the values of Cohen's d and effect size r. Similarly, the second row in the table shows these values for the 3 hour period before and after the Democratic debate. For both the debates, we find that the differences in output bias before and after the debate are statistically significant with medium to large effect sizes. We observed similar statistically significant temporal differences for other debate-related queries too.

We also observe another common trend in the variation of bias across different queries. The output bias for most debate related queries shifted down (towards the Republican perspective) during the Republican debate when possibly a larger number of influential or popular Republican users were actively posting on Twitter. Correspondingly, the output bias for most debate related queries shifted up (towards the Democratic perspective) during the Democratic debate. This trend is visible in Figure 5.5(b) and Figure 5.5(c) for the query *dem debate*, and we observed similar trends for most other debate-related queries. The third row in Table 5.10, shows the comparison between the output bias values for the search snapshots for the query *dem debate* during the 3 hour period during the Republican debate and the 3 hour period during the Democratic debate. Again, we observe that difference between the two is statistically significant (with medium effect size) with the output bias being lower (more Republican-leaning) during Republican debate than during the Democratic debate. Therefore, we find that which perspective is reflected more in the top Twitter search results varies with the time at which the query is issued.

Output Bias across 3-hour time periods	T1 Mean (Std_dev)	T2 Mean (Std_dev)	Paired t-test df	Paired t-test p-val	Cohen's d	Effect size r
Before Rep debate (T1) vs. after Rep debate (T2)	0.3783 (0.2025)	0.5189 (0.2415)	35	0.0380	−0.6309	−0.3008
Before Dem debate (T1) vs. after Dem debate (T2)	0.7675 (0.1133)	0.9576 (0.1164)	35	0.0000	−1.6550	−0.6375
During Rep debate (T1) vs. during Dem debate (T2)	0.4200 (0.2974)	0.6089 (0.2565)	35	0.0034	−0.6802	−0.3219

Table 5.10: Statistical analysis of temporal variation of output bias for the query *dem debate* – (i) variation in output bias across 3 hours before and 3 hours after the Republican debate on December 15, 2015, (ii) variation in output bias across 3 hours before and 3 hours after the Democratic debate on December 19, 2015, (iii) variation in output bias across the 3 hour time periods during the Republican debate (on December 15, 2015) and the Democratic debate (on December 19, 2015)

5.4 Comparing relative bias in political searches on the Web and social media

Next, we apply our bias quantification framework to compare the relative biases of political searches on two different search systems — Twitter social media search and Google Web search. This second study highlights another useful application scenario for our bias quantification framework where we can observe the output search results, but we do not have access to the input data to the ranking system (as is the case with most commercial search systems). This unavailability of input data makes it infeasible to disentangle the effect of input data and ranking system by measuring input bias and ranking bias separately, however, we can still compare the relative biases of different search systems.

Our choice of the two search systems to compare (Google and Twitter search) was driven by the fact that these are two popular channels by which internet users are finding news and information on the Web. Traditional media channels like Fox News or CNN have often been scrutinized by academics [197, 20, 41, 79, 86, 27, 161] as well as media watchdog groups (like FAIR (`fair.org`) and AIM (`aim.org`)) for fairness, accuracy and balance in the news they report. Additionally, tools have also been developed to mitigate or expose the media bias [188, 146, 178, 161, 3] to users. However, the relative biases of newer digital channels like search systems are not as well studied and documented as yet, and thus users may not be taking their relative biases into account while selecting where to get their information from. With this study, we aim to highlight the differences in the bias of these two popular search systems – Twitter social media search and Google Web search. To have a fair comparison, we compare the Google search results with Twitter 'news' search results, both of which frequently contain results from news media sources [9].

5.4.1 Query selection & data collection

5.4.1.1 Collecting Google Web search data

We collected the top 20 Google search results for the queries stated in Section 5.3.1.1.[7] The results were collected at 10-minute intervals during the period December 14–21, 2015, gathering a total of 714 distinct web-links across all the queries. As was done while collecting the Twitter search results, to minimize any personalization effects, all the Google search results were collected without logging in to Google, and from the same IP subnet in Germany.

Note that in the case of Web search, it is infeasible to gather the set of all relevant web-links for a query. Therefore we did not attempt to measure the input and ranking bias separately. Instead, we used the collected search snapshots to measure the bias in the output.

5.4.1.2 Collecting Twitter news search data

Following the methodology described in Section 5.3.1.2, we collected the first page of top 20 "news" search results for each query at 10-minute intervals for the whole period. In total, across all the selected queries, Twitter news search results contained tweets posted by $7,512$ distinct accounts, an order of magnitude more than the number of distinct web-links in the dataset. We used these output search results to measure the output bias for Twitter news search.

5.4.2 Measuring political bias of a search result

For applying the bias quantification framework, we need a methodology for inferring the political bias score of each data item. Next, we describe how we measured the political bias of Google search results and Twitter news search results.

[7] We did not consider the hashtags as queries in this case, since hashtags are usually popular only on social media.

5.4.2.1 Measuring bias of Google search results

We observed that the top Google search results for our chosen set of queries (US presidential debates/-candidates) contained a significant fraction of links from *news media websites* for which the political biases have been well-documented [27, 79, 86, 161]. We use the results from Balance study [161] which identified the political bias of a large number of popular news media sources, to infer the political bias of the news media links in the web search results. We mapped the URLs in the search results to media sources in the Balance list [162], by considering the longest matching substring.

Apart from links from news media sources, Google search results also frequently contain Wikipedia articles, and personal websites and social media accounts of the political candidates (as also observed in [227]). We considered all Wikipedia URLs to have a zero or neutral bias[8], all personal websites of the candidates to have their own leanings (*e.g.,trump.com*, the website of Donald Trump, gets labelled as Republican), and all the social media profile links of the candidates to have their own leanings (*e.g.,* the links to the Facebook, Twitter, Instagram accounts of Bernie Sanders are labelled as Democratic). Following this procedure, we were able to infer bias for 86% of the top Google search results on an average across all the queries. The rest of the domains, for which we did not attempt to infer bias, are mostly political facts websites (*e.g.,* ontheissues.org, ballotopedia.org), informative websites (*e.g.,* biography.com), or government websites (*e.g.,* *.gov pages).

5.4.2.2 Measuring bias of Twitter news search results

To have a fair comparison between Google and Twitter news search results, we switch our methodology to infer the political leaning of Twitter results in this section and utilize Balance scores [162] for them too. We observed that the 7512 accounts which were included in the Twitter news search results include not only news media sources and journalists, but also other users like politicians and even academicians; hence, there was no way to match all these accounts to Balance scores. Therefore, we ranked these accounts based on their frequency of occurrence in the Twitter news search results for all the queries and tried to manually map the top 200 accounts (which account for 63% of all the Twitter news search results) to Balance scores. Additionally, we attempted to match the 100 of the most influential media

[8]Given Wikipedia's policy of neutral point of view [11], we make this simplifying assumption. Though sometimes Wikipedia does contain misinformation, prior work [128] has shown that most hoaxes are quickly detected and have little impact on Wikipedia.

accounts on Twitter [40] to Balance scores as well. Twitter news results also contained posts from journalists and political workers, and there was no way to map them to Balance score, so we manually labeled such accounts with their self-declared leaning from their profile bios (whenever available). Finally, as before, we marked the Twitter accounts of the presidential candidates with the candidate's own bias. By following this methodology, we were able to get the bias annotations for 155 media accounts on Twitter, which cover 45% of the Twitter news search results on average across the different queries.[9]

5.4.3 Comparing relative biases of Google search and Twitter news search

Our analysis shows three interesting ways in which the search bias for political queries on Google web search differs from that for Twitter social media search: (i) first, we investigated the temporal dynamics of the bias in the search results on the two systems and found the bias in social media search results to be significantly more dynamic across time, (ii) next, we compared their time-averaged output bias values to capture the overall trend and observed that for Google search the bias for most queries matches the leaning of the person or event being queried for, while the bias of Twitter news search for most queries is Democratic-leaning, and (iii) finally, we noticed that on Google search, a much higher fraction of search results are candidate-controlled sources (*e.g.,* candidate's website or social media accounts), leading to more favorable results for the candidates on Web search than on social media search. Next, we elaborate on each of our findings about the differences in bias of Google Web search and Twitter social media search.

5.4.3.1 Temporal variation in search bias

We began by comparing the two search systems along the temporal aspect by computing the standard deviation in the output biases of search result snapshots across time, for the different queries. We observe that the Google web search results are much more stable over time with a mean standard deviation of 0.046 in the output bias across all snapshots of all queries, while the standard deviation for

[9] The political leaning inferred by the source bias method and the Balance score based method match for 76% of these 155 media accounts. Here, we ignored the 9% of cases where our source bias methodology inferred the political leaning as neutral, which lead to a mismatch since the Balance Score does not output a neutral leaning.

Twitter news search results is an order of magnitude higher at 0.452, highlighting their highly dynamic nature in comparison.

5.4.3.2 Higher Democratic bias in Twitter news search results

Next, to compare the overall trend in relative biases of the two search systems, we computed the time-averaged output bias (TOB) for all the queries on Google and Twitter news search, which are shown in Table 5.11. As can be observed from the table, there is a striking difference between the two – the TOB values for Twitter news search are positive (*i.e.,* more Democratic-leaning) for most of the queries, including many of the Republican candidates, while the TOB values for the Google search results in most cases match the leaning of the candidate or event being searched for. So although the average TOB values for Democratic candidates are Democratic-leaning for both systems, the average output bias for Republican candidates is Republican-leaning (TOB = -0.264) for Google, while it is on the positive side (TOB = 0.083) for Twitter news search results.

This difference between Google and Twitter news search results may be due to the larger fraction of Democratic-leaning users on Twitter as indicated by the Democratic-leaning corpus bias we computed in Section 5.3, as well as the Democratic-leaning input bias TIB values for most queries reported in Table 5.7. These bias values mean that not only are there more Democratic-leaning users on Twitter, but the users tweeting about many of our queries are also Democratic-leaning. These results hint at the tremendous influence that corpus and input data have on determining the final output bias.

5.4.3.3 Favorable bias on Google search via candidate-controlled sources

When we dug deeper, we found that another potential reason for the differences in the relative bias in Google search and Twitter news search results for a particular candidate is the difference in the *fraction of search results that come from sources controlled by the candidate themselves*. For the Google search results, a significant fraction – 24.48% on average across all queries – of the results for the presidential candidates are from sources they control, *i.e.,* either their personal websites or their social media profile links (*e.g.,* for Donald Trump, we consider the webpage *trump.com* and his Twitter profile link *https://twitter.com/realDonaldTrump* to be sources controlled by him). A similar result is also reported in [227]. This fraction is much smaller for most candidates on Twitter – across all the presidential candidates, only 7.14% of the Twitter news search results are from their own Twitter

Query	Google TOB	Twitter news TOB
Queries related to events		
democratic debate	-0.039	0.271
dem debate	0.016	0.881
republican debate	-0.224	0.216
rep debate	0.073	0.07
Queries related to Democratic candidates		
Hillary Clinton	0.766	0.3
Bernie Sanders	0.577	0.42
Martin O'Malley	0.552	0.701
Average	0.631	0.473
Queries related to Republican candidates		
Donald Trump	-0.524	0.542
Ted Cruz	-0.543	0.288
Marco Rubio	-0.055	0.253
Ben Carson	-0.259	0.191
Chris Christie	-0.105	-0.286
Jeb Bush	-0.201	0.236
Rand Paul	-0.642	-0.006
Carly Fiorina	-0.487	0.09
John Kasich	-0.364	0.442
Mike Huckabee	0.006	0.058
Rick Santorum	-0.229	-0.041
Lindsey Graham	-0.183	-0.12
George Pataki	-0.259	0.125
Jim Gilmore	0.138	-0.608
Average	-0.264	0.083

Table 5.11: Comparing time-averaged output bias TOB in (i) Google search results, (ii) Twitter news search results.

account. However, there are a few exceptions like *Martin O'Malley*, *Chris Christie* and *Jim Gilmore*, for whom 16.46%, 14.62% and 19.65% respectively of their Twitter news search results come from their own Twitter accounts. And correspondingly, the search results for these candidates show a strong bias towards their own perspective (as shown in Table 5.11). But, for most other candidates, the fractions of such tweets is much lower, and the bias in the Twitter news search results towards their own perspective is also lower.

The above observations about web search, including lower dynamicity over time and the candidates having favorable biases due to controlling a significant fraction of the links which come up in their top search results, make it easier for candidates to manipulate the Web search results in their own favor. While, the results on Twitter are much more dynamic and affected more by popular users on Twitter, rather than the candidates themselves, making them much harder to manipulate.

5.5 Comparing Relative Bias of Twitter's Different Ranking Systems

In this work, we measure the output bias of two different ranking systems of Twitter search – 'top' and 'news' search filters – for the same set of queries. Since the input biases for the two are the same, we can compare the relative ranking biases for these two different ranking systems of Twitter. When we consider the average biases for the Republican candidates, we find that the input bias is slightly Republican-leaning (average TIB = 0.07, shown in Table 5.7), the Twitter 'top' search ranking system adds a Republican-leaning bias making the output bias Republican-leaning (average TOB = -0.11, shown in Table 5.7). While the Twitter 'news' search ranking system adds a little Democratic-leaning ranking bias making the output bias even more Democratic-leaning (average TOB = 0.083, shown in Table 5.11). This comparison of their relative ranking biases indicates that the 'news' filter of Twitter search highlights much more Democratic-leaning posts than the 'top' search filter.

5.6 Generalizability of the search bias quantification framework

Having presented our results from applying our search bias quantification framework to measure the bias in political searches on Twitter social media search and Google Web search in the context of US politics, we now present a brief discussion of how our bias quantification framework can be generalized to scenarios of multiple perspectives, limited search data, and other search systems.

5.6.1 Extending to multiple perspectives scenario

In this work, we have focused on US politics, and we have applied our bias quantification framework to this two-perspective scenario. However, it is possible to extend our framework to multiple perspective scenarios, for instance with p different perspectives. These p different perspectives could correspond to the bias towards p different socio-political issues, or they could correspond to p different political parties. Our framework can be extended to p different perspectives, by associating a p-dimensional bias vector with each item, rather than a scalar bias score, as we did currently. More formally, the bias vector for the i-th data item would be given by V_i = [v_i^1, v_i^2,..., v_i^p], where v_i^j gives a measure of how biased the i-th data item is along the j-th perspective, with values in the range of $[-1, 1]$. Here a value of $v_i^j = 1$ could indicate support for the j-th perspective, $v_i^j = -1$ could indicate opposition, whereas $v_i^j = 0$ could indicate that the item is neutral with respect to that perspective. By converting Equations 5.1 to 5.4, to their vector addition formulations, we can measure the input, output and ranking biases for this p-dimensional scenario. The primary challenge for pursuing this direction in the future is the development of a methodology to capture these bias vectors.

5.6.2 Extending to limited data availability scenario

In many (if not most) cases, it may not be possible or feasible to either access or collect the input dataset of all items containing the selected queries. In such scenarios, we can adopt one of the following two approaches for applying our quantification framework for estimating the search bias:

Compare relative biases of two different search systems that function on similar input data: For many modern IR systems, the items in the corpus are directly ranked according to their relevance for a query, without explicitly extracting an intermediate relevant item set. For such systems, we can compute the relative ranking biases of the two systems assuming them to operate upon similar input sets. For instance, we could compare the relative ranking of different web search engines (*e.g.,* Google vs. Bing vs. Yahoo), by observing the output bias for the same set of queries.

Approximate the input bias from the output search result snapshots: A simple approximation of the input bias based on the output search snapshot could be computed by taking an unweighted average of the bias scores of the items in the output set. This naive approximation can be improved by averaging over items in multiple search snapshots (*e.g.,* n search snapshots), or averaging over items in a larger snapshot with more search results (*e.g.,* top-$10k$ instead of top-k results).

5.6.3 Extending to other search systems

Our bias quantification framework follows a black box approach and does not require the knowledge of the internal details of retrieval and ranking systems to quantify the search bias. As a result, it can be easily applied to study the bias of a wide range of search systems, as long as a methodology for computing the bias of an individual item (*e.g.,* web-pages, tweets, posts) is available. Measuring the bias of an individual item in a search system is a context-dependent task, and since each platform is different, this in itself requires a significant effort. In this paper, we have delineated bias measurement techniques for tweets (Section 5.3) and web-links (Section 5.4). Also, in Section 5.1, we have briefly described prior work which has developed techniques for measuring the bias of users [189, 144, 70, 82, 54, 52, 180, 33, 249] or content [260, 242] on social media as well as blogs and news stories [15, 253, 264, 41, 161] on the Web. In the future, when bias quantification schemes are developed for other search systems, for instance for videos (*e.g.,* Youtube search) or music (*e.g.,* Spotify), these methodologies can be plugged into our bias quantification framework and be used to analyze the bias of these other search systems.

5.7 Signaling bias in search results

In this work, we have shown that both social media search, as well as Web search results, display varying degrees of bias. Next, we briefly discuss some solutions for tackling the bias, though their in-depth evaluation is left for the future.

Designing bias-aware ranking systems: A potential solution to address search bias is to design bias-aware ranking systems, which trade-off other metrics like relevance, popularity or recency with the bias of the search results. For instance, this could be achieved by minimizing the overall bias of search results by interleaving results with different biases using methods similar to the ones used for injecting diversity in results [245, 257]. However, this may lead to a degradation of the quality of search results along these relevance metrics, and finding an optimal trade-off point might be domain and user specific.

Making bias transparent in search interface design: An alternative method for addressing search bias could be to make the bias of each result transparent to the user by incorporating it into the search engine's front-end design. Such a nudging practice has been used widely in the literature for purposes like delivering multiple aspects of news in social media [178] and encouraging reading of diverse political opinions [166, 164]. In a recent field study, it has been shown that by showing users alerts about the ranking bias in the search results can suppress the impact of the ranking bias on undecided voters' voting preferences and also encourage them to read lower ranked results [65].

Hybrid approach - Split search: A hybrid approach of the above two methods could also be proposed, which not only shows the bias of each search result, but also separates the results from the two political perspectives (Republican and Democratic) and shows them as *distinct ranked lists*, with each distinct list retaining the ranking of the results in the original ranked list. This solution can be particularly effective in the cases where re-designing the algorithm and reaching a trade-off point for considering both bias and other relevance factors in an algorithm's design is infeasible. This method is similar to how several product companies like Amazon separate product reviews into positive and negative reviews, such that a user searching for that product can read the perspectives of others who either liked or disliked the product. By preserving the original search engine's ranking within each list, this methodology ensures that the quality of the top search results does not degrade across other metrics such as relevance, popularity, and recency.

We have deployed the proposed split search methodology as a live Twitter-based search service [4], at **https://twitter-app.mpi-sws.org/search-bias-split-view/** which allows users to log in with their Twitter credentials and do real-time searches for political queries on Twitter. The top search results are presented to the user as two distinct ranked lists containing Democratic- and Republican-leaning tweets, with each list maintaining the relevance rankings of the original search results returned by Twitter. Figure 5.6 shows a snapshot of the tool for the search term 'abortion'. Besides showing the bias of each search result, this split search design helps users to understand what fraction of the top results are related to each political leaning. For example, Figure 5.6 shows that there are more Democratic-leaning search results for the query 'abortion' than Republican-leaning ones amongst the first page of top search results. Such differences can nudge users to notice which is the dominant political leaning for the top search results for a search query and encourage them to read more results from the other political side to gain more balanced information about a topic. A similar system has been developed by Wall Street Journal [116] which presents posts from the most biased news publishers on Facebook as

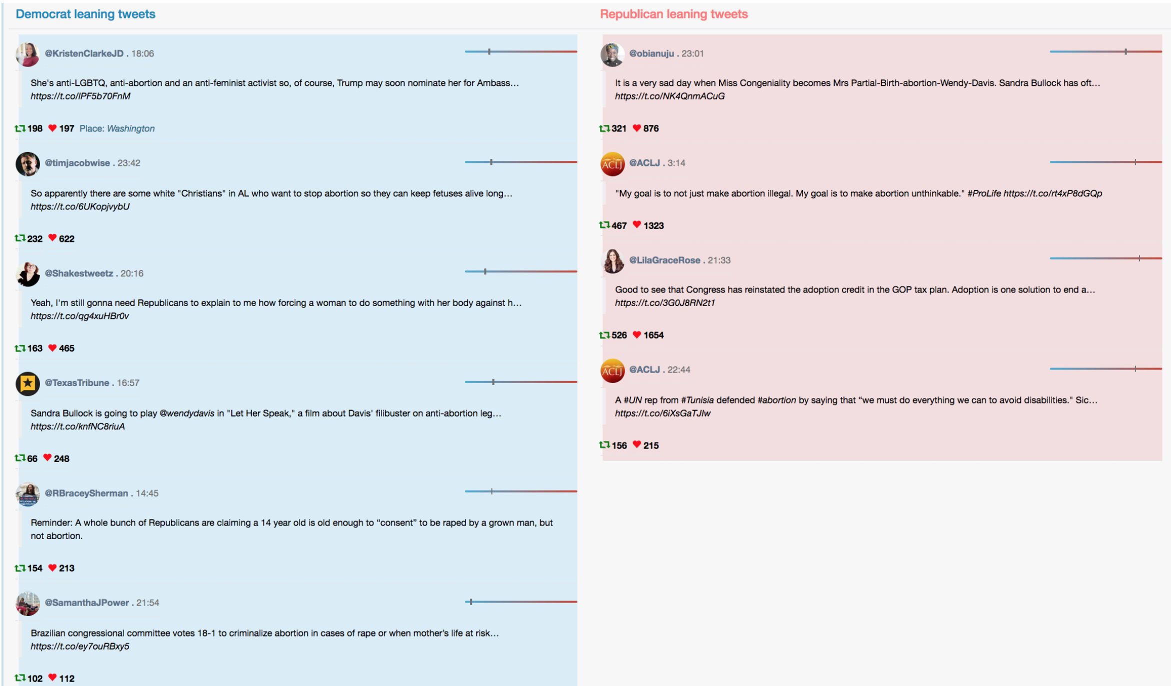


Figure 5.6: Screenshot of our Twitter-based split search service showing the results for the search term 'abortion'. A widget adjacent to each result shows its bias measure.

chronological lists, with the aim of showing both sides of the stories. However, how users interact with such alternative search interface designs remains to be investigated and is left for future work.

Further details about the functionalities of our Twitter-based split search service and a pointer to a demo video can be found in Appendix A.3.

5.8 Limitations & future directions

In this work, we focussed on a limited set of queries that were either related to a political event or a political candidate, due to our data collection limitations. Extending our query set to include more general political queries on polarizing topics like gun control or immigration could be done in the future to understand how the search systems are biasing the discourse about these popular debates in the society.

Another limiting factor in our study was using the simplifying assumption of considering a user as either neutral, pro-Democrat or pro-Republican. Under this assumption we can not have a user who is partially both pro-Republican and pro-Democrat. However, we should clarify that for doing this classification, we still considered two scores for each user, one which captures the similarity to

Republicans and the other to Democrats. Currently, to give the user a final leaning, we consider the difference between these similarities. However, in the future, we can use these two similarities to determine the extent to which a user is pro-Democrat as well as pro-Republican to have a more nuanced view of political leanings of users.

Also, the bipolar nature of US politics makes for a conducive environment for our bias measurement methodology. Extending our methodology for a multidimensional (political) space is likely to be quite challenging. Since our bias quantification framework can as easily work with a different methodology for inferring the bias of an individual item, future advances in measuring multidimensional bias could be plugged into our framework to quantify search bias for more nuanced and complex multidimensional bias search scenarios.

In this work, we have focussed our attention on non-personalized search results, by adopting measures to mitigate the personalization effects as described earlier in the paper. We do acknowledge that in reality, most searches made by users are personalized. Therefore our results may not be representative of the searches mostly done in the wild. However, we believe that the personalization is most likely to exacerbate the biases we observe and report in this thesis. In the future, our bias quantification framework can be applied to study bias in personalized search scenarios as well. By performing carefully controlled experiments [90, 119], along with our framework, the different sources of bias in personalized search scenarios can potentially be discerned. We leave the detailed design and implementation of such a study for the future.

And lastly, while we have discussed some potential solutions for signaling political bias in search results, and we have implemented our proposed split search as a Twitter application, however, we have not done a user study to investigate the effect of this signaling on the users' search experience. This exploration is an important follow up of our current work.

5.9 Conclusion

To our knowledge, this work presents the first search bias quantification framework which not only quantifies the bias in the output search results but also discerns the contributions of two sources of bias – input data and ranking system. We have applied our framework to investigate the sources of bias for political searches on Twitter social media and found both input data and the ranking system to be prominent contributors of the final bias seen by the users in the output ranked list of search results. We found that factors such as the topic of the query, the phrasing of query and the time at which a query is issued also impact the bias seen by the users. We also applied our framework to compare the relative biases of Google Web search and Twitter social media search and found that Web search results are typically more favorable for the candidates from the two parties because many of the top results include links to candidate-controlled sources like their own or their party's websites and social media accounts.

While we do measure and report the bias introduced by the ranking systems of Twitter and Google search engines, we do not claim that these biases are intentionally added by the platform. In fact, we did not find evidence of any systemic bias, *i.e.,* the platforms consistently ranking the items from one political leaning higher than the other, or consistently making the search results more polarizing by adding a Democratic-leaning bias to Democratic party related queries and Republican-leaning bias to Republican party related queries.

Our work lays the groundwork for the design of new mechanisms for making the users more aware of search bias, for instance by making the potential biases in the search results transparent to the users. For users, this awareness can lead to more intelligent use of the system to mitigate the effects of search bias. For system designers, the search bias framework can be used to audit their systems, especially in cases when the bias is introduced by the ranking system and not the input data. And lastly, researchers and watchdog organizations can utilize our framework to audit and compare the biases of different search platforms, especially to unearth cases where the search bias may be ending up misleading the users.

CHAPTER 6

Concluding discussion

With an increasing number of people around the world relying on online social media platforms like Twitter and Facebook to consume news and information about the world around them, there has a been a paradigm shift in the way news and information is exchanged in our society – from traditional mass media to online social media. This paradigm shift has lead to three fundamental changes in the way that people are exchanging information: (i) Unlike the subscribers of mass media, online social media users are not just passive consumers of information, but they are also active publishers of content on these platforms. (ii) Social media users curate personalized word of mouth channels for themselves by creating social links to their preferred sources, and therefore unlike broadcast mass media where all subscribers receive the same information, individual users on social media might receive information that is very different from what other users are consuming. (iii) Social media users often rely on automated retrieval algorithms like personalized recommendation and search systems deployed by the social media platform providers to discover interesting information from the deluge of user-generated content published and shared on these platforms.

The areas of journalism and media studies have traditionally focused mostly on broadcast mass media. With the changing environment, it is also essential to study the news and information consumption of social media users and to audit how the automated algorithms are modifying what the social media users consume. In this thesis, we fulfilled this high-level goal by following a two-fold approach. First, we proposed the concept of *information diets* – which is the composition of information being produced or consumed – to measure and reason about the bias and diversity in information production and consumption on social media platforms. We then *quantified the diversity and bias in the information diets* that social media users consume via the three primary consumption channels on social media platforms: (a) *word of mouth* channels that users select for themselves by creating social links, (b) *recommendations* that the social media platform providers give to the users, and (c) the *search* systems that users use to find interesting information on these platforms. We measured the information diets of social media users along the dimensions of *topics*, *geographic source diversity*, and *political perspectives*.

We began examining the impact of the paradigm shift by exploring the word of mouth diets that social media users are producing and consuming. First, we investigated the topical diversity in the word of mouth diets of Twitter users by proposing a novel and scalable author-based crowdsourced topic inference methodology for tweets. We observed that social media users (including news organizations) produce very focused diets that predominantly comprise of just one or two topics. Moreover, the consumed diets of these users are even more skewed towards just a couple of topics. Therefore, we observed that information production on social media has become unbundled with each source producing a very focused diet and that users do not select a topically balanced diet for themselves. We also developed a Twitter-app (deployed at **http://twitter-app.mpi-sws.org/information-diets/**), where users can log in with their Twitter credentials and view their own diets, as well as, examine the diets being produced by other users by searching for them. Then, we proceeded to examine the geographic source diversity in the diets of users on social media. We performed the analysis at the granularity of countries and found that a substantial majority of the information consumed is contributed by the country itself or a small number of geographically or linguistically close countries.

Having observed the unbundling of content production and the skewed content consumption of social media users, we proceeded to examine the impact of the algorithmic channel of recommendations on the topical diversity of the diets users consume. We utilized test accounts which mimic the social neighborhoods of real users to collect the social recommendations given to them by Twitter (which typically rely on user's neighborhood to recommend content to them). We surprisingly discovered that social recommendations somewhat mitigated the topical imbalances in the consumed diets of users by exposing them to a more heterogeneous set of topics.

Lastly, we directed our focus towards the algorithmic channel of search and proposed a search bias quantification framework to determine the impact of search on the information diets of users in the context of political searches about the US 2016 Presidential Primaries. To our knowledge, this is the first framework that not only quantifies the bias in the output search results but also distinguishes how much of the bias is due to the input data and how much is due to the ranking system. First, we applied the framework to study bias in search results on Twitter social media, for which we proposed a novel and scalable author-based methodology for inferring the political leaning of an individual tweet. We also developed a Twitter app which lets users log in with their Twitter credentials and view their inferred political leaning, as well as, search for other users and see their inferred political leanings [184] (deployed at **http://twitter-app.mpi-sws.org/search-political-bias-of-users/**). We observed that multiple factors, including the topic of the query, the phrasing of the query, and the time of querying, impact the bias seen by the users via the search on Twitter social media. We also applied our framework to compare the relative biases of Google Web search and Twitter social media search. We discovered that as compared to Twitter social media search results, the Web search results for the candidate name queries are typically more favorable to the candidates due to many web-links to candidate-controlled sources like the candidate's or their party's websites and social media accounts coming up in the top search results. However, we did not find any evidence of systemic bias, *i.e.,* the platforms consistently ranking the items with one political leaning higher than the other, or consistently making the results more polarizing.

We envisage that our work will not only create awareness among social media users about potential imbalances and biases in their information diets but will also lay the groundwork for the design of future information discovery, curation, and recommendation systems for social media. For users, this awareness can lead to a more intelligent and balanced use of the social media platforms for information consumption. For system designers, our bias and diversity measurement frameworks can be used to audit their algorithmic search and recommendation systems. And lastly, researchers and watchdog organizations can utilize our bias and diversity measurement frameworks to audit and compare different social media platforms and use them to discover scenarios where the imbalances in the users' diets may be misleading users and warn them.

6.1 Future research directions

Our work lays the foundation and takes the first steps towards studying information diets of social media users. However, there are still some open research questions that need to be addressed in the future. In this section, we briefly discuss some research directions that build upon our work and can be pursued in the future.

Developing normative guidelines for balanced information diets: In this thesis, we have presented information diets as a descriptive metric, and we do not address the question of what constitutes a "good" or "balanced" information diet. In the future, it is going to be increasingly important to follow a multi-disciplinary approach to addressing this question to provide some normative guidelines for balanced information diets. Another aspect which makes this question quite hard to answer is the fact that what constitutes a good diet may differ from person to person. Moreover, it is unclear who should prescribe this normative definition of a good diet, for individuals and the society at large. In the past, this role was played by public service media, with one of their main aims being ensuring diversity in news coverage and exposure. Therefore, there have been some suggestions [98, 100, 216] for technologically implementing the diversity principle for strengthening public service media, for instance by integrating "algorithmic diversity diet" [216] into the electronic program guides [98].

Characterizing information diets of social media users from different demographic groups: In this thesis, we focused on studying the information diets of average social media users. However, it is very likely that diets of users from different demographic groups may suffer from different biases. Thus, in the future, it would be interesting to compute and compare the diets of different demographic groups. The groups could be formed across different demographic dimensions, for instance, based on topics of interests of the users, based on the popularity of the users in terms of how many followers they have, or based on their age on the network. For instance, in Section 3.1.4, we have seen that users who are interested in different topics (interests identified as the topics which contribute most to their diets), have diets which are skewed to different extents and this could be explored more deeply in the future. Alternatively, in prior work [84], Grabowicz *et al*. examined how the characteristics of audiences and follower links change as a user gets more popular over time. We could similarly investigate how the bias and diversity in the information diets of users change as the user ages on the network, or as the user gets more popular on the network by doing a longitudinal analysis while controlling for other factors impacting the diets.

Conducting cross-platform measurement of information diets: In this thesis, we have predominantly focused on examining the *within-platform* diets of social media users. However, in today's vast and complex media ecosystem, most users rely on different platforms, both online and offline, to satisfy their diverse information needs. To get a more holistic view of users' information diets and to evaluate the full impact of algorithmic retrieval systems on what users consume, it will be imperative to extend the study of users' usage and consumption across different platforms in the future. Such *cross-platform* studies are likely to give a more comprehensive answer to questions like whether users are trapped in information bubbles or not.

Developing mechanisms to enable users to control their diets: Many of our findings in this thesis raise the need for better information curators (human editors or automated recommendation systems) on social media that provide a more balanced information diet. While we have developed a descriptive Twitter-app (**http://twitter-app.mpi-sws.org/information-diets/**), which makes the users aware of their information diets, there is further scope in the future to develop prescriptive mechanisms and tools to give the users better control on their diets. For instance, if a user observes their diet to be too biased towards a few topics or particular political ideology, they may wish for a more balanced diet. In such cases, a personalized recommendation system can be designed to supplement or balance their diet. While there has been a lot of prior work [161, 165, 163, 167, 168, 257, 50] on introducing diversity in recommendations, none of the earlier work, to our knowledge, has looked at the issue of providing personalized recommendations to individual users that balance the diets they are consuming.

Understanding the impact of information diets on user's opinion formation: Finally, much future work remains to be done in understanding the impact of the information diets – both word of mouth and algorithmic diets – on the shaping of the opinions of users. More specifically, it would be interesting to study if and how the personalization of diets reinforces biases, viewpoints, prejudices and political choices of the users. This study would require developing methods for investigating, detecting and predicting the effects of personalization.

Appendices

APPENDIX A

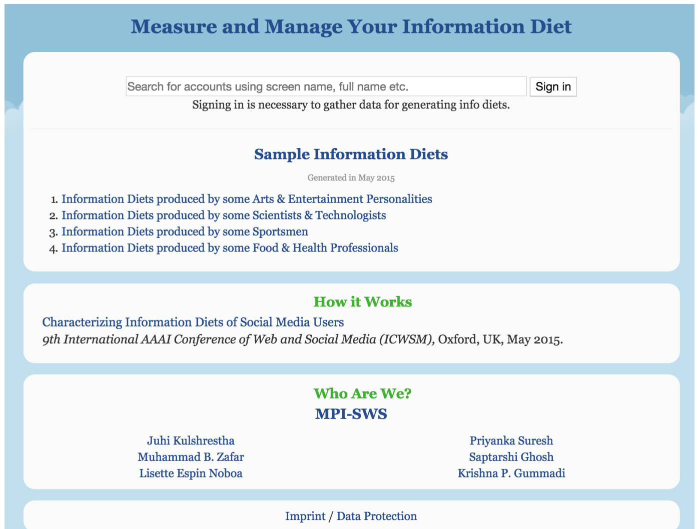
Screenshots & videos demonstrating the functionalities of publicly deployed Twitter applications

A.1 Making users aware of their information diets

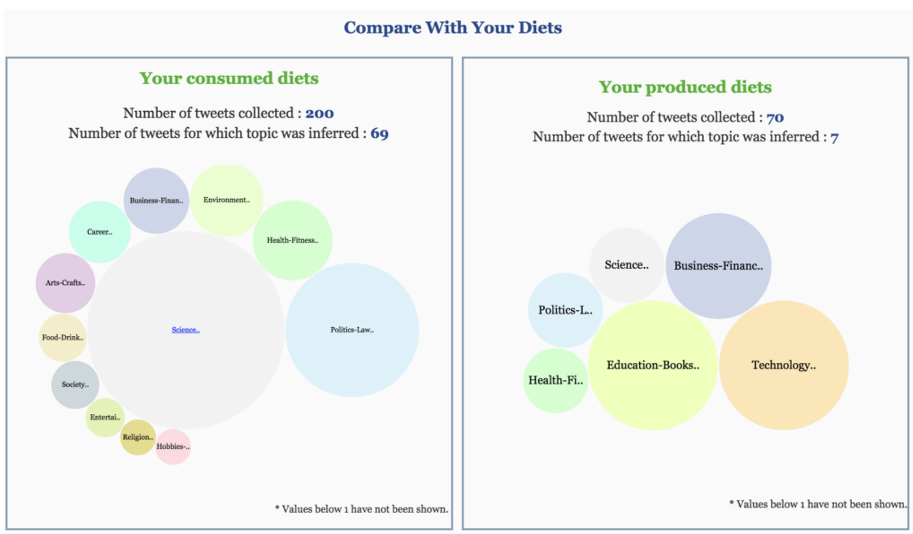
We have publicly deployed a Twitter application to make users more aware of the information diets they are producing and consuming on Twitter social media, at **http://twitter-app.mpi-sws.org/information-diets/**. Using this system, the users can explore the diets they are consuming on Twitter, as well as filter the tweets they are consuming by the topics and selectively read them. Additionally, they can also search for other Twitter users to examine the diets they are producing.

Figure A.1 depicts the functionality of our Twitter application as screenshots. In Figure A.1 (a), we show the log-in screen of our application where the users log in with their Twitter credentials. Upon login, the users are shown the diets they are consuming and producing on Twitter (shown in Figure A.1 (b)). The users can also search for other users (*e.g.,* '@fifaworldcup') and view the diets produced by them (shown in Figure A.1 (c)). And finally, the users can selectively read the produced or consumed tweets on different topics. For instance, in Figure A.1 (d), we show the tweets produced by '@fifaworldcup' on the topic of sports.

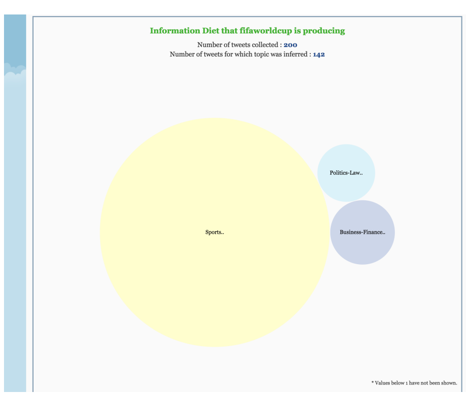
Demo video: We have uploaded a video demonstrating the functionalities of our application at **https://twitter-app.mpi-sws.org/whats-my-info-diet/demo_video.php**.



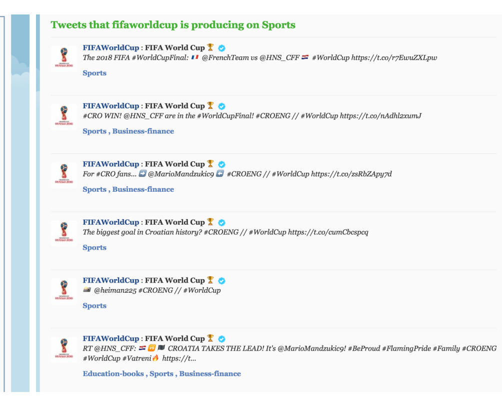
(a) **Log-in screen**



(b) **Diets consumed and produced by the logged-in user**



(c) **Diet produced by '@fifaworldcup'**



(d) **Tweets produced by '@fifaworldcup' on the topic of Sports**

Figure A.1: Screenshots of our Twitter application for making users aware of their information diets showing the (a) log-in screen, (b) diets consumed and produced by the logged-in user, (c) diet produced by searched account '@fifaworldcup', (d) Tweets produced by '@fifaworldcup' on the topic of Sports.

A.2 Inferring political leaning of Twitter users

We have publicly deployed the source bias inference methodology described in Section 5.3.2 in the form of a Twitter application [184], at **http://twitter-app.mpi-sws.org/search-political-bias-of-users/**. One can log in to the application using their Twitter credentials, and see their inferred political affiliation. One can also search for other Twitter users to check out their inferred political leaning.

Figure A.2 depicts the functionality of our Twitter application as screenshots. In Figure A.2 (a), we show the log-in screen of our application where the users log in with their Twitter credentials. Upon login, the users are shown their inferred political leaning (shown in Figure A.2 (b)). They can also search for other users to view their political leanings. Figure A.2 (c) and Figure A.2 (e) show the search results for the queries 'barack obama' and 'donald trump' respectively, while Figure A.2 (d) and Figure A.2 (f) show their inferred political leanings.

Demo video: We have uploaded a video demonstrating the functionalities of our application at **https://twitter-app.mpi-sws.org/search-political-bias-of-users/demo_video.php**.

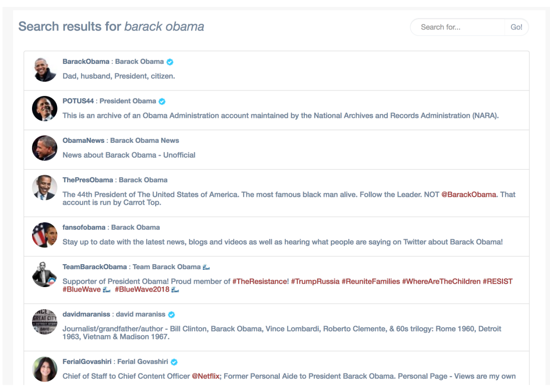
A.3 Twitter search split by the political leaning

We have deployed the split search methodology proposed in Section 5.7 as a live Twitter-based search service [4], at **https://twitter-app.mpi-sws.org/search-bias-split-view/** which allows users to log in with their Twitter credentials and do real-time searches for political queries on Twitter. The top search results are presented to the user as two distinct ranked lists containing Democratic- and Republican-leaning tweets, with each list maintaining the relevance rankings of the original search results returned by Twitter.

Figure A.3 depicts the functionality of our Twitter application as screenshots. In Figure A.2 (a), we show the log-in screen of our application where the users can log in with their Twitter credentials and search for political queries on Twitter. For instance, in Figure A.2 (b) and Figure A.2 (d), show the user searching for the query terms 'guns' and 'abortion', while Figure A.2 (c) and Figure A.2 (e) show the search results split by the political leaning. A widget adjacent to each result shows its political bias measure.

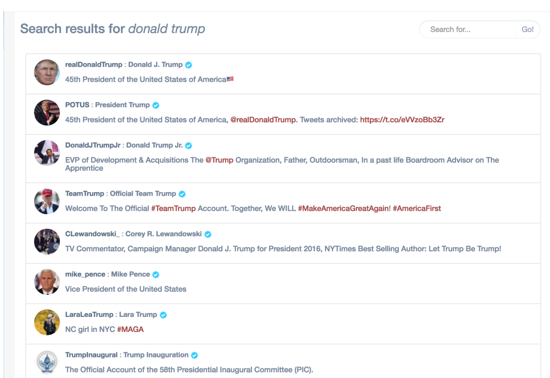
(a) **Log-in screen**

(b) **Inferred political leaning of logged-in user**

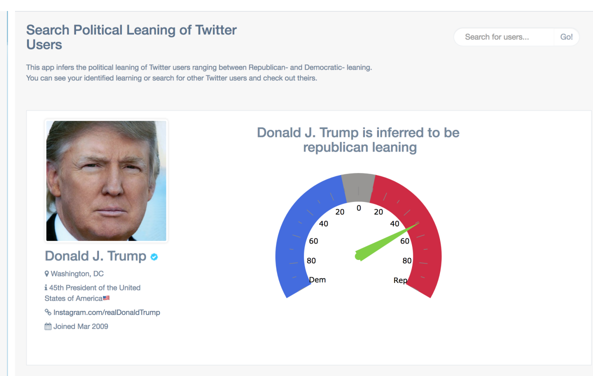


(c) **Search results for 'barack obama'**

(d) **Inferred political leaning for @BarackObama**



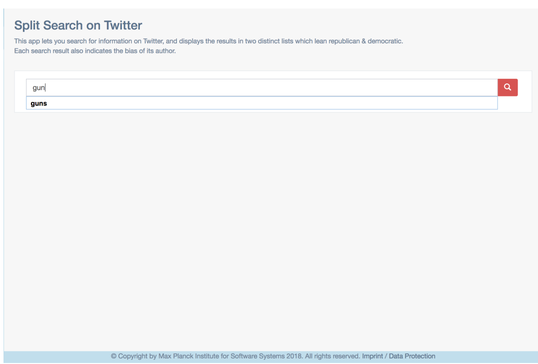
(e) **Search results for 'donald trump'**



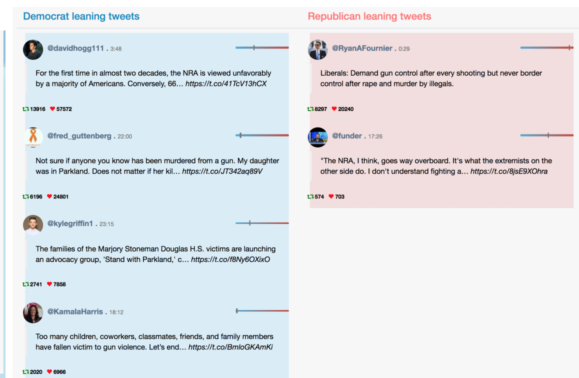
(f) **Inferred political leaning for @realDonaldTrump**

Figure A.2: Screenshots of our Twitter application for inferring political leaning of users showing the (a) log-in screen, (b) inferred political leaning of logged-in user, (c) search results for 'barack obama' , (d) inferred political leaning for @BarackObama, (e) search results for 'donald trump', (f) inferred political leaning for @realDonaldTrump.

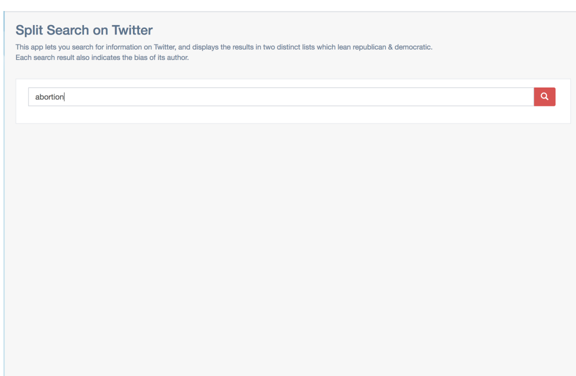
(a) **Log-in screen**



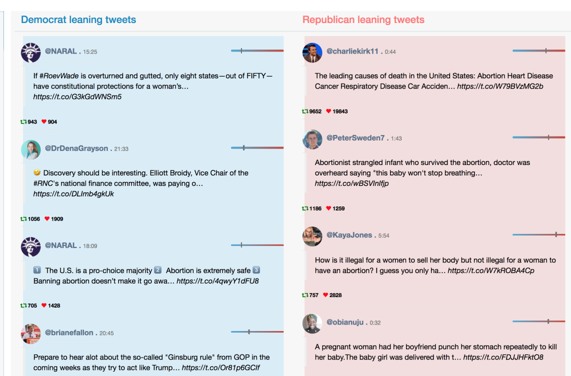
(b) **Search for 'guns'**



(c) **Split search results for 'guns'**



(d) **Search for 'abortion'**



(e) **Search results for 'abortion'**

Figure A.3: Screenshot of our Twitter-based split search service showing the showing the (a) log-in screen, (b) search screen for 'guns', (c) search results for 'guns' , (d) search screen for 'abortion', (e) search results for 'abortion'. A widget adjacent to each result shows its political bias measure.

Demo video: We have uploaded a video demonstrating the functionalities of our application at **https://twitter-app.mpi-sws.org/search-bias-split-view/demo_video.php**.

BIBLIOGRAPHY

[1] Intro to Facebook Search. *https://code.facebook.com/posts/422334894553532/intro-to-facebook-search/*.

[2] List of countries and dependencies by population. **https://en.wikipedia.org/wiki/List_of_countries_and_dependencies_by_population**.

[3] Media Bias / Fact Check - The most comprehensive media bias resource. *https://mediabiasfactcheck.com*.

[4] Twitter Application: Make Users Aware of the Biases in Search. *http://twitter-app.mpi-sws.org/search-bias-split-view/*.

[5] Twitter Blog: Building a complete Tweet index. *https://blog.twitter.com/engineering/en_us/a/2014/building-a-complete-tweet-index.html*.

[6] Twitter Blog: Search Relevance Infrastructure at Twitter. *https://blog.twitter.com/engineering/en_us/topics/infrastructure/2016/search-relevance-infrastructure-at-twitter.html*.

[7] Twitter Blogs | Discover: Improved personalization algorithms and real-time indexing. *http://tinyurl.com/twitter-discover-tab*.

[8] Twitter Help Center: Search result FAQs. *https://help.twitter.com/en/using-twitter/top-search-results-faqs*.

[9] Twitter Search Home. *https://twitter.com/search-home*.

[10] Where do People You May Know suggestions come from? *https://www.facebook.com/help/163810437015615?helpref=faq_content*.

[11] Wikipedia : Neutral point of view. *https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view*.

[12] *European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14 (ETS 5)*. Council of Europe, 1950.

[13] A. I. ABRAMOWITZ. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. Yale University Press, 2010.

[14] L. A. Adamic and N. Glance. The political blogosphere and the 2004 us election: Divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, LinkKDD '05, pages 36–43, New York, NY, USA, 2005. ACM.

[15] L. A. Adamic and N. Glance. The political blogosphere and the 2004 u.s. election: Divided they blog. In *In Proceedings of LinkKDD*, 2005.

[16] E. Agapie and S. A. Munson. Social cues and interest in reading political news stories. In *Ninth International AAAI Conference on Web and Social Media*, 2015.

[17] AMT. **https://www.mturk.com/**.

[18] Get better results with less effort with Mechanical Turk Masters – The Mechanical Turk blog. **http://bit.ly/amt-masters**.

[19] G. Anthes. Search engine agendas. *Communications of the ACM*, 59(4):19–21, 2016.

[20] M. Babaei, J. Kulshrestha, A. Chakraborty, F. Benevenuto, K. P. Gummadi, and A. Weller. Purple feed: Identifying high consensus news posts on social media. 2018.

[21] L. Backstrom, E. Sun, and C. Marlow. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th int'l conference on World wide web*, 2010.

[22] C. A. Bail, L. Argyle, T. Brown, J. Bumpuss, H. Chen, M. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky. Exposure to opposing views can increase political polarization: Evidence from a large-scale field experiment on social media, Mar 2018.

[23] E. Bakshy, S. Messing, and L. Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 2015.

[24] P. Barberá. A new measure of party identification in twitter. evidence from spain. In *2nd Annual General Conference of the European Political Science Association (EPSA), Berlin, June. EPSA*, 2012.

[25] P. Barberá. Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis*, 23(1):76–91, 2014.

[26] S. Barocas and A. D. Selbst. Big data's disparate impact. *Available at SSRN 2477899*, 2014.

[27] D. P. Baron. Persistent media bias. *Journal of Public Economics*, 90(1-2):1–36, 2006.

[28] M. S. Bernstein, B. Suh, L. Hong, J. Chen, S. Kairam, and E. H. Chi. Eddi: Interactive topic-based browsing of social status streams. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 303–312, New York, NY, USA, 2010. ACM.

[29] P. Bhattacharya, S. Ghosh, J. Kulshrestha, M. Mondal, M. B. Zafar, N. Ganguly, and K. P. Gummadi. Deep twitter diving: Exploring topical groups in microblogs at scale. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, pages 197–210, New York, NY, USA, 2014. ACM.

[30] P. Bhattacharya, M. B. Zafar, N. Ganguly, S. Ghosh, and K. P. Gummadi. Inferring user interests in the twitter social network. In *Proceedings of the 8th ACM Conference on Recommender Systems*, RecSys '14, pages 357–360, New York, NY, USA, 2014. ACM.

[31] B. Bishop and R. G. Cushing. *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart*. Tantor Audio, 2017.

[32] V. Blondel, J. Guillaume, R. Lambiotte, and E. Mech. Fast unfolding of communities in large networks. *J. Stat. Mech*, page P10008, 2008.

[33] R. Bond and S. Messing. Quantifying social media's political space: Estimating ideology from publicly revealed preferences on facebook. *American Political Science Review*, 109(01):62–78, 2015.

[34] I. Bordino, Y. Mejova, and M. Lalmas. Penguins in sweaters, or serendipitous entity search on user-generated content. In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, CIKM '13, pages 109–118, New York, NY, USA, 2013. ACM.

[35] J. Borge-Holthoefer, W. Magdy, K. Darwish, and I. Weber. Content and network dynamics behind egyptian political polarization on twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, pages 700–711, New York, NY, USA, 2015. ACM.

[36] E. Borra and I. Weber. Political insights: Exploring partisanship in web search queries. *First Monday*, 17(7), 2012.

[37] A. Boutet, H. Kim, and E. Yoneki. What's in your tweets? I know who you supported in the UK 2010 general election. In *Proceedings of the Sixth International Conference on Weblogs and Social Media, Dublin, Ireland, June 4-7, 2012*, 2012.

[38] E. Bozdag. Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3):209–227, Sept. 2013.

[39] J. Breeman, V. Breeman, and N. Helberger. On the regulator's plate: Exposure diversity in a changing media environment: Workshop report and highlights of an expert discussion. *Journal of Information Policy*, 1:370–377, 2011.

[40] N. Bremmen. The 100 most influential news media Twitter accounts. *https://memeburn.com/2010/09/the-100-most-influential-news-media-twitter-accounts/*.

[41] C. Budak, S. Goel, and J. M. Rao. Fair and balanced? quantifying media bias through crowd-sourced content analysis. *Public Opinion Quarterly*, 80(S1):250–271, 2016.

[42] D. R. Carney, J. T. Jost, S. D. Gosling, and J. Potter. The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, 29(6):807–840, 2008.

[43] O. Celma and P. Herrera. A new approach to evaluating novel recommendations. In *Proceedings of the 2008 ACM conference on Recommender systems*, pages 179–186, 2008.

[44] M. Cha, F. Benevenuto, H. Haddadi, and K. P. Gummadi. The world of connections and information flow in twitter. *IEEE Trans. Systems, Man, and Cybernetics, Part A*, 42:991–998, 2012.

[45] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi. Measuring user influence in twitter: The million follower fallacy. In *Proceedings of International AAAI Conference on Weblogs and Social (ICWSM '10)*, 2010.

[46] J. Chen, R. Nairn, L. Nelson, M. Bernstein, and E. Chi. Short and tweet: experiments on recommending content from information streams. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 1185–1194, 2010.

[47] L. Chen, A. Mislove, and C. Wilson. Peeking beneath the hood of uber. In *Proceedings of the 2015 ACM Conference on Internet Measurement Conference*, pages 495–508. ACM, 2015.

[48] Z. Cheng, J. Caverlee, and K. Lee. You are where you tweet : A content-based approach to geo-locating twitter users. In *Proceedings of the 19th ACM int'l conference on Information and knowledge management*, 2010.

[49] C.-F. Chiang and B. Knight. Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies*, 78(3):795–820, 2011.

[50] M. D. Choudhury, S. Counts, and M. Czerwinski. Find Me the Right Content! Diversity-Based Sampling of Social Media Spaces for Topic-Centric Search. In *ICWSM*, 2011.

[51] R. Cohen and D. Ruths. Classifying political orientation on twitter: It's not easy! 2013.

[52] M. Conover, B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer. Predicting the political alignment of twitter users. In *In Proceedings of IEEE SocialCom*, 2011.

[53] M. Conover, J. Ratkiewicz, M. Francisco, B. Goncalves, F. Menczer, and A. Flammini. Political polarization on twitter. In *International AAAI Conference on Web and Social Media*, 2011.

[54] M. Conover, J. Ratkiewicz, M. Francisco, B. Gonçalves, F. Menczer, and A. Flammini. Political polarization on twitter. In *In Proceedings of AAAI ICWSM*, 2011.

[55] T. J. A. Covert and P. C. Wasburn. Measuring media bias: A content analysis of time and newsweek coverage of domestic social issues, 1975–2000. *Social science quarterly*, 88(3), 2007.

[56] E. Daehnhardt, Y. Jing, and N. Taylor. Cultural and geolocation aspects of communication in twitter. In *ASE Internation Conference on Social Informatics*, volume 2014, 2014.

[57] A. Datta, M. C. Tschantz, and A. Datta. Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *CoRR*, abs/1408.6491, 2014.

[58] Delia Baldassarri and Andrew Gelman. Partisans without constraint: Political polarization and trends in american public opinion. *American Journal of Sociology*, 114(2):408–446, 2008.

[59] D. DellaPosta, Y. Shi, and M. Macy. Why do liberals drink lattes? *American Journal of Sociology*, 120(5):1473–1511, 2015.

[60] P. DiMaggio, J. Evans, and B. Bryson. Have american's social attitudes become more polarized? *American Journal of Sociology*, 102(3):690–755, 1996.

[61] A. Dimitrov, A. Olteanu, L. McDowell, and K. Aberer. Topick: Accurate topic distillation for user streams. In *Data Mining Workshops (ICDMW), 2012 IEEE 12th International Conference on Data Mining Workshops*, pages 882–885. IEEE, 12 2012.

[62] B. Edelman. Hard-Coding Bias in Google "Algorithmic" Search Results. *http://www.benedelman.org/hardcoding/*, November 2010.

[63] R. Epstein and R. E. Robertson. The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *In Proceedings of of the National Academy of Sciences (PNAS)*, 112(33):E4512–E4521, 2015.

[64] R. Epstein, R. E. Robertson, D. Lazer, and C. Wilson. Suppressing the search engine manipulation effect (seme). *Proceedings ACM Hum.-Comput. Interact.*, 1(CSCW):42:1–42:22, Dec. 2017.

[65] R. Epstein, R. E. Robertson, D. Lazer, and C. Wilson. Suppressing the Search Engine Manipulation Effect (SEME). *Proceedings of the ACM : Human-Computer Interaction*, 1(2), November 2017.

[66] S. Eskens, N. Helberger, and J. Moeller. Challenged by news personalisation: five perspectives on the right to receive information. *Journal of Media Law*, 9(2):259–284, 2017.

[67] M. Eslami, K. Karahalios, C. Sandvig, K. Vaccaro, A. Rickman, K. Hamilton, and A. Kirlik. First i "like" it, then i hide it: Folk theories of social feeds. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2371–2382. ACM, 2016.

[68] M. Eslami, A. Rickman, K. Vaccaro, A. Aleyasen, A. Vuong, K. Karahalios, K. Hamilton, and C. Sandvig. I always assumed that i wasn't really that close to [her]: Reasoning about invisible algorithms in news feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 153–162. ACM, 2015.

[69] U. Executive Office of the President. Big Data: A Report on Algorithmic Systems, Opportunity,and Civil Rights. *http://tinyurl.com/Big-Data-White-House*, 2016.

[70] A. Fang, I. Ounis, P. Habel, C. Macdonald, and N. Limsopatham. Topic-centric classification of twitter user's political orientation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '15, pages 791–794, 2015.

[71] D. A. Ferguson and E. M. Perse. Media and audience influences on channel repertoire. *Journal of Broadcasting & Electronic Media*, 37(1):31–47, 1993.

[72] D. A. Ferguson and E. M. Perse. The world wide web as a functional alternative to television. *Journal of Broadcasting & Electronic Media*, 44(2):155–174, 2000.

[73] P. Ferragina, F. Piccinno, and R. Santoro. On analyzing hashtags in twitter, 2015.

[74] M. P. Fiorina, S. J. Abrams, and J. Pope. *Culture war?: the myth of a polarized America*. Boston, MA: Longman, 2011.

[75] S. Flaxman, S. Goel, and J. M. Rao. Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80(S1):298–320, 2016.

[76] S. Fortunato, A. Flammini, F. Menczer, and A. Vespignani. Topical interests and the mitigation of search engine bias. *In Proceedings of of the National Academy of Sciences (PNAS)*, 103(34):12684–12689, 2006.

[77] V. R. K. Garimella and I. Weber. Co-following on twitter. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, HT '14, pages 249–254, New York, NY, USA, 2014. ACM.

[78] M. Ge, C. Delgado-Battenfeld, and D. Jannach. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In *Proceedings of the fourth ACM conference on Recommender systems*, pages 257–260, 2010.

[79] M. Gentzkow and J. Shapiro. What drives media slant? evidence from u.s. daily newspapers. *Econometrica*, 78(1):35–71, 2010.

[80] S. Ghosh, N. Sharma, F. Benevenuto, N. Ganguly, and K. Gummadi. Cognos: Crowdsourcing search for topic experts in microblogs. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '12, pages 575–590, New York, NY, USA, 2012. ACM.

[81] J. Golbeck and D. Hansen. Computing political preference among twitter followers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1105–1108, New York, NY, USA, 2011. ACM.

[82] J. Golbeck and D. Hansen. Computing Political Preference Among Twitter Followers. In *ACM SIGCHI*, 2011.

[83] Official Google blog: New ways to stay informed about presidential politics. **http://tinyurl.com/OfficialGoogleBlog**.

[84] P. Grabowicz, M. Babaei, J. Kulshrestha, and I. Weber. The road to popularity: The dilution of growing audience on twitter. In *International AAAI Conference on Web and Social Media*, ICWSM '16, May 2016.

[85] L. A. Granka. The politics of search: A decade retrospective. *The Information Society*, 26(5):364–374, 2010.

[86] T. Groseclose and J. Milyo. A measure of media bias. *The Quarterly Journal of Economics*, 120(4):1191–1237, 2005.

[87] P. Gupta, A. Goel, J. Lin, A. Sharma, D. Wang, and R. Zadeh. Wtf: The who to follow service at twitter. In *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13, pages 505–514, New York, NY, USA, 2013. ACM.

[88] P. Gupta, V. Satuluri, A. Grewal, S. Gurumurthy, V. Zhabiuk, Q. Li, and J. Lin. Real-time twitter recommendation: Online motif detection in large dynamic graphs. volume 7, pages 1379–1380. VLDB Endowment, Aug. 2014.

[89] A. T. Hadgu, K. Garimella, and I. Weber. Political hashtag hijacking in the u.s. In *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13 Companion, pages 55–56, New York, NY, USA, 2013. ACM.

[90] A. Hannak, P. Sapiezynski, A. Molavi Kakhki, B. Krishnamurthy, D. Lazer, A. Mislove, and C. Wilson. Measuring personalization of web search. In *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13, pages 527–538, New York, NY, USA, 2013. ACM.

[91] A. Hannak, G. Soeller, D. Lazer, A. Mislove, and C. Wilson. Measuring price discrimination and steering on e-commerce web sites. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, IMC '14, pages 305–318, New York, NY, USA, 2014. ACM.

[92] U. Hasebrink and J. Popp. Media repertoires as a result of selective media use. a conceptual approach to the analysis of patterns of exposure. *Communications*, 31(3):369–387, 2006.

[93] B. Hecht and D. Gergle. The tower of babel meets web 2.0: user-generated content and its applications in a multilingual context. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 291–300. ACM, 2010.

[94] B. Hecht, L. Hong, B. Suh, and E. H. Chi. Tweets from justin bieber's heart: the dynamics of the location field in user profiles. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 237–246. ACM, 2011.

[95] C. Heeter. Program selection with abundance of choice. *Human Communication Research*, 12(1):126–152, 1985.

[96] C. Heeter and B. S. Greenberg. *Cableviewing*. Ablex Publishing Corporation, 1988.

[97] N. Helberger. Diversity by design. *Journal of Information Policy*, 1:441–469, 2011.

[98] N. Helberger. Diversity by design. *Journal of Information Policy*, 1:441–469, 2011.

[99] N. Helberger. Exposure diversity as a policy goal. *Journal of Media Law*, 4(1):65–92, 2012.

[100] N. Helberger. Public service medial merely facilitating or actively stimulating diverse media choices? public service media at the crossroad. *International Journal of Communication*, 9:17, 2015.

[101] J. L. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *ACM CSCW*, pages 241–250, 2000.

[102] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1):5–53, Jan 2004.

[103] I. Himelboim, S. McCreery, and M. Smith. Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on twitter. *Journal of Computer-Mediated Communication*, 18(2):40–60, 2013.

[104] L. Hong, G. Convertino, and E. H. Chi. Language matters in twitter: A large scale study. In *In Proceedings of AAAI International Conference on Weblogs and Social Media (ICWSM '10)*. The AAAI Press, 2010.

[105] L. Hong and B. D. Davison. Empirical study of topic modeling in twitter. In *Proceedings of the First Workshop on Social Media Analytics*, SOMA '10, pages 80–88, New York, NY, USA, 2010. ACM.

[106] Information Diets of Twitter Users. **http://twitter-app.mpi-sws.org/information-diets/**, 2015.

[107] L. D. Introna and H. Nissenbaum. Shaping the web: Why the politics of search engines matters. *The information society*, 16(3):169–185, 2000.

[108] F. J. Zuiderveen Borgesius, D. Trilling, J. Moeller, B. Bodo, C. de Vreese, N. Helberger, and I. Policy Review. Should we worry about filter bubbles? Volume 5, 03 2016.

[109] A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, WebKDD/SNA-KDD '07, pages 56–65, New York, NY, USA, 2007. ACM.

[110] C. A. Johnson. *The Information Diet: A Case for Conscious Consumption*. O'Reilly Media, Inc., 2012.

[111] S. N. Jomini. Polarization and partisan selective exposure. *Journal of Communication*, 60(3):556–576, 2010.

[112] P. R. C. Journalism and Media. Americans' online news use is closing in on TV news use. **http://www.pewresearch.org/fact-tank/2017/09/07/americans-online-news-use-vs-tv-news-use/**, 7 2017.

[113] P. R. C. Journalism and Media. News Use Across Social Media Platforms 2017. **http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/**, 9 2017.

[114] P. R. C. Journalism and Media. Newspapers Fact Sheet. **http://www.journalism.org/fact-sheet/newspapers/**, 6 2018.

[115] M. Kay, C. Matuszek, and S. A. Munson. Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3819–3828. ACM, 2015.

[116] J. Keegan. Blue feed, red feed - see liberal facebook and conservative facebook, side by side. http://graphics.wsj.com/blue-feed-red-feed/, 2017.

[117] S. Kim, I. Weber, L. Wei, and A. Oh. Sociolinguistic analysis of twitter in multilingual societies. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, HT '14, pages 243–248, New York, NY, USA, 2014. ACM.

[118] S. Kim, I. Weber, L. Wei, and A. Oh. Sociolinguistic analysis of twitter in multilingual societies. In *Proceedings of the 25th ACM conference on Hypertext and social media*, pages 243–248. ACM, 2014.

[119] C. Kliman-Silver, A. Hannak, D. Lazer, C. Wilson, and A. Mislove. Location, location, location: The impact of geolocation on web search personalization. In *Proceedings of the 2015 Internet Measurement Conference*, IMC '15, pages 121–127, New York, NY, USA, 2015. ACM.

[120] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, H. Soncu, and C. Newell. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4-5):441–504, Oct. 2012.

[121] D. Kotkov, S. Wang, and J. Veijalainen. A survey of serendipity in recommender systems. *Knowledge-Based Systems*, 111(C):180–192, Nov. 2016.

[122] D. Koutra, P. N. Bennett, and E. Horvitz. Events and controversies: Influences of a shocking news event on information seeking. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, pages 614–624, Republic and Canton of Geneva, Switzerland, 2015. International World Wide Web Conferences Steering Committee.

[123] B. Krishnamurthy, P. Gill, and M. Arlitt. A few chirps about twitter. In *Proceedings of the First Workshop on Online Social Networks*, WOSN '08, pages 19–24, New York, NY, USA, 2008. ACM.

[124] J. Kulshrestha, M. Eslami, J. Messias, M. B. Zafar, S. Ghosh, K. P. Gummadi, and K. Karahalios. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 417–432, New York, NY, USA, 2017. ACM.

[125] J. Kulshrestha, F. Kooti, A. Nikravesh, and K. P. Gummadi. Geographic dissection of the twitter network. In *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media*, ICWSM '12, June 2012.

[126] J. Kulshrestha, M. B. Zafar, S. Ghosh, and K. P. Gummadi. Evaluating social recommendation systems: A case-study on twitter's personalised suggestions. In *In poster session at ACM Conference on Online Social Networks (COSN)*, October 2014.

[127] J. Kulshrestha, M. B. Zafar, L. E. Noboa, K. P. Gummadi, and S. Ghosh. Characterizing information diets of social media users. In *Proceedings of the 9th International AAAI Conference of Web and Social Media*, ICWSM '15, May 2015.

[128] S. Kumar, R. West, and J. Leskovec. Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of the 25th International Conference on World*

Wide Web, WWW '16, pages 591–602, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee.

[129] M. Kunaver and T. Porl. Diversity in recommender systems a survey. *Knowledge-Based Systems*, 123(C):154–162, May 2017.

[130] L. R. Kurtz. Culture wars: The struggle to define america. james davison hunter. *American Journal of Sociology*, 99(4):1125–1128, 1994.

[131] H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web (WWW '10)*, 2010.

[132] P. Lahoti, K. Garimella, and A. Gionis. Joint non-negative matrix factorization for learning ideological leaning on twitter. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 351–359. ACM, 2018.

[133] F. Lardinois. Mashable: A brief guide to using Twitter's 'Discover' tab. **https://techcrunch.com/2012/05/01/twitter-improves-discover-tab-to-surface-more-interesting-content-promises-to-make-it-magical/**, May 2012.

[134] H. T. Le, Z. Shafiq, and P. Srinivasan. Scalable news slant measurement using twitter. In *Proceedings of the Intl. AAAI Conference on Web and Social Media (ICWSM)*, 2017.

[135] Q. V. Liao, W.-T. Fu, and M. Strohmaier. # snowden: Understanding biases introduced by behavioral differences of opinion groups on social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3352–3363. ACM, 2016.

[136] J. Lichterman. New Pew data: More Americans are getting news on Facebook and Twitter. **http://www.niemanlab.org/2015/07/new-pew-data-more-americans-are-getting-news-on-facebook-and-twitter/**, 2010.

[137] M. D. Lieberman and J. J. Lin. You are where you edit: Locating wikipedia contributors through edit histories. In *ICWSM*, 2009.

[138] J. Lin, R. Snow, and W. Morgan. Smoothing techniques for adaptive online language models: Topic tracking in tweet streams. In *Proceedings of the 17th ACM SIGKDD International*

Conference on Knowledge Discovery and Data Mining, KDD '11, pages 422–429, New York, NY, USA, 2011. ACM.

[139] L. Liu, N. Mehandjiev, and D.-L. Xu. Multi-criteria service recommendation based on user criteria preferences. In *Proceedings of the fifth ACM conference on Recommender systems*, pages 77–84, 2011.

[140] Z. Liu and I. Weber. Is twitter a public sphere for online conflicts? a cross-ideological and cross-hierarchical look. In *International Conference on Social Informatics*, pages 336–347. Springer, 2014.

[141] C. G. Lord and L. Ross. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, pages 2098–2109, 1979.

[142] H. Lu, J. Caverlee, and W. Niu. Biaswatch: A lightweight system for discovering and tracking topic-sensitive opinion bias in social media. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, 2015.

[143] J. Mahmud, J. Nichols, and C. Drews. Home location identification of twitter users. *ACM Transactions Intelligent Systems Technology*, 5(3):47:1–47:21, July 2014.

[144] A. Makazhanov and D. Rafiei. Predicting political preference of twitter users. In *In Proceedings of ASONAM*, 2013.

[145] C. D. Manning, P. Raghavan, and H. Schutze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.

[146] News Media Bias Monitor Discover the Demographics of News and Media Outlets. **https://twitter-app.mpi-sws.org/media-bias-monitor/**, 2018.

[147] E. Meij, W. Weerkamp, and M. de Rijke. Adding semantics to microblog posts. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, WSDM '12, pages 563–572, New York, NY, USA, 2012. ACM.

[148] G. Meikle and S. Young. *Media convergence: Networked digital media in everyday life*. Palgrave Macmillan, 2011.

[149] P. Metaxas and Y. Pruksachatkun. Manipulation of Search Engine Results during the 2016 US Congressional Elections. In *Proceedings of of the ICIW 2017*, Venice, Italy, 2017.

[150] M. Michelson and S. A. Macskassy. Discovering users' topics of interest on twitter: A first look. In *Proceedings of the Fourth Workshop on Analytics for Noisy Unstructured Text Data*, AND '10, pages 73–80, New York, NY, USA, 2010. ACM.

[151] A. G. Miller, J. W. McHoskey, C. M. Bane, and T. G. Dowd. The attitude polarization phenomenon: Role of response measure, attitude extremity, and behavioral consequences of reported attitude change. *Journal of Personality and Social Psychology*, 64(4):561, 1993.

[152] A. Mitchell. Key findings on the traits and habits of the modern news consumer. *Pew Research Center*, July 2016.

[153] A. Mitchell, J. Gottfried, J. Kiley, and K. Matsa. Political polarization and media habits. *Pew Research Center*, 21, 2014.

[154] D. Mocanu, A. Baronchelli, N. Perra, B. Gonçalves, Q. Zhang, and A. Vespignani. The twitter of babel: Mapping world languages through microblogging platforms. *PloS one*, 8(4):e61981, 2013.

[155] J. Moeller, D. Trilling, N. Helberger, K. Irion, and C. D. Vreese. Shrinking core? exploring the differential agenda setting power of traditional and personalized news media. *info*, 18(6):26–41, 2016.

[156] J. Moeller, D. Trilling, N. Helberger, and B. van Es. Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity. *Information, Communication & Society*, 21(7):959–977, 2018.

[157] F. Morstatter, J. Pfeffer, H. Liu, and K. Carley. Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. In *International AAAI Conference on Web and Social Media*, 2013.

[158] A. Mowshowitz and A. Kawaguchi. Assessing bias in search engines. *Information Processing & Management*, 38(1):141–156, 2002.

[159] A. Mowshowitz and A. Kawaguchi. Measuring search engine bias. *Information Processing and Management*, 41(5):1193–1205, 2005.

[160] G. D. Munro and P. H. Ditto. Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin*, 23(6):636–653, 1997.

[161] S. Munson, S. Chhabra, and P. Resnick. BALANCE - Tools for improving your news reading experience. *http://balancestudy.org/*.

[162] S. Munson, S. Chhabra, and P. Resnick. BALANCE - Tools for improving your news reading experience - List of Classified sources. *http://balancestudy.org/whitelist-classifiable.html*.

[163] S. Munson, S. Lee, and P. Resnick. Encouraging reading of diverse political viewpoints with a browser widget. In *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, pages 419–428, Boston, USA, 2013. AAAI press.

[164] S. A. Munson, S. Y. Lee, and P. Resnick. Encouraging reading of diverse political viewpoints with a browser widget. In *ICWSM*, 2013.

[165] S. A. Munson and P. Resnick. Presenting diverse political opinions: How and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1457–1466, New York, NY, USA, 2010. ACM.

[166] S. A. Munson and P. Resnick. Presenting diverse political opinions: how and how much. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1457–1466. ACM, 2010.

[167] S. A. Munson, D. X. Zhou, and P. Resnick. Designing interfaces for presentation of opinion diversity. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 3667–3672. ACM, 2009.

[168] S. A. Munson, D. X. Zhou, and P. Resnick. Sidelines: An algorithm for increasing diversity in news and opinion aggregators. In *ICWSM*, 2009.

[169] P. M. Napoli. Exposure diversity reconsidered. *Journal of Information Policy*, 1:246–259, 2011.

[170] G. E. Natali Helberger. Media diversity from the user's perspective: An introduction. *Journal of Information Policy*, 1:241–245, 2011.

[171] T. T. Nguyen, P.-M. Hui, F. M. Harper, L. Terveen, and J. A. Konstan. Exploring the filter bubble: The effect of using recommender systems on content diversity. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 677–686, New York, NY, USA, 2014. ACM.

[172] S. U. Noble. *Algorithms of Oppression: How search engines reinforce racism*. NYU Press, 2018.

[173] P. Norris. *Digital divide: Civic engagement, information poverty, and the Internet worldwide*. Cambridge University Press, 2001.

[174] A. Pal and S. Counts. Identifying topical authorities in microblogs. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pages 45–54, New York, NY, USA, 2011. ACM.

[175] B. Pan, H. Hembrooke, T. Joachims, L. Lorigo, G. Gay, and L. Granka. In google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, 12:801–823, 2007.

[176] A. Pansari and M. Mayer. Google Official Blog: This is a test. This is only a test. **http://googleblog.blogspot.in/2006/04/this-is-test-this-is-only-test.html**, Apr 2006.

[177] E. Pariser. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Group , The, 2011.

[178] S. Park, S. Kang, S. Chung, and J. Song. NewsCube: delivering multiple aspects of news to mitigate media bias. In *In Proceedings of ACM CHI*, 2009.

[179] M. J. Pazzani and D. Billsus. The adaptive web: Methods and Strategies of Web personalization. chapter Content-based recommendation systems, pages 325–341. 2007.

[180] M. Pennacchiotti and A.-M. Popescu. Democrats, republicans and starbucks afficionados: User classification in twitter. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 430–438, New York, NY, USA, 2011. ACM.

[181] M. Pennacchiotti and A.-M. Popescu. A machine learning approach to twitter user classification, 2011.

[182] Twitter Reaction to Events Often at Odds with Overall Public Opinion. *http://www.pewresearch.org/2013/03/04/twitter-reaction-to-events-often-at-odds-with-overall-public-opinion/*, 2013.

[183] J. Poirier, W. L. Bennett, G. J. Jerome, N. G. Shah, M. Lazo, H.-C. Yeh, J. M. Clark, and N. K. Cobb. Effectiveness of an activity tracker-and internet-based adaptive walking program for adults: a randomized controlled trial. *Journal of medical Internet research*, 18(2), 2016.

[184] Search Political Leaning of Twitter Users. **http://twitter-app.mpi-sws.org/search-political-bias-of-users/**, 2017.

[185] RealClearPolitics – Election 2016 – 2016 Republican Presidential Nomination. **http://tinyurl.com/us-republican-polling-data**, 2015.

[186] A. Popescu and G. Grefenstette. Mining user home location and gender from flickr tags. In *ICWSM*, 2010.

[187] ProCon.org. Differences in Conservative and Liberal Brains. *http://2012election.procon.org/view.resource.php?resourceID=004818*, February 2015.

[188] Purple Feed: Identifying High Consensus News Posts on Social Media. **https://twitter-app.mpi-sws.org/purple-feed/**, 2018.

[189] M. Purver and S. Karolina. Twitter Language Use Reflects Psychological Differences between Democrats and Republicans. *PLoS ONE*, 10(9):e0137422, 2015.

[190] C. Puschmann. How significant is algorithmic personalization in searches for political parties and candidates? *https://aps.hans-bredow-institut.de/personalization-google/*, 2017.

[191] C. Puschmann. How significant is algorithmic personalization in searches for political parties and candidates? (Part II). *https://aps.hans-bredow-institut.de/personalization-google-part-ii/*, 2017.

[192] D. Quercia, H. Askham, and J. Crowcroft. Tweetlda: Supervised topic classification and link prediction in twitter. In *Proceedings of the 4th Annual ACM Web Science Conference*, WebSci '12, pages 247–250, New York, NY, USA, 2012. ACM.

[193] K. R Canini, B. Suh, and P. Pirolli. Finding relevant sources in twitter based on content and social structure. 01 2010.

[194] D. Ramage, S. Dumais, and D. Liebling. Characterizing microblogs with topic models. In *Proceedings ICWSM 2010*. American Association for Artificial Intelligence, May 2010.

[195] D. Rao and D. Yarowsky. Detecting latent user properties in social media. In *In Proceedings of the NIPS MLSN Workshop*, 2010.

[196] J. Reagan. The "repertoire"; of information sources. *Journal of Broadcasting & Electronic Media*, 40(1):112–121, 1996.

[197] F. N. Ribeiro, F. B. Lucas Henrique, A. Chakraborty, J. Kulshrestha, M. Babei, and K. P. Gummadi. Media bias monitor: Quantifying biases of social media news outlets at large-scale. In *Proceedings of the 12th International AAAI Conference of Web and Social Media*, ICWSM '18, May 2015.

[198] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor. *Recommender Systems Handbook*. Springer-Verlag, Berlin, Heidelberg, 1st edition, 2010.

[199] R. Robertson, D. Lazer, and C. Wilson. Auditing Politically-Related Search Rankings and Suggestions. In *Proceedings of the 27th International Web Conference (WWW 2018)*, Lyon, France, April 2018.

[200] A. Sadilek, H. Kautz, and J. P. Bigham. Finding your friends and following them to where you are. In *Proceedings of the fifth ACM int'l conference on Web search and data mining*, 2012.

[201] A. Samoilenko, F. Karimi, D. Edler, J. Kunegis, and M. Strohmaier. Linguistic neighbourhoods: explaining cultural borders on wikipedia through multilingual co-editing activity. *EPJ data science*, 5(1):9, 2016.

[202] A. Samoilenko, F. Karimi, J. Kunegis, D. Edler, and M. Strohmaier. Linguistic influence patterns within the global network of wikipedia language editions. In *Proceedings of the ACM Web Science Conference*, page 54. ACM, 2015.

[203] C. Sandvig, K. Hamilton, K. Karahalios, and C. Langbort. Auditing algorithms: Research methods for detecting discrimination on internet platforms. *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 2014.

[204] C. Sandvig, K. Hamilton, K. Karahalios, and C. Langbort. Automation, algorithms, and politics | when the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication*, 10(0), 2016.

[205] D. Schkade, C. R. Sunstein, and R. Hastie. What happened on deliberation day? *California Law Review*, 95(3):915–940, 2007.

[206] J.-H. Schmidt. Twitter and the rise of personal publics. *Twitter and society*, pages 3–14, 2014.

[207] J.-H. Schmidt. Twitter friend repertoires: Introducing a methodology to assess patterns of information management on twitter. *First Monday*, 21(4), 2016.

[208] D. Sears and J. L. FREEDMAN. Selective exposure to information: A critical review. 31, 06 1967.

[209] B. C. Semaan, S. P. Robertson, S. Douglas, and M. Maruyama. Social media supporting political deliberation across multiple public spheres: towards depolarization. In *In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1409–1421. ACM, 2014.

[210] G. Shani and A. Gunawardana. Evaluating Recommender Systems. Technical Report MSR-TR-2009-159, Microsoft Research, November 2009.

[211] N. K. Sharma, S. Ghosh, F. Benevenuto, N. Ganguly, and K. Gummadi. Inferring who-is-who in the twitter social network. *ACM SIGCOMM Computer Communication Review*, 42(4):533–538, Sept. 2012.

[212] Y. Shi, K. Mast, I. Weber, A. Kellum, and M. Macy. Cultural fault lines and political polarization. In *Proceedings of the 2017 ACM on Web Science Conference*, WebSci '17, pages 213–217, New York, NY, USA, 2017. ACM.

[213] P. J. Shoemaker, T. P. Vos, and S. D. Reese. Journalists as gatekeepers. *The handbook of journalism studies*, 73, 2009.

[214] S. Shorey and P. N. Howard. Automation, algorithms, and politics| automation, big data and politics: A research review. *International Journal of Communication*, 10:24, 2016.

[215] L. M. Smith, L. Zhu, K. Lerman, and Z. Kozareva. The role of social media in the discussion of controversial topics. In *Social Computing (SocialCom), 2013 International Conference on*, pages 236–243. IEEE, 2013.

[216] J. K. Sorensen and J.-H. Schmidt. An algorithmic diversity diet?: Questioning assumptions behind a diversity recommendation system for psm, 9 2016.

[217] A. Springer, V. Hollis, and W. Steve. Dice in the black box: User experiences with an inscrutable algorithm. In *The AAAI 2017 Spring Symposium on Designing the User Experience of Machine Learning Systems*. AAAI, 2017.

[218] H. Steck. Item popularity and recommendation accuracy. In *Proceedings of the Fifth ACM Conference on Recommender Systems*, RecSys '11, pages 125–132, New York, NY, USA, 2011. ACM.

[219] N. Stroud. *Niche News: The Politics of News Choice*. Oxford University Press., 2011.

[220] C. R. Sunstein. The law of group polarization. *Journal of Political Philosophy*, 10(2):175–195, 2002.

[221] L. Sweeney. Discrimination in online ad delivery. *Queue*, 11(3):10, 2013.

[222] Y. Takhteyev, A. Gruzd, and B. Wellman. Geography of twitter networks. *Social Networks*, 34(1):1–25, 2011.

[223] H. Taneja, J. G. Webster, E. C. Malthouse, and T. B. Ksiazek. Media consumption across platforms: Identifying user-defined repertoires. *New Media & Society*, 14(6):951–968, 2012.

[224] J. Tang, X. Hu, and H. Liu. Social recommendation: a review. *Social Network Analysis and Mining*, 3(4):1113–1133, Dec 2013.

[225] H. Tavani. Search engines and ethics. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2014 edition, 2014.

[226] J. Teevan, D. Ramage, and M. R. Morris. #twittersearch: A comparison of microblog search and web search. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pages 35–44, New York, NY, USA, 2011. ACM.

[227] D. Trielli, S. Mussenden, and N. Diakopoulos. Why Google Search Results Favor Democrats. *http://tinyurl.com/google-search-favor-dems*, Dec 2015.

[228] Z. Tufekci. Algorithmic harms beyond facebook and google: Emergent challenges of computational agency. *J. on Telecomm. & High Tech. L.*, 13:203, 2015.

[229] A.-M. Turner. Mashable: A brief guide to using Twitter's 'Discover' tab. **https://mashable.com/2015/01/26/twitter-discover-tab/#.OmYIUx2o5qw**, January 2012.

[230] Twitter Search. **https://twitter.com/search/**.

[231] Twitter Blogs | Discover: Improved personalization algorithms and real-time indexing. **https://blog.twitter.com/2012/discover-improved-personalization-algorithms-and-real-time-indexing**.

[232] Twitter promoted accounts. **https://business.twitter.com/products/promoted-accounts**.

[233] Twitter Help Center: About tailored suggestions. **https://support.twitter.com/articles/20169421**.

[234] E. Van Couvering. *Search engine bias: the structuration of traffic on the World-Wide Web*. PhD thesis, The London School of Economics and Political Science, 2010.

[235] K. Van Rees and K. Van Eijck. Media repertoires of selective audiences: The impact of status, gender, and age on media use. *Poetics*, 31(5-6):465–490, 2003.

[236] S. Vargas and P. Castells. Rank and relevance in novelty and diversity metrics for recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems*, RecSys '11, pages 109–116, New York, NY, USA, 2011. ACM.

[237] L. Vaughan and M. Thelwall. Search engine coverage bias: Evidence and possible causes. *Information Processing and Management*, 40(4):693–707, May 2004.

[238] C. Wagner, V. Liao, P. Pirolli, L. Nelson, and M. Strohmaier. It's not in their tweets: Modeling topical expertise of twitter users. In *2012 International Conference on Privacy, Security, Risk and Trust, PASSAT 2012, and 2012 International Confernece on Social Computing, SocialCom 2012, Amsterdam, Netherlands, September 3-5, 2012*, pages 91–100, 2012.

[239] A. Watch. DatenspendeBTW17. *https://datenspende.algorithmwatch.org/*, 2017.

[240] I. Weber and V. R. K. Garimella. Using co-following for personalized out-of-context twitter friend recommendation. 2014.

[241] I. Weber, V. R. K. Garimella, and E. Borra. Mining web query logs to analyze political issues. In *Proceedings of the 4th Annual ACM Web Science Conference*, WebSci '12, pages 330–334, New York, NY, USA, 2012. ACM.

[242] I. Weber, V. R. K. Garimella, and A. Teka. Political hashtag trends. In *European Conference on Information Retrieval*, pages 857–860. Springer, 2013.

[243] J. G. Webster and T. B. Ksiazek. The dynamics of audience fragmentation: Public attention in an age of digital media. *Journal of communication*, 62(1):39–56, 2012.

[244] J. G. Webster and P. F. Phalen. The mass audience: Rediscovering the dominant model (book review). *Journalism and Mass Communication Quarterly*, 74(4):898, 1998.

[245] M. J. Welch, J. Cho, and C. Olston. Search result diversity for informational queries. In *In Proceedings of WWW*, 2011.

[246] J. Weng, E.-P. Lim, J. Jiang, and Q. He. Twitterrank: Finding topic-sensitive influential twitterers. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, pages 261–270, New York, NY, USA, 2010.

[247] M. C. Willemsen, M. P. Graus, and B. P. Knijnenburg. Understanding the role of latent feature diversification on choice difficulty and satisfaction. *User Modeling and User-Adapted Interaction*, 26(4):347–389, Oct 2016.

[248] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang. Quantifying political leaning from tweets and retweets. In *In Proceedings of AAAI ICWSM*, 2013.

[249] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang. Quantifying political leaning from tweets, retweets, and retweeters. *IEEE Transactions on Knowledge and Data Engineering*, 2016.

[250] S. Wu, J. M. Hofman, W. A. Mason, and D. J. Watts. Who says what to whom on twitter. In *Proceedings of the 20th International Conference on World Wide Web*, WWW '11, pages 705–714, New York, NY, USA, 2011. ACM.

[251] Z. Xu, L. Ru, L. Xiang, and Q. Yang. Discovering user interest on twitter with a modified author-topic model. In *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01*, WI-IAT '11, pages 422–429, Washington, DC, USA, 2011. IEEE Computer Society.

[252] S.-H. Yang, A. Kolcz, A. Schlaikjer, and P. Gupta. Large-scale high-precision topic modeling on twitter. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, pages 1907–1916, New York, NY, USA, 2014. ACM.

[253] T. Yano, P. Resnik, and N. A. Smith. Shedding (a thousand points of) light on biased language. In *Proc NAACL HLT Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT)*, 2010.

[254] S. Yardi and D. Boyd. Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*, 30(5):316–327, 2010.

[255] S. Yigit-Sert, I. S. Altingovde, and Ö. Ulusoy. Towards detecting media bias by utilizing user comments. In *Proceedings of the 8th ACM Conference on Web Science*, pages 374–375. ACM, 2016.

[256] E. Yilmaz and J. A. Aslam. Estimating average precision with incomplete and imperfect judgments. In *In Proceedings of ACM CIKM*, 2006.

[257] E. Yom-Tov, S. Dumais, and Q. Guo. Promoting Civil Discourse Through Search Engine Diversity. *Social Science Computer Review*, 32:145–154, 2013.

[258] E. Yuan. News consumption across multiple media platforms: A repertoire approach. *Information, Communication & Society*, 14(7):998–1016, 2011.

[259] E. J. Yuan and J. G. Webster. Channel repertoires: Using peoplemeter data in beijing. *Journal of Broadcasting & Electronic Media*, 50(3):524–536, 2006.

[260] M. B. Zafar, K. P. Gummadi, and C. Danescu-Niculescu-Mizil. Message impartiality in social media discussions. In *Proceedings AAAI ICWSM*, 2016.

[261] F. A. Zamal, W. Liu, and D. Ruths. Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors, 2012.

[262] A. Zeng, C. H. Yeung, M.-S. Shang, and Y.-C. Zhang. The reinforcing influence of recommendations on global diversification. *EPL (Europhysics Letters)*, 97:18005, Jan. 2012.

[263] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li. Comparing twitter and traditional media using topic models. In *Proceedings of the 33rd European Conference on Advances in Information Retrieval*, ECIR'11, pages 338–349, Berlin, Heidelberg, 2011. Springer-Verlag.

[264] D. X. Zhou, P. Resnick, and Q. Mei. Classifying the political leaning of news articles and users from user votes. In *In Proceedings of AAAI ICWSM*, 2011.

[265] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J. R. Wakeling, and Y.-C. Zhang. Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences*, 107(10):4511–4515, 2010.

[266] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, pages 22–32, 2005.

[267] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web*, WWW '05, pages 22–32, New York, NY, USA, 2005. ACM.