

Public Digital Media and Democracy

Constructing Instruments for Measuring Expressions of Support
for Democracy on Twitter

by
Barend Frederik Lutz

*Thesis presented in partial fulfilment of the requirements for the degree of
Master of Arts (International Studies)
in the Faculty of Arts and Social Sciences
at Stellenbosch University*



Supervisor: Prof Pierre du Toit
MARCH 2013

Declaration

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

November 2012

Copyright © 2012 Stellenbosch University
All rights reserved

Abstract

The constant growth in information and communications technology (ICT) broadens the capabilities of researchers to understand and interpret the social world. New multidisciplinary methods of measuring social realities are constantly being developed and improved. This methodological study aims to incorporate novel methods of automated, computerised content- and sentiment analysis software in order to produce a usable instrument for measuring public expressions on democracy from the social network site, Twitter. Adapting methods developed by O'Connor *et al.* (2010) and using an online platform called Chatterbox Analytics Ltd., this study attempts to complement traditional survey research data on democracy. The creation of this method allows researchers to automatically and without deep programming knowledge, extract and analyse opinionated data from a substantial segment of the population, namely Twitter users. This is the first study, as far as this researcher is aware, that develops automated instruments for the measurement of expressions on democracy from online social networks.

The rationale for the development of this instrument lies in the apparent recent rise in negative sentiment on democracy in academic spheres and political realities alike. Throughout history support for democracy has swayed, which in turn placed democracy on a rocky path of development. As negative opinions on democracy became overwhelming this system of governance was forced to adapt by changing its core structure. This could be the case again today. Authors such as Crozier, Huntington and Watunaki (1975) have already warned of a crisis of democracy in the 1970s. More recently authors such as Dalton (2004) and van Beek (2012) expressed opinions indicating that globally citizens are steadily becoming more critical of political parties, politicians, political policies and the whole democratic process. These negative sentiments might be indicative of what van Beek (2012:12) calls “an erosion in support for democracy”.

Traditional survey projects examining opinions on democracy, such as the World Values Survey, offer insights into global expressions on democracy, but are time consuming, costly and complicated to conduct. The method developed in this study allows researchers to complement traditional survey data with insights that are automatically gathered and analysed from the influential social media network, Twitter.

This introductory study finds that automated analysis of expressions on democracy from Twitter is indeed feasible. In order to express the extent to which the developed method is feasible, this study offers an example case, examining expressions on democracy from Twitter for the period of 1 May to 31 July 2012. Furthermore the study offers a quantitative manual evaluation of the accuracy of the developed measurement instrument. With the present level of content- and sentiment analysis technology this study finds that the accuracy of the results from this method, though informative, is still limited. The study therefore concludes with an advisory section highlighting methods for future studies to improve on the accuracy of this measurement instrument.

Opsomming

Deur die konstante groei in inligting- en kommunikasietegnologie (IKT) word die veld waaruit navorsers kan put om die sosiale wêreld te verstaan en te interpreteer al groter. Nuwe multidissiplinêre metingsmetodes om die sosiale werklikheid te verstaan, word konstant ontwikkel en verbeter. Hierdie metodologiese studie beoog om nuwe metodes van geoutomatiseerde, gerekenaariseerde, inhouds- en sentimentanalisesagteware saam te voeg om 'n bruikbare instrument vir die meting van publieke uitsprake oor demokrasie, op die sosiale netwerk Twitter te skep. Metodes wat deur O'Connor *et al.* (2010) en die aanlyn platform, Chatterbox Analytics Ltd ontwikkel is, is aangepas om 'n instrument te skep wat gebruik kan word om by te dra tot tradisionele opname-navorsingsdata oor demokrasie. Hierdie nuwe metode sal navorsers toelaat om outomaties en sonder veel programmeringsvaardighede opiniegelaai data van 'n beduidende segment van die samelewing, naamlik Twitter gebruikers, te verkry en te analiseer. Hierdie is die eerste studie, sover hierdie navorsers bewus is, wat geoutomatiseerde instrumente ontwikkel vir die meting van uitsprake oor demokrasie op sosiale netwerke.

Die beweegrede vir die ontwikkeling van hierdie instrument vloei uit die oënskynlike onlangse styging in negatiewe sentiment oor demokrasie in sowel akademiese- as politieke sirkels. Die ondersteuning vir demokrasie het dwarsdeur die geskiedenis sy hoogste en laagtepunte gehad en hierdie onsekerheid het die ontwikkelingspad van demokrasie taamlik met dorings besaai. Die hele kernstruktuur van hierdie regeringstelsel was geforseer om hewig aan te pas elke keer wanneer die negatiewe opinies oor demokrasie oorweldigend geraak het. Dit wil voorkom asof demokrasie weer op die rand van verandering is. Skrywers soos Crozier, Huntington en Watunaki (1975) het alreeds in die 1970's gewaarsku teen 'n krisis rondom demokrasie. Meer onlangs het skrywers soos Dalton (2004) en van Beek (2012) opinies gelug dat burgers reg oor die wêreld stadig maar seker besig is om meer krities op politieke partye, politici en oor politieke beleide en die demokratiese proses in geheel raak. Hierdie negatiewe sentimente mag dalk 'n aanduiding wees van "'n erosie in die ondersteuning van demokrasie,'" soos Van Beek (2012:12) dit noem.

Tradisionele opname-navorsingsprojekte wat na opinies oor demokrasie kyk, soos die 'World Values Survey', bied wel 'n blik op globale uitsprake oor demokrasie, maar is duur, tydsaam en moeilik om deur te voer. Die metode wat in hierdie studie ontwikkel is, kan bykomende insigte tot tradisionele opname-navorsingsprojekte bring. Dié insigte word outomaties ingesamel en geanaliseer vanuit die invloedryke sosiale netwerk, Twitter.

Hierdie inleidende studie bevind dat geoutomatiseerde analise van uitsprake oor demokrasie op Twitter inderdaad moontlik is. Ten einde die mate van sukses van hierdie metode wat ontwikkel is te illustreer, analiseer hierdie studie 'n datagreep van uitsprake oor demokrasie uit Twitter oor die tydperk 1 Mei tot 31 Julie 2012. Hierbenewens bied die studie ook 'n kwantitatiewe waardasie per hand wat die akkuraatheid van die ontwikkelde meetingsinstrument toets. Met die huidige ontwikkelingsvlak van inhouds- en sentiment-analisetegnologie vind hierdie studie dat die akkuraatheid van hierdie metode, alhoewel informatief is, tog beperk is. Die studie sluit af met 'n afdeling wat advies oor hoe verdere studies die akkuraatheid van hierdie meetinstrument kan verbeter, toelig.

Acknowledgements

I would like to express my heartfelt gratitude towards everyone that has supported me during the writing of this thesis. I would like to start by showing my most humble gratitude to the Lord for giving me the strength, courage and ability to complete this thesis. To Professor du Toit, my inspiring supervisor, who has motivated and continuously guided my work, I extend my profound gratitude. Also to Derica Lambrechts who, with her keen logical approach, provided truly helpful insight. Next, to Professor Anthony Leysens and the Political Science Department for taking such an interest in my work and allowing me to represent the Department at the South African Association of Political Science Conference. I would also like to express the deepest appreciation especially to Professor Gert-Jan van Rooyen and Dr Herman Engelbrecht; the MIH Media Lab and Naspers, who have not only financed this thesis, but also provided the stimulating environment and continues support, which made it a pleasure to write. Thank you also to Chatterbox Analytics, and specifically Dr Stuart Battersby, who allowed me to use their platform and provided helpful insight throughout the thesis. I would then like to extend copious amounts of thanks to my family, who have supported me throughout my life and my studies. I would be in danger if I did not note my appreciation to my brother who has provided his zone of econometrical expertise in support of this thesis. Last, but far from least, I would like to thank my lovely Marina for her continuous support, inspiration and editorial help.

Table of Contents

Declaration.....	ii
Abstract.....	iii
Opsomming.....	iv
Acknowledgements.....	v
List of Abbreviations.....	x
List of Figures.....	xi
Chapter 1: Introduction.....	1
1.1 INTRODUCTION TO THE TOPIC.....	1
1.1.1 The Growth of Information and Communications Technology and the study of Democracy 1	
1.1.2 Research Problem.....	2
1.1.3 Research Aim.....	4
1.1.4 Rationale and Significance of the Study.....	4
1.1.4.1 Why look at expressions on Twitter?.....	6
1.1.4.2 Why measure expressions on democracy?.....	9
1.2 THEORETICAL FRAMEWORK.....	10
1.3 RESEARCH METHODOLOGY.....	11
1.3.1 Research Design.....	11
1.3.1.1 Research Question(s).....	11
1.3.1.2 Hypothesis.....	12
1.3.2 Units of Analysis.....	13
1.3.3 Levels of Analysis.....	13
1.3.4 Time Dimension.....	13
1.3.5 Data Collection and Analysis.....	13
1.4 CONCEPTUALISATION OF CORE TERMINOLOGY.....	14
1.4.1 To what ‘extent’.....	15

1.4.2	Public Expression.....	15
1.4.3	Democracy	15
1.4.4	Interest Articulated Sentiments.....	16
1.4.5	Tweet.....	16
1.4.6	Content and Sentiment Analysis Software.....	17
1.4.7	The Public Sphere	18
1.4.8	Twitter.....	19
1.4.9	Chatterbox Analytics Platform.....	19
1.4.10	Big Data	19
1.5	LIMITATIONS AND DELIMITATIONS OF THE STUDY	20
1.6	STRUCTURE AND CHAPTER OUTLINE	22
1.7	CONCLUSION.....	22
Chapter 2: An Examination of Democracy.....		23
2.1	INTRODUCTION	23
2.2	UNDERSTANDING THE TERM DEMOCRACY	23
2.3	THE TURBULENT DEVELOPMENT OF DEMOCRACY	24
2.3.1	The Formative Years of Democracy: Athens 508 BCE to the Renaissance (1700+-).....	24
2.3.2	The Renaissance and the Birth of Liberalism	27
2.3.3	The Three Waves of Democracy.....	30
2.4	A CONTEMPORARY VIEW OF DEMOCRACY: EFFECTIVE DEMOCRACY (WHY PEOPLE MIGHT BE NEGATIVE).....	32
2.4.1	Empowered Citizens	34
2.5	EVOLVING VIEW OF DEMOCRACY – THE RISE OF NEGATIVE SENTIMENT.....	35
2.6	CONCLUSION.....	38
Chapter 3: The Development of Measuring Expressions on Democracy		39
3.1	INTRODUCTION	39
3.2	THE VALUE OF MEASURING EXPRESSIONS ON DEMOCRACY	39

3.3	TRADITIONAL METHODS OF MEASURING PUBLIC EXPRESSION ON DEMOCRACY	40
3.3.1	Global Research Projects	41
3.3.1.1	The World Values Survey	41
3.3.1.2	Voice of the People	43
3.3.2	Regional Research Projects.....	45
3.3.2.1	AfroBarometer	45
3.3.2.2	LatinoBarometer (Latinobarómetro).....	47
3.4	TWITTER AS A NEW PUBLIC SPHERE.....	49
3.5	EMERGING METHODS OF MEASURING PUBLIC EXPRESSIONS	50
3.6	CONCLUSION.....	52
Chapter 4: Operationalisation and Measurement.....		53
4.1	INTRODUCTION	53
4.2	OPERATIONALISATION.....	53
4.2.1	Extent	54
4.2.2	Public Expressions	54
4.2.3	Sentiment Classification	54
4.2.4	The Measurement Instrument: Chatterbox Analytics Platform	55
4.3	DATA COLLECTION	59
4.4	REPRESENTING MEASURED EXPRESSIONS ON DEMOCRACY FROM TWITTER.....	59
4.4.1	Utilising Histograms	60
4.4.2	Utilising Scatter Graphs	63
4.4.3	Utilising a Bar Chart	65
4.5	ACCURACY OF THE MEASUREMENT INSTRUMENT	66
4.5.1	Determining the Accuracy of Measurement Instrument	66
4.5.2	Manual Classification of Random Sample.....	67
4.6	CONCLUSION.....	71

Chapter 5: Findings and Recommendations	73
5.1 INTRODUCTION	73
5.2 BRIEF OVERVIEW OF ARGUMENT	73
5.3 CASE STUDY: PUBLIC EXPRESSIONS ON DEMOCRACY FROM TWITTER	75
5.4 ACCURACY OF MEASUREMENT INSTRUMENT	77
5.5 FEASABILITY OF MEASURING PUBLIC EXPRESSIONS ON DEMOCRACY FROM TWITTER.....	79
5.6 OTHER POTENTIAL (INTERESTING) FINDINGS	80
5.7 FUTURE RESEARCH	82
5.8 CONCLUSION.....	86
Bibliography	87
APPENDIX A: CALCULATIONS	99
APPENDIX B: EXAMPLE OF MANUAL CLASSIFICATION	101
APPENDIX C: INTERESTING FOR FUTURE CONSIDERATION	106
APPENDIX D: GROUPING NEGATIVE SENTIMENT	108

List of Abbreviations

9/11	September 11, 2001
API	Application Programming Interface
CDD	Centre for Democratic Development
FHI	Freedom House Index
HTTP	Hypertext Transfer Protocol
ICT	Information and Communications Technology
IDASA	Institute for Democracy in South Africa
NLP	Natural Language Processing
USA	United States of America
VOTP	Voice of the People
WEF	World Economic Forum
WVS	World Values Survey

List of Figures

Figure 1: Growth in Information Created – 2005-2011.....	5
Figure 2: Social Media Growth – 2006-2011.....	8
Figure 3: Top 20 countries in terms of Twitter accounts, 2012.....	8
Figure 4: Levels of classification.....	14
Figure 5: Information gained from a Tweet object.....	17
Figure 6: Global Trends in Governance, 1800-2011.....	31
Figure 7: VOTP Global Satisfaction with Democracy.....	44
Figure 8: VOTP Global Sentiment on Democracy.....	44
Figure 9: Chatterbox Analytics Platform.....	57
Figure 10: Chatterbox map representation of global tweets on Democracy.....	57
Figure 11: Chatterbox sentiment classification histogram.....	60
Figure 12: Chatterbox sentiment classification monthly density spread.....	61
Figure 13: Chatterbox sentiment classification scatter graph.....	62
Figure 14: Chatterbox sentiment classification monthly scatter graphs.....	64
Figure 15: Aggregated Chatterbox sentiment classification.....	65
Figure 16: Manual Classification of Negative Chatterbox classified Tweets.....	68
Figure 17: Manual Classification of Positive Chatterbox classified Tweets.....	69
Figure 18: Estimation of Actual Sentiment Levels.....	70
Figure 19: Aggregated Chatterbox sentiment classification.....	75
Figure 20: Estimation of Actual Sentiment Levels.....	77
Figure 21: Example Graph: Key Word Search.....	83

Chapter 1: Introduction

1.1 INTRODUCTION TO THE TOPIC

1.1.1 The Growth of Information and Communications Technology and the study of Democracy

The growth in information and communications technology (ICT) broadens the capabilities of researchers to understand and interpret the social world. New methods of measuring social realities are constantly being developed. Online social media networks, a form of ICT, have, in recent years, received quite a lot of attention from the academic and the corporate world alike. With the growth of online social media networks, the availability of freely available, user-generated data is increasing. The second biggest online social network, after Facebook, is Twitter¹ and the data created on this network are freely available to the public. Computer scientists and social scientists alike are currently exploring the limits of the usability, value and applicability of the data that are constantly being created online². This is also the goal of this study.

Taking advantage of specific technological advancements in the field of automated content and sentiment analysis, it might be possible to gain a deeper understanding of certain important aspects of political social life. This is where the study of democracy comes to play. Within the academic school of political culture, there exists the belief that the functioning and persistence of democracy and democratic institutions is integrally linked with dominant mass tendencies in individual-level attitudes and value orientations (Lerner, 1958; Almond and Verba, 1963; Eckstein, 1966, Inglehart and Welzel, 2005). What these studies show is that what people think about democracy and democratic institutions is important to the functioning of democracy as a political system of governance.

The rationale for the link between opinions of citizens and the functioning of a democracy is based in the idea of a public sphere. This idea will be further elaborated, but in brief it refers to a physical or virtual space where the public can meet in order to discuss, debate and deliberate on public affairs. The public sphere lies between state and society, and is an essential component of socio-political organisation. It provides a form of legitimacy and accountability to a government and a space for citizens to participate in public affairs (Castells, 2008: 78-79).

¹ Twitter is an online social networking- and microblogging service that enables its users to send and read text-based messages of up to 140 characters, known as "tweets". This service will be discussed in more detail throughout the thesis.

² See for example: Hopkins and King (2010); Diakopoulos and Shamma (2011); O'Connor *et al.* (2010) and Bollen *et al.* (2009).

In this study the social media site, Twitter, will be examined as a type of public sphere. Within this public sphere, global citizens are constantly discussing a multitude of issues. One of these issues that receives constant discussion on Twitter is democracy. If one can measure public expressions on democracy, one would in theory be able to gain deeper insights on the health of democracy. There are however various ways in which public expressions on democracy can be gauged, as will be shown in Chapter Three. In the past, most studies that attempted to measure expressions on or opinions about democracy have used surveys consisting of static questionnaires. The belief of this researcher, however, is that the insights gained from these survey studies can be complimented by additional expression data gathered from online social media sites. If one views Twitter as a public sphere and examines public deliberation and discussion regarding the topic of democracy, one might gain valuable insights as to the global opinion on democracy. In this thesis the methodology of measuring public expressions on democracy from Twitter will be developed. This methodology will be built upon, tested and expressed through a descriptive analysis of expressions on democracy from Twitter for the period of 1 May to 31 July 2012.

1.1.2 Research Problem

It has now been almost forty years since the beginning of what Samuel P. Huntington (1991), famously called the *Third Wave of Democracy*. Within these forty years it has been stated that human socio-cultural evolution has reached a pinnacle, culminating in liberal democracies and free market capitalism (Fukuyama, 1989; 1992). This radical idea has been criticised heavily; most notably by Huntington (1993) in his paper, *The Clash of Civilizations*. With the events of 9/11 the dominance of liberal democracies and free market capitalism seemed to be strongly challenged. The global financial crises that began in 2008 and the ongoing crises in the Eurozone have, according to authors such as van Beek (2012:12), finally ended the dominance of the powerful developed democracies with their free market capitalist ideals. Van Beek (2012:12) states that these events might even start to erode global confidence in democracy.

This warning by van Beek resonates with the ideas of the political culture school of political science. Tocqueville *et al.* (1945:299) stated that the notions and opinions (“habits of the heart”) - which form the “whole moral and intellectual condition of a people” - could be influenced by real-world events. Diamond (1993:1-4) has also reiterated these sentiments, stating that political culture undergoes certain changes which could move a political system towards or away from the current ideal of democracy.

It is however important to note that the current ideal of democracy is constantly evolving. Countless writers³ have examined and expressed their opinions on the rocky development that the ideal of democracy has experienced throughout history. This road of development will be briefly examined in Chapter Two of this study.

With the evolution - and the potential recent erosion - of democracy in mind, this thesis aims to examine public expressions on democracy in order to gauge the 'habits of the heart' that are prevalent in contemporary society. There is however an ongoing debate on measuring mass attitudes on democracy and defining democracy (Hadenius and Teorell, 2006; Inglehart and Welzel, 2005; Welzel, 2007; Welzel and Inglehart, 2008; Knutsen, 2010). There are many empirical research programmes that aim to measure democracy and expressions on democracy, such as the Freedom House Index, the AfroBarometer, the European Values Survey, the New Russia Barometer, the LatinoBarometer, and the World Values Survey. What these programmes have in common, however, is more than the measurement of democracy - they all use survey data in the form of static questionnaires in order to gain standardised information.

The strength of the survey data, with regards to democratic measurement, is that it attempts to provide equal consideration to all opinions (Verba, 1996). Survey data measurement is however tedious and expensive to gather. There are also a number of problems, such as social pressure, that are being placed on respondents of questionnaires (Berinsky, 1999:1210). Survey data therefore, by no means, provides perfectly objective and perfectly efficient measurements of public expressions. This is problematic because, as noted above, it is paramount to the functioning of democracy to understand mass attitudes on the subject.

It has to be stated clearly that it is not the goal of this thesis to streamline survey data measurement. It is however the opinion of this author that additional insight on public expressions could be gained from examining user-generated content from online sources. As will be shown in subsequent sections of this chapter, online user-generated content represents expressions from a large proportion of the human population. These online users are constantly generating opinion-rich data that might add to one's understanding of mass attitudes. This study views the expressions on Twitter as a form of discourse within a Habermasian⁴ (1989) public sphere. The expressions on democracy from this sphere could then be indicative as to the legitimacy of democracy as a whole (Habermas, 1976; CommGAP, 2012).

³ Their ideas will be discussed throughout the thesis, but to name a few: Crozier, Huntington and Watunaki (1975); Arendt (1990); Dahl (1998); Held (2006); Welzel and Inglehart (2008) and (Dryzek & Dunleavy, 2009).

⁴ This concept will be discussed in more detail subsequently, but refers to a type of public sphere discussed by Habermas (1989), referring to a virtual or physical place (sphere) wherein public opinion is generated.

The central problem of this research is that traditional survey research methods cannot access (i.e. describe, analyse and interpret) the dimension of the public sphere that is created by the digital media and expressed through social networks. There might exist a mass erosion of attitudes on democracy within this digital realm, due to events such as the 2008-9 financial crisis, and this researcher believes that novel methods of measuring public expressions could highlight this erosion. There are various methods in which this potential erosion of mass attitudes might be measured, but this thesis proposes a novel method of automated content analysis on online user-generated data. In this thesis, the methodology of measuring these mass attitudes from Twitter will be developed. The methodology, which will be developed in Chapter Three, aims to allow one to automatically gather opinion-rich data on democracy from Twitter. This opinion-rich data are then analysed in the descriptive section of this thesis in order to see whether expressions on democracy from Twitter are more negative or positive. If it can be shown that expressions on democracy are more negative it would serve to support claims such as that of van Beek (2012:12) on the potential erosion of democracy. This more ontological, descriptive analysis section exemplifies and tests the methodology that will be created in this thesis.

1.1.3 Research Aim

This is an empirical study that aims to explore and analyse primary data that are publicly made available through Twitter. The rationale for this methodological and descriptive research is to gain a better understanding of public expressions on democracy and also to develop a method for examining public expressions from online sources that can be utilised in subsequent studies. This research therefore looks at both the ontological question as to the nature of expressions on democracy and the methodological question as to how an inquirer should seek knowledge on expressions on democracy from Twitter.

The aim of this study can be summarised as an attempt to develop and test instruments with which to measure public expression of support for democracy between 1 May and 31 July 2012 according to interest-articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software.

1.1.4 Rationale and Significance of the Study

The ability to effectively and systematically categorise and analyse text documents has been an important part of history. In the late 1600s all non-religious texts were tracked by the Catholic Church, for example (Krippendorff, 2004). This ability has also been used in the social sciences. Social scientists such as Berelson, Waples, Bradshaw, Brelson and de Grazia were some of the early adopters of what they called “content analysis” in the 1940s (Hopkins and King, 2010: 229). Our ability to effectively analyse ever larger data sets of texts has also developed through the years (Neuendorf, 2002). Today there is an

explosive increase in blogs, social networks, web pages, digitised books and articles, and many other forms of electronic documents, which means that there is much more information available for a variety of content analysis applications (Lyman and Varian, 2003). The graph below (*Figure 1*), from a 25 February 2010 Economist article, shows how this rapid increase in information is actually overtaking our ability to store all the information. In 2011 it was estimated that about 1750 exabytes of information was created, but the global data storage capacity was only about 800 exabytes, for example.

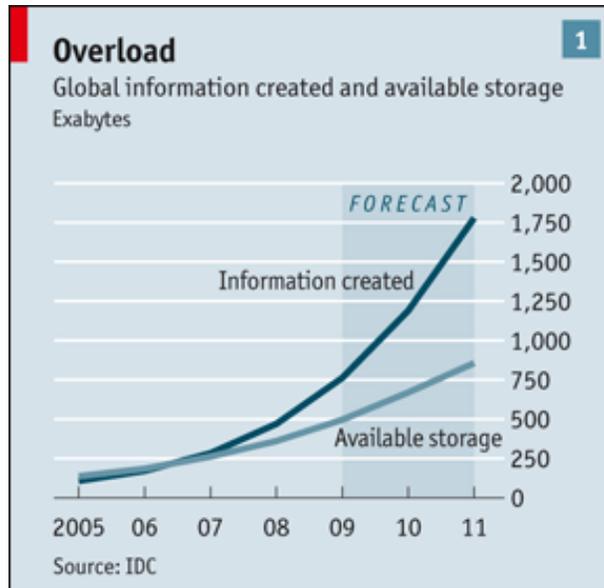


Figure 1: Growth in Information Created – 2005-2011 - Source: (Economist, 2010).

There have been studies that manually examine thousands of blogs in order to gain a better understanding of political events such as the US elections (Adamic and Glance, 2005). Manual methods are time intensive and work if one wishes a small snapshot of opinions on the blogosphere, but as the authors admit, a more extensive analysis of the blogosphere is needed (Adamic and Glance, 2005: 15). This is where automated data analysis methods offer help.

The rise in the quantity of information necessitates more efficient, automated methods of data analysis. Within the field of computer science there are already a number of automated content analysis applications which are well suited for computer science purposes. Authors such as Hopkins and King (2010), however, believe that these applications should be and can be adapted to fit social sciences better (Hopkins and King, 2010: 229).

With these computerised methods documents can be divided and generalised into specific themes which give one as a political scientist the ability to examine documents that are from a specific area of interest.

In this thesis, this function of a content analysis programme will be paramount as it will highlight the information that is relevant to the chosen topic of democracy.

Within blogs and micro blogs (such as Twitter) a large number of people are voicing their opinions on various issues. One could view these opinions as a visible public expression of opinions, or interest articulation, which might be more valuable to political scientists than attitudes and non-attitudes expressed in traditional survey responses (Hopkins and King, 2010: 231);(Ginsberg, 1986). Not all scholars however agree with the opinion that these visible expressions of opinion are more valuable than the poll data. In a critical review of Ginsberg's book, *The Captive Public: How Mass Opinion Promotes State Power*, Smith (1987:520-522) highlights some of the problems such as overemphasised opinions of certain actors or strong views and the inability to determine whether an opinion comes from a knowledgeable person. With regard to this discussion on the preference of polling or visible expression methods of measuring public opinion, it is important to remember that neither of these techniques is infallible. It is however encouraging to see that, as O'Conner *et al.* (2010) note: "...expensive and the time-intensive polling can be supplemented or supplanted with the simple-to-gather text data that are generated from online social networking."

The O'Conner *et al.* (2010) study focused on comparing public opinion measured from polls with sentiment measured from text. They correlated sentiments gathered from Twitter during 2008-2009 with traditional approval polling and consumer confidence polling results for the same period. Their findings showed that, even with a relatively simple sentiment detector, using Twitter data, the results were comparable to much more time-intensive (and expensive) polling methods.

1.1.4.1 Why look at expressions on Twitter?

At the World Economic Forum (WEF) meeting in Davos, Switzerland, in January 2012 "Big Data"⁵ was heralded as a new type valuable economic asset (WEF, 2012). The sentiment that is rife in the academic and business world alike, is that Big Data and the rapidly advancing ICT surrounding Big Data can and will "open the door to a new approach to understanding the world and making decisions" (Lohr, 2012). On the one hand the availability and quantity of data are increasing rapidly. Social media sites, mobile devices, digital sensors, electrical metres and countless other interconnected devices and networks are generating data at a growing rate (40 percent projected annual growth) (Nie, 2011: 1). Furthermore, the growth in technologies such as natural language processing programmes, data mining programmes,

⁵ Big Data refers to the analysis of massive data sets (today even terabytes, petabytes and exabytes in size) which require exceptional technologies in order to be processed within reasonable times. The analysis of data gathered from social media sites such as Twitter can be seen as Big Data.

sentiment/opinion mining/analysis programmes and the computational capabilities of computers enables one to analyse and use this Big Data more effectively than ever before.

This rapid growth in the availability of data and the capacity of ICT can also be of great value to the field of political science. This author believes, as a number of scholars before him (O'Connor *et al.*, 2010; Zhao *et al.*, 2011; Diakopoulos and Shamma, 2010) that mass attitudes could indicate public opinion and can be extracted from social media sites and used as a supplement to traditional “poll measured”, survey data. In this thesis, data mined from Twitter from *1 May and 31 July 2012* (104 037 Tweets) will be analysed. This analysis is conducted through the use of an automated content and sentiment analysis platform called Chatterbox Analytics Ltd. The aim is to gauge the public expressions of democracy during this period in terms of sentiment. This will subsequently be discussed in more detail.

The popularity and usage of the social media sites have grown rapidly during the last couple of years as can be seen from *Figure 2* below. Millions of people from across the world are using these sites and expressing their feelings and opinions on a vast range of subjects. Twitter, as the second biggest social networking site, has grown from its infancy in 2006 to more than 300 million users in 2011. In July 2012 - the final month within which data were collected for this study - Twitter surpassed the half a billion mark, with approximately 517 million Twitter user profiles (Semiocast, 2012).

According to the French Internet brand management company, Semiocast (2011), about 39 percent of the users are English - therefore about 201 630 000. Most of the users are situated in the US (141.8 million) as can be seen from *Figure 3* below. The 517 million users create approximately 400 million Tweets every day (Farber, 2012). These Tweets contain 140 characters of personalised information each.

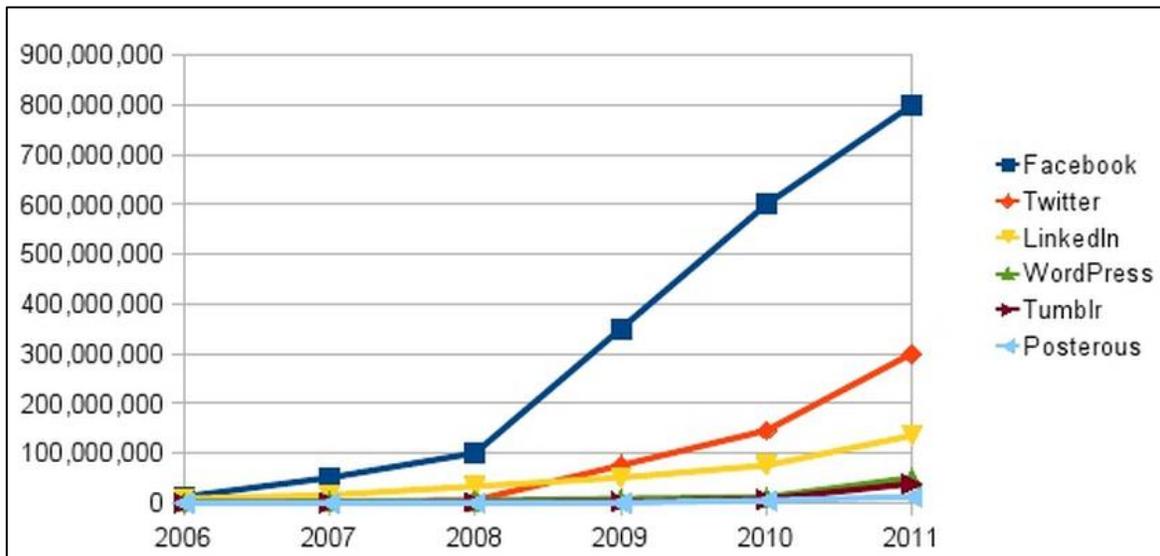


Figure 2: Social Media Growth – 2006-2011 - Source: (DStevenWhite, 2011).

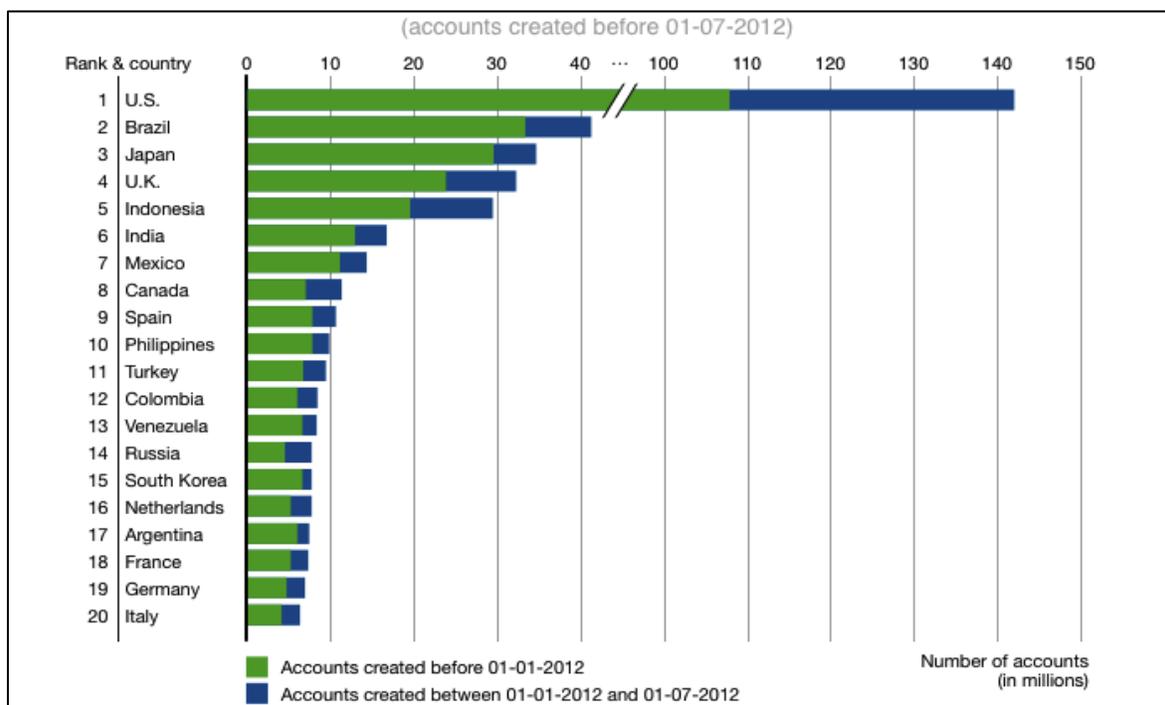


Figure 3: Top 20 countries in terms of Twitter accounts, 2012 - Source: (Semiocast, 2012).

For the period 1 May to 31 July 2012, according to Tweets gathered for this research thesis, there were approximately 100 000 Tweets every day relating to democracy⁶. This study can however only access one

⁶ Relating in this sense refers to a Tweet containing the word democracy in any of its colloquial forms specified in the conceptualisation section of this chapter.

percent of the generated Tweets, but this will be discussed in more detail in the limitations and delimitations section of this chapter. Each user profile does not necessarily represent one person as one can have multiple accounts and other entities such as organisations can also have accounts. If however only half of the users represent a real person, the expressions on Twitter are already coming from 3.7 percent of the world's population (Census, 2012).

This Tweet information contains expressions of millions of people that could be of substantial value for assessing the public's opinion on specific subjects (O'Conner, 2010). The Tweet information itself is publically available and easily analysed with automated computerised methods. Twitter information therefore gives one an ideal case for developing a method for analysing public expression on democracy from a social media site. This study aims to develop this methodology that will utilise automated content and sentiment analysis software in order to collect and analyse specific Tweets relating to the subject matter of democracy.

1.1.4.2 Why measure expressions on democracy?

The political subject matter that is studied in this study is democracy. The author would like to gauge public expression on democracy by looking at the 104 037 Tweets from Twitter discussed above. Within democratic states, the will of the people is important and this "will" needs to be represented if a government wants legitimacy (Barber, 1996: 349). The public will can be examined by looking at the political sphere and opinions that are represented within the public sphere. Scholars such as Habermas (1989) and Arendt (1963) have discussed the public sphere and how this sphere is constantly changing. Habermas (1989) has looked specifically at the influence of the media on political opinion in the public sphere.

Using the ideas from Habermas (1989), a number of scholars such as Papacharissi (2009) have examined social media sites, and the Internet as a whole, as new types of public spheres where public opinion is being generated and discussed. The creator of Twitter, Jack Dorsey, stated at the 2012 Digital Life Design Conference in Munich, that Twitter is a public medium that hosts public conversation in a way that is simple enough for anyone to use (VentureBeat, 2012). If this statement is an accurate representation of Twitter, the discussions on Twitter might be comparable to a type of Habermasian public sphere and the opinions generated here can be of the utmost value in gauging public opinion. This thesis will treat Twitter as such a public sphere.

As discussed above, the descriptive section of this research will examine the potential erosion of democracy that has been discussed by authors such as van Beek (2012: 12). It will be shown in the second chapter of this thesis that the idea of democracy is constantly evolving and that this evolution in many

cases throughout history has been driven by sentiments expressed by the general populace. If one can therefore show a negative sentiment that is present within a section of the populace regarding democracy; this might support ideas as to the potential erosion of democracy.

1.2 THEORETICAL FRAMEWORK

There are a number of theoretical underpinnings that are central to this study. Firstly, this study is positioned within the school of political culture. More specifically this study falls within the study of the political culture of democracy. This field has developed out of the seminal work of Gabriel Almond and Sidney Verba (1963), *The Civic Culture*. This work was based on the idea of Lipset (1959) on the link between democratic consolidation and popular support for democratic institutions. In Almond and Verba's book, political attitudes and democracy are examined within five nations, namely the USA, UK, Germany, Italy and Mexico. The purpose of their study was to quantify the attitudes and ideals that compose the civic culture of certain democratic systems (Almond and Verba, 1963:11). The conclusion Almond and Verba came to was that individual level support for democracy and a strong supportive political culture contributed to the consolidation of democracy (Almond and Verba, 1963:168-9). This study gave birth to countless further studies⁷ that examined what effect attitudes towards democracy and democratic institutions within countries have on the functioning and consolidation of the democracy within a country. A section on the development of studies examining mass attitudes towards democracy can be found in Chapter Three of this thesis.

There are studies⁸ that do not adhere as strongly to the idea that a positive political culture towards democracy will lead to democratic consolidation. This thesis, however, does not examine this question, but rather accepts that individual attitudes on democracy do in fact affect democracy and that positive mass attitudes towards democracy positively affect democratic consolidation - and vice versa, that negative attitudes could erode democracy and affect democratic consolidation negatively.

As stated previously, this thesis also examines Twitter as a form of Habermasian (1989) public sphere. This idea is discussed in more detail in the conceptualisation section of this chapter. It should however be noted that the expressions on democracy from Twitter will be examined as discourse that might lead to the formation of public opinion.

This study is grounded within the theories of democracy and democratic governance. In the second chapter the erratic developmental path of democracy will be examined. It will be shown that democracy has evolved throughout the years and that public opinion on democracy has had an influence on this

⁷ Studies such as: Lipset, 1994; Diamond, Linz and Lipset, 1995; Linz and Stepan, 1996; Diamond *et al.*, 1997.

⁸ Studies such as: Dahl, 1972; Rustow, 1970; Lijphart, 1977.

evolution (Crozier, Huntington and Watunaki, 1975; Dahl, 1998; Dalton, 2004). The ideas of authors such as Welzel (2002); Welzel and Inglehart (2008) and van Beek (2012) are also examined in order to highlight the potential of a global erosion of confidence in democracy.

Finally, the methodological framework for this thesis is developed by examining a number of traditional and new methods of measuring expressions on democracy. Firstly, a number of traditional research projects, such as the World Values Survey, are examined. Following this, novel methods⁹ of measuring public expressions from social media sites - especially Twitter - are examined. The study by O'Connor *et al.* (2010), *From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series*, is used as a broad framework for the methodology that is then developed in order to measure public expressions on democracy from Twitter.

1.3 RESEARCH METHODOLOGY

1.3.1 Research Design

This study is both a methodological and a descriptive study as it includes a main methodological research question, but also supplementary descriptive research questions. It follows quantitative methods, seeing that automated computerised and statistically verifiable methods will be used to classify Tweets according to numerical sentiment scores. Statistically verifiable methods will also be used in measuring the accuracy of these methods. This quantitative study will use automated content analysis methods, adapted from methods used in the O'Connor (2010) and the Hopkins and King (2010) studies. Furthermore, a platform called Chatterbox Analytics will be used for the actual data collection and analysis.

1.3.1.1 Research Question(s)

The main methodological question that this research aims to address is:

Main Research Question: *To what extent can one measure public expressions on democracy from interest articulated sentiments gathered from a sample of "Tweets", processed with content analysis and sentiment analysis software?*

In order to answer this main question two supplementary questions are utilised. This key question is firstly supplemented by the descriptive question of:

⁹ This section includes works such as: Pang and Lee (2008); Asur and Huberman's (2010); Bollen *et al.* (2011); Lindsay (2008); Tumasjan *et al.* (2010); Gayo-Avello, 2012; Ritterman *et al.* (2009) and Hopkins and King (2010).

Supplementary Question 1: *What are public expressions on democracy between 1 May and 31 July 2012, according to interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software?*

The goal of this supplementary question is to provide a test case for implementing the methodology and analysing its success. A second supplementary research question is also needed. This final question reads:

Supplementary Question 2: *How accurate is the content analysis and sentiment analysis software used in measuring public expressions on democracy between 1 May and 31 July 2012 from Twitter?*

This final supplementary question aims to evaluate the accuracy of the measurement instrument. In this thesis the measurement that will be used is an online platform called Chatterbox Analytics, but this will be subsequently discussed in more detail.

1.3.1.2 Hypothesis

A hypothesis that this study will test, with regards to the first supplementary descriptive question, reads:

The public expression of democracy between 1 May and 31 July 2012, according to interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis, and sentiment analysis software is more negative than positive.

The researcher believes that this quantitative study can integrate the available technologies in automated content and sentiment analysis in order to effectively measure public expressions on democracy from Twitter. Furthermore, this quantitative study will provide scientific falsifiable and representative results as it is expanding on available/tested methods. The thesis is novel as expressions of democracy have never (as far as the researcher is aware) been studied in this manner. It is hoped that this method might supplement traditional methods of measuring expressions on democracy, such as survey data.

The study follows an inductive approach examining the individual Tweets of micro bloggers on Twitter, categorising the sentiments found within their Tweets and then revealing the public expression on democracy.

If one looks at three principles of classification that Babbie and Mouton (2001:76) recommend for a research design one can see that: this thesis is an empirical study following a primary, textual data design as new primary data will be collected from the real world; the data are textual in that Tweets will be analysed, processed and documented in order to gauge the public expressions on democracy; and opinions of micro bloggers will be analysed.

1.3.2 Units of Analysis

The objects that are analysed in this study are the cultural objects, Twitter micro blogs (Tweets). The units of analysis are therefore Tweets. Each Tweet has various characteristics that will be explained in the conceptualisation section of this chapter. The Tweets relating to democracy for the chosen period will be analysed, processed and documented. These analysed and processed Tweets will then represent interest articulated expressions on democracy for the chosen population of Tweets.

1.3.3 Levels of Analysis

In order to examine public expressions on democracy from Twitter, this study will make use of macro-level measures. This is the case, seeing that expressions on democracy on Twitter for all Twitter users across the globe will be analysed.

1.3.4 Time Dimension

This exploratory study will utilise a cross-sectional approach. A snapshot of the constantly created Tweet data has been collected. This snapshot lies between 1 May 2012 and 31 July 2012 (92 days). This recent period has been chosen in order to represent a relevant and up-to-date section of data. Twitter is growing quite rapidly, as has been shown in *Figure 2*. Using recent data therefore not only keeps this study up to date, but also allows one to work with a much bigger dataset. The researcher is also in possession of a dataset of Tweets for almost the whole of 2010. This dataset contains about 8000 Tweets relating to democracy. The recent three-month period selected for this research, in comparison, contains 104 037 relevant Tweets which provides one with more than 95 000 extra data points to work with.

1.3.5 Data Collection and Analysis

In order to answer the research questions this study makes use of naturally occurring data in the form of Tweets that exist in the online, public realm. A dataset of 104 037 Tweets has been collected using the Chatterbox Analytics platform (Chatterbox, 2012). This platform automatically gathers Tweets related to a user-specified topic, harnessing the Twitter stream application programming interface (API). Without going into too much technical detail, the Twitter stream API allows one to access ‘streams’ of public data flowing through Twitter. This is done by accessing what Twitter calls the public *firehose*. Twitter however limits public access to the firehose stream to the first one percent of the available Tweets at any moment. It is therefore possible to view and analyse one percent of the publicly generated Tweets every ‘streaming second’ into a database (Twitter Developers, 2012). This is commonly called Twitter data mining. It is necessary to mention that there are companies such as DataSift, Collecta, CrowdEye, Tap11 and of course Twitter, that have access to the full firehose stream. DataSift can even provide a full historic

set of Tweets dating back to January 2010 (DataSift, 2012). The one percent sample publicly available is however sufficient for this introductory methodological and descriptive study.

This massive collected Tweet dataset is processed using a number of automated computerised methods by Chatterbox. These methods will be discussed in more detail in the operationalisation section of this thesis. The methods include content analysis, sentiment analysis, and key word extraction and measurement. The goal of the analysis is to determine whether and to what extent the Twitter dataset can answer the research question.

Figure 4 below provides an introductory diagram of how the data analysis method will function. As discussed above, one percent of all Tweets for the period are available. From this one percent this paper will only focus on the English Tweets, which pertain to democracy. These Tweets will be split into those that contain positive sentiments on democracy and those that contain negative sentiment. A neutral section will also finally be added in order to represent tweets that have been collected, but that do not contain direct sentiment on democracy.

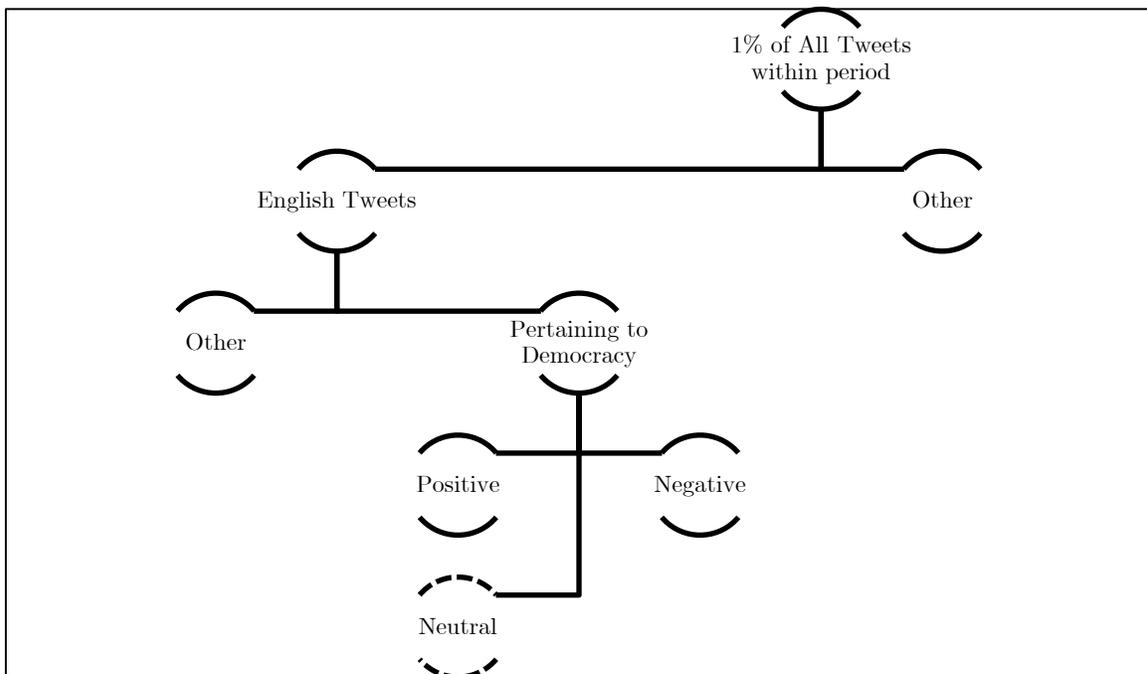


Figure 4: Levels of classification

1.4 CONCEPTUALISATION OF CORE TERMINOLOGY

There are a number of explicit and implicit concepts that will be used in this study that warrant further clarification. In this section the concepts used in this study will therefore be clearly defined and clarified.

1.4.1 To what 'extent'

The main methodological research question of this study asks: “*to what extent can one measure public expressions on democracy...*”. The explicit concept of ‘extent’ warrants further conceptualisation. The extent to which one can measure public expressions on democracy using the novel methods proposed within this thesis will be tested in two ways. Firstly, a case study in the form of the cross-sectional analysis discussed above will be conducted. This case study will aim to answer the first supplementary descriptive question and prove or disprove the hypothesis expressed previously.

Secondly, a succinct subjective analysis of the accuracy of the measurement instrument will be conducted. This analysis will aim to answer the second supplementary research question.

The aim is to firstly show the practical applicability of the method that has been developed. Next, the accuracy and validity of this method is tested. Both these tests add up in order to show the extent to which one can measure public expressions on democracy from Twitter within the chosen timeframe with the automated computerised methods.

1.4.2 Public Expression

The next explicit concept is that of public expression. Public expression in this study will refer to the sentiment within the statements (Tweets) that have been collected for the chosen period. The aim of this study is not to measure the public expression of the whole world on democracy. The goal is only to show whether the collected 104 037 Tweets regarding democracy are more negative or more positive.

The goal is therefore to aggregate individual expressions, from Twitter, on democracy, in order to gain a “picture” of a specific population’s attitude on the subject of democracy. In this study focus will be placed on interest articulation through Tweets and this will provide a broad public expression of sentiments rather than a classical aggregated public opinion.

1.4.3 Democracy

It should also be noted that this study will view democracy as a broad encompassing term. In this introductory study attention will not be placed on the multitude of meanings that the term democracy might refer to. All public expressions on democracy will therefore be grouped. In order to clarify what is meant by a broad encompassing view of democracy a few examples will suffice. This researcher has found that expressions on democracy from Twitter mostly fall within one of three categories. Firstly, there are expressions referring to the broad idea of democracy such as:

#Democracy always produces unpopular candidates and forces you to choose one. Don't submit to democracy if you're against this notion.

Then there are the expressions that refer to the state of democracy within another country:

RT @ForeignAffairs Bad judges! #Pakistan's supreme court is undermining democracy

Finally, there are expressions that refer to the state of democracy within one's own country:

RT @sunshineejc: Why do Americans Allow corporations to co-opt our democracy by granting them the rights of legal personhood and defining money as speech #p2

These are only three classifications that can be placed upon the expressions on democracy. There are many other classification methods that could be used in future studies. For example, one might classify the expressions into classic models of democracy. This introductory study, however, has grouped all the expressions on democracy into a broad encompassing view of democracy.

The word democracy itself will be used as a keyword for the automated data collection, but this will be explained further in the operationalisation section of this thesis. The term democracy is however not only used in its standard form online, but also in a number of colloquial forms such as: *demockerycy*, *deMockracy*, *democratees*, *democrabeep*, *democrackcy*, *democrack*, *democrap*, *democrapsy*, *democrASS*, *D3m0cr@cy*, *plutocracy*, *demo-crazy*, etc. (Urban Dictionary, 2012). In order to collect as many possible relevant expressions on democracy a number of these colloquial forms of the word have been added as keywords.

1.4.4 Interest Articulated Sentiments

In this study interest articulated sentiments will refer to the positive or negative sentiment that can be measured from publicly expressed Tweets discussed above.

1.4.5 Tweet

Another explicit concept is Tweets or Tweet objects. In this thesis, Tweets are the objects of study and refer to all the micro blog entries that appear on Twitter. A Tweet is a 140 character message that is posted by a user on the user's Twitter profile. These Tweets are public, meaning that anyone is able to see what the user is writing. From the Twitter *firehose* one is able to download one percent of all the created Tweet objects which each contain quite a large amount of information. A Tweet contains more than just the Tweet text. *Figure 5* is a table that highlights relevant information that can be gained from a Tweet object.

Information on the user	Information on the Tweet	Other information
Username + Screen name (online nickname)	Tweet ID	Browser and operating system used for Tweet
Picture of the user can be acquired in most cases	Is it a reply and if so to whom and with regards to what discussion	IP address of device used for Tweeting
Location – can include all or some of: country/region/city/neighbourhood (if geotags are activated)	Tweet text 140 characters max	type of device used for Tweet can be determined in some cases
Date account was created	Number of re-Tweets	
Language selected by user	Date and time created and time zone	
Number of followers user has	How many people have marked the Tweet as a favourite	

Figure 5: Information gained from a Tweet object - Source: (Meshak, 2012)

1.4.6 Content and Sentiment Analysis Software

Content analysis, also called textual analysis, refers to a methodology for studying the content of communication (Babbie, 2010:333). In this study the content of communications on the social media site, Twitter, will be studied. The content that will be collected and analysed with the Chatterbox platform will also undergo sentiment analysis. Sentiment analysis aims to determine the contextual polarity of a certain text. The text is automatically analysed using various machine learning techniques. Each word is analysed within a sentence in order to determine the overall sentiment of the sentence. The machine learning techniques used by Chatterbox will be seen as a black box for this study. The reason for this is that it would fall outside the scope of this study and the ability of this author to discuss the field of machine learning¹⁰.

The Chatterbox platform that is used in executing these methods will be discussed in Chapter Three, but will greatly be treated only as a tool (in the same way that Microsoft Excel or the Statistical Package for

¹⁰ Studies such as Pang *et al.* (2002) contain more in depth information on machine learning and sentiment analysis software.

the Social Sciences is used as a tool) and therefore a black box with which the data are analysed¹¹. The validity and effectiveness of these types of programmes is discussed in literature such as: Zhao *et al* (2011), Bollen *et al* (2009), O’Conner *et al* (2010), Hopkins and King (2010) and Pang and Lee (2008).

1.4.7 The Public Sphere

The implicit concept of the public sphere is central to this thesis. The public sphere refers to a place, be it physical or virtual, where citizens gather and exchange views on public affairs. These spheres are ‘places’ where citizens can meet and discuss, debate and deliberate on current affairs which eventually leads to the formation of public opinion (Castells, 2008). The contemporary view of the public sphere is mainly based on the works of Jürgen Habermas (1989). This German sociologist defines the public sphere as a “network for communicating information and points of view ... the streams of communication are, in the process, filtered and synthesized in such a way that they coalesce into bundles of topically specified public opinions (Habermas, 1996:360).”

The public sphere is a normative idea as it is an ideal for good and accountable governance. The rationale is that if a state has a well-functioning and democratic public sphere, then citizens would in a participatory fashion assert their influence on political decisions and hold governments responsible (CommGAP, 2012). According to Habermas (1976), a crisis of legitimacy could develop if the citizens feel that they are not being recognised in the institutions of society. This in turn could lead to a crisis of authority and culminate in a redefinition of power relationships within a state (Sassen, 2006; Castells, 2008:80).

Throughout the development of democracy there have been versions of public spheres promoting political accountability. Examples are the *agora* of Greek City States, where citizens gathered and discussed public affairs. Tribal gatherings and certain church congregations throughout history could also be seen as a form of public sphere. In European monarchies the royal court was a public sphere but it was dominated by the monarch. During the 17th and 18th centuries, salons (France), coffeehouses (England) and table societies (Germany) could be seen as public spheres that were limited to the aristocracy. As technology developed and mass newspapers became distributed the discussions from these coffeehouses and salons became truly public (CommGAP, 2012; Castells, 2008).

Today, with the rapid development of technology, there are those¹² who argue that new forms of public spheres are developing with the help of ICTs. People do not have to gather physically in order to discuss

¹¹ The MIH Media Lab is currently also working on the development of these types of programmes. They can be contacted at: info@ml.sun.ac.za

public affairs anymore. The development of new technology based public spheres will be discussed in more detail in Chapter Four. What is important to note is that this study will treat Twitter as a version of an online public sphere. In this study the public expressions on Twitter will therefore be examined as a form of public debate and discussion by global citizens that could lead to public opinion formation.

1.4.8 Twitter

The next implicit concept is Twitter. Twitter is an online social networking service that was created by Jack Dorsey in 2006. It allows users to share short messages (Tweets) publicly on its online platform. Twitter has grown rapidly since its inception and was estimated to have over 500 million users by March 2012 (Dugan, 2012). This study will use expressions on Twitter, in the form of Tweets, as its units of analysis.

1.4.9 Chatterbox Analytics Platform

The Chatterbox Analytics Ltd. platform is the chosen platform for automatically collecting and analysing the Tweet data. This type of platform is traditionally aimed at market research, allowing brands to see what their customer base is saying about a brand on social media sites. This allows the specific brand to respond to customer complaints and gauge public sentiment on their specific brand instantly. In this study democracy is treated as such a brand and analysed with the Chatterbox Analytics platform in the same way.

1.4.10 Big Data

Big Data is a popular term referring to the dynamic or activity of a new generation technologies and architectures that can economically extract value from massive volumes of data at a high velocity. As storage is becoming less expensive and processing power increases the analysis of Big Data is becoming more feasible. More rich data for analysis are also being created from social media solutions such as Facebook and Twitter. This information is user-generated and flows in a near continuous stream (Grantz and Reinsel, 2011).

In 2010 the annual amount of data created digitally exceeded one zettabyte¹³ for the first time. In 2011 it is estimated that this number grew to 1.8 zettabytes. Around 70 percent of this information is created by individuals and it is estimated that the amount of data created doubles every two years. This is according to a five-year study of the digital universe, by the IDC (Grantz and Reinsel, 2011). Big Data is the activity of analysing this vast amount of created data in order to gain valuable insights from it.

¹² Such as Castells (2008), Papacharissi (2002), Downey and Fenton (2003), Dahlberg (2001), Gruicic (2011), Trenz (2009)

¹³ Zettabyte = 1 trillion gigabytes.

1.5 LIMITATIONS AND DELIMITATIONS OF THE STUDY

Important limitations to this study flows out of the delimitation of focussing only on English Tweets that are collected from the Twitter firehose with the Chatterbox platform. As stated above in the conceptualisation of public expression section, the aim of this study is not to measure the public expression of the whole world on democracy. The goal is only to show whether the collected 104 037 Tweets regarding democracy are more negative or more positive. This will be determined by the measurement instrument, namely the Chatterbox platform.

This study will treat this novel platform used for data analysis broadly as a black box. Much in the same way that a statistician does not describe the inner workings of the statistical package that they use, this study will not examine the inner functioning of the Chatterbox platform that it uses for analyses.

A limitation to this study is that the Chatterbox platform was still in its beta testing phase when this study was conducted. This meant that the platform was still being developed while the Tweets were collected and analysed for this study. The developmental nature of the platform in turn meant that there were a number of small gaps (a couple of hours of Tweets) that were not collected while Chatterbox updated their programme. Gaps exist between 11:52 on 31 May and 14:44 on 1 June, also between 02:51 and 04:54 on 15 June, and between 23:52 on 30 June and 02:49 on 1 July. Also, the sentiment classification method that Chatterbox uses was updated on 11 July 2012. Seeing that Chatterbox keeps details on its sentiment classification method private and the fact that this study views the classification method as a black box, the specific change in the method is not described here. It should however be noted that the change in the method of classification is a limitation of this study.

For the introductory methodological research question of this study, the fact that the measurement instrument is still in its *beta stages* does not pose a concern. This study does show the inaccuracies within the measurement instrument and advises on future improvements and other methods that might be used to improve the instrument's capability to measure expressions on democracy from Twitter. With regards to the descriptive question on the expressions on democracy, it should be stated that the inaccuracy and developmental nature of the measurement instrument does pose limitations to the conclusions that can be drawn from the results. The results from this analysis only represent the 104 037 Tweets and can only provide a rough estimation (found in Chapter Four) of the level of positive and negative sentiment within this population. This is however an introductory study that aims to open the field of measuring public expressions from social media sites and future studies might improve on the measurement accuracy and draw larger or more specific conclusions on specific sections of society.

In August 2012, after the data for this study were collected, Chatterbox again changed its sentiment classification method. This change meant that Tweets were now classified not only as positive or negative, but a neutral (weak sentiment) category was also added. Tweets with a sentiment score falling between -0.25 and 0.25 were now classified as having a weak sentiment. As a delimitation this weak or neutral sentiment category was not included in the original measurement by Chatterbox¹⁴. This researcher however decided to add a neutral section in the final analysis of the accuracy of the measurement instrument¹⁵.

As explained previously, Twitter only allows individuals access to one percent of public Tweets. The measurement instrument (Chatterbox) only has access to this one percent. Chatterbox at present also only looks at English Tweets. This study is therefore limited to only English Tweets within one percent of global public Tweets. For the period of 1 May to 31 July this provides this study with 104 037 Tweets - which is sufficient to develop and test the methodology.

An important delimitation of this study is that it does not aim to link expressions on Twitter to specific users. The Tweets are seen as the units of analysis. Future studies may use some of the extra criteria that one can gather from Tweet objects in order to view Tweets, for example, according to the location from which it originated. In this introductory study aspects regarding the person or entity that has created the Tweet will however be one such delimitation.

Another delimitation of this study regards the demographics of the users on Twitter. This researcher is aware of the following: that Twitter users at present are mostly from developed countries; are mostly below the age of 35; that there are more women than men; that the Internet penetration rate in developing countries limits their access to Twitter; and various other demographical factors which limit the users on social media sites in general (Skelton, 2012). The methodological research question of this study, however, only looks at the possibility and accuracy of measuring expressions on democracy from Twitter and the conclusions from the descriptive question is limited to the 104 037 Tweets. The demographic factors of the users on Twitter therefore do not affect this study.

There are debates as to whether new media such as Twitter constitutes a platform for a public sphere. This debate will be examined in more detail in Chapter Three. However, the author will delimit this study and treat Twitter as a public sphere. Expressions on democracy will therefore be viewed as discussion and debate that might lead to the formation of public opinion.

¹⁴ With regards to the descriptive supplementary research question 1.

¹⁵ With regards to the supplementary research question 2.

Finally, this thesis includes delimitation as to what is meant by the term democracy. There are countless versions and theories explaining and describing the phenomenon of democracy. In this study, however, the idea of democracy will, as explained in the conceptualisation section, refer to quite a broad definition of democracy.

1.6 STRUCTURE AND CHAPTER OUTLINE

Following on from this introduction to the topic and intentions for the research, Chapter Two will focus on the turbulent development of democracy and contemporary views on democracy. It will be shown that democracy has adapted and evolved throughout the ages of its existence in order to stay a relevant system of governance. A potential negative contemporary view on democracy will be highlighted in order to show that the system of democracy today might again be calling for adaptation. Chapter Three will illustrate the importance of public opinion on democracy. Furthermore, it shows how these opinions were traditionally gauged; and examine novel methods of gauging opinions on democracy. Chapter Four takes the form of an operationalization and measurement chapter. In this chapter, the novel method for measuring public expressions on democracy from Twitter will be developed and the actual measurement on the proposed population of Tweets will be conducted. Finally, in Chapter Five the findings of the measurement from Chapter Four will be examined. The relevance of these findings will be discussed and recommendations for future studies, based on these findings, will be made.

1.7 CONCLUSION

This chapter sought to provide a clear outline of the intentions for this thesis as well as introduce the methodology, structure and theory that will be used to inform this analysis. Attention has been given to relevant literature that attempts to contextualise this introductory chapter in terms of the existing scholarship on the topic as well as with regards to applicable contemporary developments. In the subsequent chapter an overview of the development of democracy followed by a contemporary view on democracy can be found.

Chapter 2: An Examination of Democracy

2.1 INTRODUCTION

The aim of this chapter is to introduce the term democracy, which is the system of governance that this thesis aims to examine. Firstly the term democracy will be briefly discussed. Following this section, a historic overview on the development of democracy can be found. It will be shown that democracy has evolved and developed through a disjointed and rocky path. It will also be shown that negative sentiment has followed the development of democracy and in a sense guided its evolution. Next a contemporary view of democracy will be explored in order to show that contemporary perceptions on democracy might again be indicating an oncoming evolutionary shift in the functioning of democracy. In the final section of this chapter a rising negative view on democracy is examined in a more detail. This researcher is of the opinion that this critical view on democracy could indicate that democracy again is at a precipice of development or adaptation. The goal of this study is to develop novel methods of measuring if this is indeed the case.

2.2 UNDERSTANDING THE TERM DEMOCRACY

The English word ‘democracy’ was first used in the 16th century and was adapted from the French word *démocratie*. The etymology of the French word is however derived, via Late Latin, from the Greek word *demokratia*, which can be broken up into *demos* (the people) and *kratos* (rule/strength/power) (Miller *et al.*, 1991; Oxford Dictionary, 2012). If one is to measure public expression on democracy, as this thesis intends to do, it is important to understand what these expressions are referring to. One should therefore understand what ‘democracy’ is. This is however not only a case of providing a standardised definition of democracy, as there are many ways in which this word/phenomenon can be interpreted. Throughout history one finds almost as many real world examples of regimes that claim to be democracies as there are alternative interpretations (Held, 2006:1). The term has developed and evolved throughout the history of political thought and refers to vastly differing political systems.

In order to better understand the term democracy, this study will look at the historic development of the political system of governance. Following the example of a number of authors¹⁶ who studied the development of democracy in the past, this thesis aims to approach the development of this system of

¹⁶ Dahl (1998); Held (1987, 1996, 2006); Harrison (1993) for example.

governance from a chronological perspective. The aim of this chronological structure of examination is twofold. Firstly I aim to highlight the main developments in the historical development of democracy and democratic theory. Emphasis will be placed on certain key aspects in the development of democracy. These aspects will show the developments that are paramount to specific contemporary understandings of democracy. Factors such as the ideals of liberty and the idea of representativeness will therefore form key sections in the subsequent discussion.

The second aim of the chronological structure of the examination of democracy is to show how sentiments about the term democracy have oscillated throughout history. It will be shown that since the origins of democracy in ancient Athens, negative and positive sentiments can be found. The development of the political system of democracy seems, as Dahl (1998:9) states in his book, *On Democracy*, not to be a tranquil history with a smooth progression to the modern form of democracy, but a bumpy development. It will be shown that negative sentiments have followed the progression of democracy since ancient times. The aim of this revelation is to show that the negative sentiment that one might find in expressions on democracy, in works such as this thesis, is not new and has in fact been part of the development of the term. Many of these historical negative sentiments have in fact led to the evolution, progression and development of democracy.

2.3 THE TURBULENT DEVELOPMENT OF DEMOCRACY

2.3.1 The Formative Years of Democracy: Athens 508 BCE to the Renaissance (1700+-)

There are still debates over the exact origin of the first democracy, but most scholars believe that Cleisthenes instituted the first democracy in Athens in the year 508/7 BCE (Hansen, 1992: 25-27; Raaflaub, 2008: 2). This first democracy was the result of continuous change over many generations, and the democracy that followed was also not stagnant as it continued evolving. The Athenian democracy followed some of the ideals of democracy¹⁷ that are still adhered to today, such as equality among citizens¹⁸, respect for law and justice, and liberty. This type of democracy, however, also differed greatly from the type of democracy one sees in the contemporary world (Held, 2006:12-13). The closest link to a contemporary form of democracy is to that of direct democracy (Dryzek and Dunleavy, 2009: 19).

Sentiments on democracy from the Athenian times vary. One famous general of Athens, Pericles¹⁹, was specifically positive in his view of democracy. However, throughout these early years of democracy there were also many critics of this political system. Even in the Greek times there have been warnings about

¹⁷ The basis of these Athenian ideals were: Isnomia – Political Rights; Isegoria – Free Speech; Isocratia – Political Power; Iogonia – Birth Right; Isopolitea – Legal Rights (Dryzek & Dunleavy, 2009: 19).

¹⁸ Citizens were however mostly only free adult males of strictly Athenian decent (Held, 2006: 13).

¹⁹ As found in Pericles' Funeral Oration, recorded by Thucydides (Palmer, 1992).

democracy from philosophers such as Plato (1974)²⁰ and Aristotle (1981)²¹. They believed that in a system of pure democracy, where the *demos* have supreme power, demagogues would tend to emerge and supersede the law, thereby turning the democracy into a form of despotism (Diamond, 1999:2-3; Harrison, 1993: 33). It was therefore believed that democracy would lead to some form of rabble politics that would ultimately culminate in an ochlocracy²².

In time the growth of Roman power restricted the Athenian democracy. This first documented Athenian democracy finally came to an end when Octavian became Emperor Augustus and the Roman Empire was born in 27 BCE (Starr, 1952: 4). During the Roman era there was a rise in a form of aristocratic republicanism in which elections were held through various assemblies (Dryzek and Dunleavy, 2009: 19).

After the Roman era the emergence of Christianity led to the formation of a faith centric political system, and shifted man towards *homo credens*. In this system, the rationality for political action shifted from support for the *polis*, as in the Athenian democracy, to a theological framework in which the goal of citizens was to submit to the will of God (Pocock, 1975: 84). During these ‘Middle Ages’ there was little extensive reflection on the nature of the democratic polis (Poggi, 1978).

Nevertheless, in the 11th century, a number of Italian city-republics²³ emerged that challenged the prevailing assumption that government should be a God-given form of lordship (Skinner, 1992: 57). These city-republics were a primitive form of republicanism, in which various councils, headed by elected *podestà*²⁴ governed. These states did not come close to the democratic values that were seen in Athenian democracy, but it was the catalyst that inspired future Europeans and Americans to question the absolutist system of government that was in place (Skinner, 1992: 58-60).

During the Middle Ages the idea of democracy was still, however, being broadly shunned. In the mid-13th century, when Aristotle’s *The Politics* reappeared in Europe, the negative view of democracy grew even further. *The Politics* was used by influential philosophers such as Thomas Aquinas to infer a pejorative connotation to the term democracy (Held, 2006: 33). This meant that the term democracy, in this era, was used in reference to “the politics of the rabble”; a government that only protected the poor, rather than the public interest and a system in which the *common people* could easily become tyrannical. This critical view of democracy only began to change with the birth and development of the liberal tradition of political thought throughout the renaissance. Before this birth of this liberal tradition one

²⁰ In works such as *The Republic*

²¹ In works such as *The Politics*

²² An ochlocracy refers to a type of mob rule wherein government is dominated by the will of a *fickle crowd* (Harrison, 1993: 33).

²³ Florence, Padua, Pisa, Milan and Siena (Held, 2006:32).

²⁴ Leaders with supreme executive and judicial power (Held, 2006:32).

notes a shift away from the church centric view of political life (Held, 2006: 33). This negative sentiment towards democracy, seen during its early years is perfectly expressed by the following quote by Finer (1997:46-47):

“The Forum polity is comparatively rare in the history of government, where the Palace polity and its variants are overwhelmingly the most common type. Only in the last two centuries has the Forum polity become widespread. Before then its appearance was, on the whole, limited to the Greek poleis, the Roman Republic, and the mediaeval European city-states. Furthermore, most of them for most of the time exhibited the worst pathological features of this kind of polity. For rhetoric read demagogy, for persuasion read corruption, pressure, intimidation, and falsification of the vote. For meetings and assemblies, read tumult and riot. For mature deliberation through a set of revising institutions, read instead self-division, inconstancy, slowness, and legislative and administrative stultification. And for elections read factional plots and intrigues. These features were the ones characteristically associated with the Forum polity in Europe down to very recent times. They were what gave the term ‘Republic’ a bad name, but made ‘Democracy’ an object of sheer horror.”

There were two main figures in history that, according to Paul Sigmund (1962), were paramount in the shift away from the antecedent tradition of the *res publica Christiana*, namely the dominance of Christianity over state governance. These figures are Marsilius of Padua²⁵ and Niccolò Machiavelli²⁶. Both these men were condemned by the ecclesiastical rulers of the era as their theories moved away from a church centric state towards a citizen centric state.

Marsilius of Padua (1980) gave profound articulation to the developmental republicanism approach to state governance, which was somewhat similar to the Athenian democracy. This approach placed focus on the intrinsic values of political participation that would develop the citizenry of a state (Held, 2006:35-37). Political participation according to this approach was therefore a good in itself. In works such as *Vindication of the Rights of Women*, by Mary Wollstonecraft (1792), one can witness a further development of this developmental republicanism approach.

The work of Machiavelli (1975, 1983) is more commonly associated with protective republicanism (Held, 2006: 37). This form of republicanism focussed closely on the instrumental value of political participation. Machiavelli, as Plato before him, conceived “the generality of man” as lazy, self-seeking and incapable of good unless forced in some way. The function of the republic was therefore to promote

²⁵ Mainly through his work, the *Defensor Pacis* (Defender of Peace) (Marsilius of Padua, 1980).

²⁶ Mainly through his works, *The Prince* and *The Discourses* (Machiavelli, 1975; 1983).

civic virtue that would force citizens to act appropriately (Machiavelli, 1983: 256-257). This approach promotes political participation as an essential condition of personal liberty and a tool for citizens to protect themselves from the domination of others. Protective republicanism was further elaborated by authors such as Montesquieu²⁷ (1952) and Madison²⁸ (1966), which will be discussed in a subsequent section (Held, 2006:37-41).

Both the works of Machiavelli and Marsilius of Padua contained some of the germs of a theory of liberal democracy. They did not propose a full liberal democracy, but it was a step closer to such a state. Machiavelli even proposed a more liberal view of governance, in which not only the elite should participate, but in which artisans and small traders should also be allowed to participate. Furthermore, the idea of self-determination as a form of self-protection and the break with the *res publica Christiana* was a fundamental moment in the development of political thought. This can be seen as a precursor to the development of a more liberal form of democracy during the renaissance (Held, 2006: 43).

2.3.2 The Renaissance and the Birth of Liberalism

The fears of the ancient philosophers regarding the tyranny of the majority that might occur in a democracy have remained present through the ages and are still present in contemporary society. In attempts to solve this problem with pure democracy, one therefore notes that democracy has been tempered with various rules and institutions throughout its development. Intellectuals such as Locke (2005), Hobbes (2008), Rousseau (2008), Montesquieu (1952) and the American Federalists (Madison, 1966) played a big role in this temperament of democracy. This moderation of democracy is done in order to protect the individual liberty of citizens and will be referred to herewith as the liberal tradition of democracy. With the birth of this liberal tradition during the Renaissance, the focus switches from man finding fulfilment as part of a community, towards man finding fulfilment from within his individual liberty and freedom (Harrison, 1993: 36-37).

The limitations that were placed on democracy in this liberal tradition include the protection of human rights, a constitution and a division of executive power. This is the birth of liberal democracy in which individual and group liberties are protected and in which there is a separation between private life and state functioning (Diamond, 1999:2-3). In the next few paragraphs, the development of the liberal tradition of democracy will be briefly examined.

²⁷ Mainly through his work, *The Spirit of Laws* (Montesquieu, 1952).

²⁸ Mainly through his work, *The Federalist Papers* (Madison, 1966).

The value of the individual was beginning to become more pronounced with the works of Thomas Hobbes²⁹ and John Locke³⁰ in the 17th century. These philosophers described individuals as self-fulfilling entities in themselves and not only according to their position as part of a community as the ancient Greek philosophers did. With this new approach to individual rights, the question that both Hobbes and Locke faced, was how much of his/her individual rights/freedom a person would or should hold in entering into a political society (Harrison, 1993: 37).

This is where intellectual constructs, such as the social contract, come into play. Many of the 17th century philosophers, including Hobbes (2008), Locke (2008) and Rousseau (2008), used this idea in order to clarify and explain the problem of political authority over free people. For this thesis, it is superfluous to examine the social contract in full, what is however relevant is to understand that the liberal ideas about individual freedom began to flourish with the developments that flowed out of these ideas.

A key figure in the development of democratic ideas was Jean-Jacques Rousseau (2008). Like Machiavelli, Rousseau was interested in republicanism and also stood between ancient and modern thought. Rousseau, writing in the late 18th century, spoke up against the absolute claims of kings. As Machiavelli challenged the state domination of the *res publica Christiana*, Rousseau was challenging the state domination of monarchies (Pocock, 1975: 504). He believed that the absolute rule of kings should give way to a system of political participation by all citizens. In his seminal work, *The Social Contract* (Rousseau, 2008), Rousseau proposed a radical system of small city-republics, in which everyone knew each other and decisions were taken by all citizens. This ideal system of Rousseau almost resembles a form of direct democracy. In this system sovereignty belonged to the collection of citizens as they were collectively responsible for every decision (Held, 2006: 43-46).

Rousseau (2008: 54) stated that a necessary condition of liberty is political participation and promoted citizen³¹ participation within a structure of protective republicanism. The welfare of the individual and the collective, according to Rousseau, is ensured via the freedom of citizens to participate in the public sphere. He (2008: 88), however, did not believe in democracy and stated that both democracy and aristocracy would degenerate into a state of anarchy. As Aristotle and Plato before him, Rousseau believed that the democracy would eventually lead to mob rule and ochlocracy.

The idea of democracy was therefore still quite a pejorative concept during the 18th century. The works of Rousseau, however, inspired the likes of Maximillian Robespierre, the French revolutionary, who was

²⁹ Mainly in the works of *De Cive* (1642) and *Leviathan* (1651)

³⁰ Mainly in his 1690, *Second Treatise of Government – An Essay concerning the True Original, Extent, and End of Civil Government*.

³¹ Citizens excluded women and the poor (Rousseau, 2008).

one of the first to use the word, *démocratie*, in the positive sense to refer to the unmediated implementation of the “will of the people” (Dryzek and Dunleavy, 2009: 20). The Jacobian rebels were unfortunately brutal in their execution of democracy and this again helped to reinforce the negative view of democracy. With the rise of Napoleon Bonaparte’s empire, it again seemed that the phenomena of democracy failed, just as its ancient predecessors. What the French revolution did however offer the theory of democracy was to promote the idea of representation. This idea would enable democracy to be feasible in large modern societies (Dryzek and Dunleavy, 2009: 20-21).

After the French revolution, in works such as Thomas Paine’s (2008) *The Rights of Man*, the ideal of representation becomes engrafted to democracy. According to Robert Dahl (1989: 28-30), this development can be seen as “the second transformation of democracy” (the first being the invention of Athenian democracy). Writers such as Destutt De Tracy (1811: 19), James Mill (2008) and James Madison (1966) also saw representation as the revolution that was necessary to bring the ideals of democracy to practical statecraft in large nation states (Dahl, 1989: 29). Mill (2008) went as far as calling the system of representation, “the grand discovery of modern times”.

What representation meant for democracy was that maintaining a democracy in a large nation state - instead of small city-states - was now more feasible. The people could now elect representatives to represent their interests *en masse*, while still maintaining liberal ideals such as personal freedom, personal autonomy and personal rights (Dahl, 1989: 30). It is today believed that the first real liberal democratic representative state was produced out of the American Revolution (Dryzek and Dunleavy, 2009: 21). The founders of this state did however not see the United States of America (USA) as a democracy, but rather as a republic. Madison (1966), for example in *The Federalist 10*, stated that:

“[D]emocracies have ever been spectacles of turbulence and contention; have ever been found to be incompatible with personal security or the rights of property, and in general have been as short in their lives as they have been violent in their deaths.”

The negative view of the term democracy was therefore still prevalent with the formation of the USA, even though this state could be seen as a democracy itself³² (Dahl (1998: 16). The important thing to note is that the idea of representation was now being enshrined into modern liberal nation states. It is also important to note that democracy still had a negative connotation during the infancy years of the USA. This negative view subsided slightly in the 19th century, especially in the USA, with leaders such as

³² Dahl (1998: 16-17) states that the terms ‘democracy’ and ‘republic’ refer to the same type of political system. The only difference between the words is that ‘democracy’ has its etymological origins in Ancient Greece and ‘republic’ in Ancient Rome. This debate over the difference between a republic and a democracy however falls outside the scope of this thesis.

Andrew Jackson, who used the term ‘democracy’ to refer to the ordinary people, compared to the plutocracy – the rich and powerful. He used a populist, anti-elitist approach to persuade this ‘democracy’ to elect him in 1828 as the seventh president of the USA. This Jacksonian era gave birth to a modern liberal democracy that was focussed on the rights of not only the elite, but also the common man (Dryzek and Dunleavy, 2009: 21-22).

2.3.3 *The Three Waves of Democracy*

The growth of liberal ideals in the French Revolution and the emergence of the USA as the first liberal representative democratic state was the root of what Samuel P. Huntington (1991) would later call the *first, long wave of democratisation*. This first wave started with Jackson and the 1828 US elections in which more than 50 percent of the US male population was eligible to vote for an executive leader (Huntington, 1991: 16). According to Huntington (1991: 17), this wave was the first in a series of waves and reverse waves of growth in modern democracies. Each wave consists of:

“...a group of transitions from non-democratic to democratic regimes, and that significantly outnumbered transitions moving in the opposite direction during that period of time”
(Huntington, 1991: 17).

There have been three waves of democratisation and the first two waves were followed by reverse waves in which the number of democracies in the world declined. Each wave will be examined briefly in this section. The first wave lasted from 1828 – 1926 and consisted mostly of North American and European states adopting modern democracies. After this first wave, there was a reverse wave from 1922-1942. During this reverse wave one saw the rise of traditional authoritarian rule and mass-based totalitarian rule (fascism). Many of the newly developed democracies now fell back into authoritarian rule. This reverse wave included the rise of Mussolini’s fascist regime in Italy and Hitler’s conquest of power in Germany. In this reverse wave, anti-democratic movements also gained momentum in countries such as France, the United Kingdom and the USA, even though these democracies did not fall (Huntington, 1991: 16-17).

With the end of the Second World War approaching, the second wave of democratisation was initiated, lasting from 1943-1962. This wave of democratisation saw Allied-occupied areas such as Germany, Italy, Korea, Austria and Japan inaugurate democratic regimes. There was also a rise in democracies in Latin America, South and South Eastern Asia and Africa, during this wave (Huntington, 1991: 18-19). After this second wave, there was again a reverse wave of democratisation. This second reverse wave lasted from 1958 – 1975. During this wave authoritarian regimes again emerged, replacing many of the new democracies. By the mid-1970s, it is estimated that a third of the 32 working democracies that existed in

1958 reverted to authoritarian rule. This reverse wave produced largely negative views on the practicality of installing democracies in newly developing countries (Huntington, 1991: 20-21).

Still, even with these negative sentiments, a third wave of democracy emerged in 1974. The end of the Portuguese dictatorship in 1974 marks the start of the rise in global democracies, which arguably is still happening today. There are those such as Diamond (1996) who argues that the third wave has already ended, but these debates are superfluous to this thesis. In the first fifteen years after 1974, approximately 30 countries from Europe, Latin America and Asia became democracies. This wave also spread to Africa and the Middle East in the late 1980s and early 1990s (Huntington, 1991: 21).

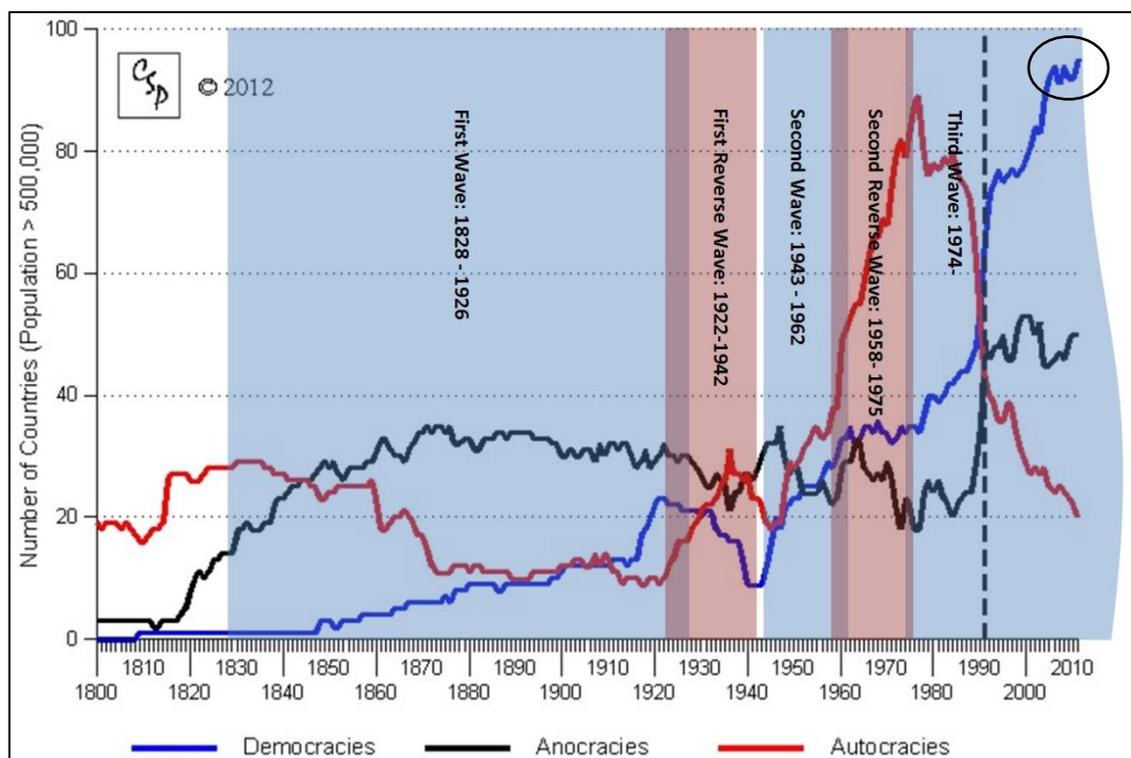


Figure 6: Global Trends in Governance, 1800-2011 - Adapted from Graph for the Polity IV Project Source: (Polity IV Project, 2012).

The graph above (*Figure 6*), which I adapted from a *Center for Systemic Peace: Polity IV Project* graph (2012), clearly shows the rise in the number of democracies since the 1800s. Blue and red highlighted areas on the graph indicate the waves and reverse waves of democratisation. The vertical black dashed line represents the fall of the Berlin Wall in 1989. One notes the first reverse wave of democratisation (1922-1942) by the clear dip in the democracy (blue line) during this time period. The second reverse wave (1958-1975) is less obvious, but one does note the growth of democracies stagnates a bit during this time. The second reverse wave is more clearly seen in the rapid rise in autocratic regimes (red line) during

this time. There is a strong rise in the number of democracies from the mid-1970s, which corresponds to the third wave of democracy (1974-). Perhaps the most interesting aspect of the graph can be found during the last few years (circled section) which seem to hint at stagnation in the growth of the number of democracies in the world during this period. This study is not the place to examine this stagnation further, but if the results of this thesis show that there is clear negative sentiment regarding democracy, future studies might examine this stagnation more closely.

If one knows that a large section of the world population has a negative view on democracy or at least certain aspects of a specific democracy, it can be argued that the time is ripe for another evolutionary step in the development of the system of governance. As has been shown in this section, there have been many developments in democracy, and democracy is certainly not static. In a subsequent section of this chapter it will be shown that there are negative sentiments on democracy that are developing in the years after the 2008 global financial crisis. These negative sentiments were the inspiration for this study and might also be an indication as to why there seems to be stagnation in the graph

Before the recent negative sentiment on democracy is examined, a short section follows on the development of theories of democracies during the so-called third wave. These theories examine how democracies are classified and what roles citizens should play in the democracy. It is the belief of this author that these developments - in the way that the world perceives democracy - lies at the heart of the recent negative sentiment on democracy. From the Tweet data and from the academic debates it would seem that people are calling for a more involved citizenship, which perhaps alludes to a more participatory form of governance. The aim of this thesis is however not to discover what system of government people would prefer in the future, but rather if people are in fact negative about the democracy that is present today, as expressed on Twitter.

2.4 A CONTEMPORARY VIEW OF DEMOCRACY: EFFECTIVE DEMOCRACY (WHY PEOPLE MIGHT BE NEGATIVE)

During the early 1990s it seemed that democracy was again at an apex of support. The third wave of democratisation (Huntington, 1991) attracted the attention of a substantial number of academics³³. Authors such as Francis Fukuyama (1989) and Pye (1990) wrote about the end of authoritarianism and the culmination of human development expressed through Western democracy and capitalism. This was at end of the Cold War and it seemed that liberal “Western democracy” was becoming the only game in town (Dalton, 2006: 245).

³³ Academics such as O’Donnell and Schmitter (1986); Higley and Gunther (1992); Casper and Taylor (1996); Linz and Stepan (1996).

With the third wave of democracy a broad ‘electoral’ view of democracy gained prominence. This definition dates back to Joseph Schumpeter (1976) and is used by Adam Przeworski and Fernando Limongi (1997) in *Modernization: Theories and Facts*. According to this broad conception of democracy, any regime that holds competitive, free, fair, and regular elections can be classified as a democracy. This form of electoral democracy gives prominence to the elite, as democracy emerges out of the strategic agreements of the elite (O’Donnell and Schmitter, 1986). According to Inglehart and Welzel (2005), the mass preferences of the broad citizenry are greatly neglected in this definition of democracy.

This electoral definition of democracy evolved out of the critique of Schumpeter (1976:250-251) on what he calls the ‘classical definition’ of democracy. Schumpeter stated that the classical definition defined democracy in terms of its institutions and the ability of these institutions to ensure that the general will of the populace is followed. He believed that this was an unsatisfactory way of defining democracy as the will of the populace is not certain. Kenneth Arrow (1951) showed in later years the impossibility of defining aggregate preferences of a big populace when many issues are at stake. This showed that it was perhaps nonsensical to view democracy as the implementation of the general will (Knutson, 2010:110). However, not all authors agreed with this sentiment, and many opposed electoral theories on democracy.

Opposing these electoral theories of democracy were more liberal theories on democracy, such as the liberal form of democracy described by Robert Dahl (1972) in *Polyarchy: Participation and Opposition*. This liberal notion of democracy included a wider set of civil and political rights as an integral part of democracy.

After the initial ‘end of history’ type studies, a more tempered view of democracy emerged, showing that democracy did not function equally well in all circumstances. Many of the newly developing democracies, for example, were showing great deficiencies. Authors³⁴ such as O’Donnell (1993) and Finer (1999) began to warn against confusing “effective democracies” with “façade democracies”. The difference between these two forms of democracy rests in the level of personal rights and freedom that is expressed through the “rule of law” in a specific country. A democracy, according to this idea, needs more than legal legislation and elections to make “freedom” rights effective³⁵. According to Christian Welzel (2002:270), the president of the World Values Survey, what democracy needs to be effective is the support of more empowered citizens. This will be discussed subsequently.

³⁴ See also Linz and Stepan (1996) and Rose (2001)

³⁵ This sentiment has also been expressed in earlier works: See Almond and Verba (1963).

2.4.1 *Empowered Citizens*

There are various social mechanisms that are present in democracies, which could hinder the citizens in practicing their rights efficiently. The most detrimental of these mechanisms, according to Welzel (2002:270), are “elite corruption” and “elite closure”. Welzel states that both these mechanisms work against the core principles of democracy. Elite corruption leads to the violation of the ‘rule of law’ and elite closure limits the equality of rights. In order to limit these mechanisms, Welzel and Inglehart (2008) argue for a more empowered citizenry.

The empowerment of ordinary citizens is the essence of democracy, according to Welzel and Inglehart (2008:128). If the officeholders do not respect the rights that exist in a democracy, the democracy will not be ‘effective’ and would lack legitimacy. Three elements have been identified that together can lead to effective democracy (Welzel and Inglehart, 2008:129). These three elements are: *action resources*, *self-expression values*, and *democratic institutions*. Welzel and Inglehart, (2008:129) call this the *Human-Empowerment Triad*.

Action resources, in this triad, refer to material and cognitive resources that the citizenry possess, which help them to govern their own lives. Aspects such as education and skills development, which make people more equipped to participate in political society, fall under this leg of the triad (Welzel and Inglehart, 2008:129).

The next leg of the triad, self-expression values, has emerged with the rise of post-industrial society and refers to the priority that citizens give to self-expression. Welzel and Inglehart (2008:129-130) state that “societies that place a high value on self-expression tend to have participatory orientations toward society and politics; support gender equality; are relatively tolerant of foreigners, homosexuals, and other out-groups; and rank high on interpersonal trust.” With a rise in self-expression values, a push for more responsive governance, improved political liberties, free choice and gender equality can also commonly be seen (Welzel and Inglehart, 2008:130).

The final leg of the triad, democratic institutions, provides the civil and political rights that allow people to shape their private and public life. The rise of action resources and self-expression in a society, inevitably leads to better and stronger democratic institutions, according to Welzel and Inglehart (2008:130).

From this Human-Empowerment Triad, Welzel and Inglehart (2008:131) identify self-expression values as the strongest predictor of effective democracy. What this means is that from a number of pro-democratic attitudes, the appearance of self-expression values is most conducive to democracy (Welzel,

2007: 397). In older studies such as Almond and Verba's (1963) *The Civic Culture*, it was shown that certain mass attitudes to democracy were conducive to democracy³⁶. This study did not however provide statistically reliable results that could show whether certain individual-level attitudes were linked with the functioning of democracy.

Using extensive³⁷ World Values Survey data, Welzel and Inglehart (2008) have demonstrated that if certain attitudes are more prevalent in a specific society, then the society tends to be more democratic. These self-expression values include support for interpersonal trust, tolerance of other groups, and political activism. According to Welzel (2006); and Welzel and Inglehart (2008:132) these values are a strong indicator for the emergence, survival and growth of democratic institutions.

It is surprising to note that mere positive expressions on democracy do not correlate with deep-rooted orientations and strong motivation for democracy according to this study. Welzel and Inglehart (2008:132) state that most pollsters around the world are willing to pay positive lip service to democracy, but that this does not indicate deep-rooted endorsement for democratic values. These positive attitudes towards democracy from survey data will be illustrated in Chapter Three of this thesis. Without going into too much detail it should however be noted that the results from this thesis show that the positive lip service that pollsters pay to democracy does not extend to expressions on Twitter. Speculations on the reason for this can be found in the final chapter of this thesis.

2.5 EVOLVING VIEW OF DEMOCRACY – THE RISE OF NEGATIVE SENTIMENT

Throughout the turbulent development of democracy the status quo of democracy does not last forever, as has been shown in this chapter. In the conclusion of the Welzel and Inglehart (2008:139) paper, for example, it is stated that when representative democracy first emerged, it accepted elements such as slave labour and the disenfranchisement of women, which would not be accepted in modern definitions of democracy. This changed with time as people became in turn disenfranchised with the old values of democracy. Today scholars and the public alike might be, once again, becoming critical of narrow definitions of democracy.

This view of an adaptive and evolving democracy is by no means novel. The prediction of the seminal work of Crozier, Huntington and Watunaki (1975), in *Crises of Democracy*, warning against the demise of democracy highlights the emergence of similar sentiment in the 1970s. In the introductory note to their

³⁶ Also see Eckstein (1966).

³⁷ This study is more statistically reliable than the older study, as it includes data on about 90% of the world's population, compared to the five countries studied in the Almond and Verba (1963) case.

book, it is stated that: “it is important for citizens of our democracies to re-examine the basic premises and the workings of our system.” If liberal democracies are to continue to prosper, they have to be willing to evolve, adapt their institutions and therefore not stagnate (Crozier, Huntington and Watunaki, 1975: 1-5; 157-158). This view relates to the picture of a constantly evolving theory of democracy that has been drawn throughout this chapter. This sentiment also fits well with the view of Dahl (1998: 25), stating that democracy follows a disjointed path of development and that it is constantly evolving.

It has been shown in this chapter, that throughout the development of democracy, one finds negative views of the term democracy itself, but also of certain aspects/institutions of democracy. Against the backdrop of the recent success (third wave) of democracy, authors such as Dalton (2004) argue that there are indications that the citizens of many countries are, once again, slowly becoming more critical of democracy. Dalton believes that citizens are also becoming weary of political parties, politicians and political policies in recent years. He states that this rising level of criticism started with US critique surrounding specific events and crises, such as the Vietnam War, the Watergate Scandal and the Camp David aftermath. After the 9/11 events in 2001, the public’s support and opinions of government rose for a while, but this was only temporary. Furthermore, he believes that the negative public sentiment on democracy is not only limited to America, but can be witnessed globally (Dalton, 2006:246).

In the aftermath of the global financial crises of 2008-9 and with the continuing Eurozone crises more critical views on democracy have arisen. As has been discussed in the introduction to this thesis, there are authors such as van Beek (2012:12) that believe that the global degradation of the global economy might be bringing the dominance of Western developed democracies to an end which might in turn lead to the erosion of global confidence in democracy. Movements such as the Spanish Indignados and Occupy Wall Street might also indicate this erosion, but it falls outside the scope of this thesis to examine these movements here.

Global erosion in the confidence of democracy seems quite harsh, but one should understand that this erosion might in fact be referring to a changing view on democracy. It has been shown that negative sentiments regarding democracy have throughout history been followed by changes in the understanding and functioning of democracy. Negative sentiment regarding democracy can be aggravated by global events such as global financial crises (Diamond, 2011:17). A final example of how negative sentiment on democracy can be aggravated by global events and lead to changes within one’s understanding of democracy can be found within the aftermath of the Great Depression in 1929 (Wnuk-Lipinski, 2012: 217).

During the years following the Great Depression there were a number of parliamentary democracies which, due to legitimacy problems, fell into autocracy. The Weimar Republic was the most prominent, and one of the consequences of this fall was the Second World War (Wnuk-Lipinski, 2012: 217). According to Dirk Berg-Schlosser (2012: 47-60), the 2008-9 recession was not as severe as in the 1930s and there have not been such strong extremist political responses to the crises as was seen after the Great Depression. In fact, with the emergence of the so-called Arab Spring, the world has seen the formation of new democracies. It still, however, remains to be seen whether these new democracies will be consolidated (Wnuk-Lipinski, 2012: 220). It therefore seems unlikely that the recent recession would lead to consequences as grave as the Second World War.

What one does however see is that issues of global governance, specifically the issue of accountability of leaders, are being raised. The response from Western developed countries, which have been hardest hit by the crisis, has been especially strong (Du Plessis, 2012: 41-44). The accountability of policymakers with regards to globally significant decisions is specifically being challenged (Wnuk-Lipinski, 2012: 218).

In his paper, *Why Democracies Survive*, Larry Diamond (2011) states that both survey data and objective trends suggests that a crisis of democracy in the short run will more likely flow out of poor quality of governance than out of financial strain. He discusses the possibility of a rollback of democracy and the end of the third wave of democracy. In the paper, he does not state that the Third Wave is at an end yet, but he does believe that there are some worrisome developments that might lead to the end of the third wave. At the heart of the issue he places governance problems that are rooted in deficient political institutions and norms.

Only the future will tell what the final outcome of global events such as the financial crisis and the Eurozone crisis will be. From the above discussion I, however, want to highlight that there is at least a possibility that global governance structures and/or democracy might be affected. As has been shown, there are those who believe that global sentiment on democracy might be eroding (Dalton, 2006; van Beek, 2012). Seeing that the opinions of citizens are paramount to the functioning of a democracy, one might therefore look at what the people are saying regarding governance and specifically democracy in order to determine if the erosion of democracy is already present in society. This leads one to the question of how one might measure what people are thinking about political concepts such as democracy. This question will be discussed in the next chapter.

2.6 CONCLUSION

This chapter has looked at the historic developmental path of democracy. It has been shown that democracy as a system of governance has adapted throughout the years in order to remain a relevant governmental system. A potential rise in negative sentiment that has been discussed by a number of authors might indicate that democracy is again at a point in history when it will have to adapt in order to stay relevant. In the next chapter contemporary and novel methods of measuring whether views on democracy are indeed negative will be examined.

Chapter 3: The Development of Measuring Expressions on Democracy

3.1 INTRODUCTION

This chapter can be divided into three broad topics. Firstly, the value of measuring expressions on democracy will be briefly examined. Traditional literature on the value of public opinions to the democratic consolidation and functioning of democracies is given in this section. Secondly, an examination of traditional methods of measuring expressions on democracy can be found. Finally, a section on novel emerging methods of measuring public expressions from online sources is found. This section will specifically look at a number of studies, such as O'Connor *et al.* (2010), which have utilised Twitter in order to gauge public expressions. The broad aim of this chapter is to show the importance of measuring expressions and opinions on democracy, to show how this measurement has been done up to now; and finally, to illustrate novel methods - such as the method developed in this study - that might lead to an alternative method of measurement. This author does not claim that this method will replace traditional methods of measuring expressions on democracy; but does believe that it could supplement the data gathered using traditional methods.

3.2 THE VALUE OF MEASURING EXPRESSIONS ON DEMOCRACY

A key hallmark of a well-functioning consolidated democracy is the unconditional embrace of the citizenry of democracy as “the only game in town” (Bratton and Mattes, 2001; Diamond, 1999; Shin, 2006, Linz, 1990; Rose, 2001). The belief of the citizens of a democracy, in the legitimacy of the democratic system, has become a central tenet of a stable democracy (Diamond, 1999:168; Fails and Pierce, 2010:174). It is therefore critical to the stability and depth of a democratic regime to have a citizenry that believes in the legitimacy of the democratic process.

With the potential of a crisis - or an erosion of democracy, as van Beek (2012:12) calls it - it is important to understand and be able to analyse the functioning of democracies across the world. With this in mind, after the third wave of democratisation, there has been a rise in studies looking specifically at the level of democratic consolidation in new and old democratic regimes. These studies can be broadly divided into two main traditions. Firstly, there are scholars such as Horowitz (1991), Lijphart (1985 and 1990) and Sisk (1994) that focus on the institutionalisation and habitualisation of political practices and institutions.

They examine how, through constitutional design, the best procedures and institutions for democratic governance can be attained.

Secondly, there are researchers such as Almond and Verba (1963) and Inglehart (2003) that examine the practices and values of civil societies, together with the attitudes and values of the citizenry, in order to gauge the level of democratic consolidation of a country. This study broadly falls under this second category. These scholars argue that the quality of the institutions in a democratic regime is not the most important aspect for democratic consolidation. They believe that one should rather look at the attitudes of the citizens participating in the democratic process, in order to understand the level of democratic consolidation. Their rationale is that if there are high levels of support for sustained democratic governance within the citizenry, democracy will tend to flourish.

The measurement of overt support for democracy, in a democratic system, has obvious face validity according to Inglehart (2003: 51) - seeing that lack of support for a political system based on majority support would raise serious questions as to the validity and the legitimacy this system. In Inglehart's (2003) study *How Solid Is Mass Support for Democracy: And How Can We Measure It?*, World Values Survey data for the period of 1981-2000 are used to show that if one examines the long-term stability of democracy, positive attitudes towards democracy are a good indicator for this stability.

In a study by Rose *et al.* (1998) political institutions are compared to the “hardware” of democratic systems, and the opinion of the citizens on these institutions is compared to the software of the system. Extending this metaphor, they state that just as a computer cannot function without the combination of both hardware and software, so a democratic system cannot function without proper institutions and public support for these institutions. In this thesis the focus is placed more on the “software”, namely the support for democracy that is gathered from the public expression of sentiment, gathered from the Tweets about democracy.

3.3 TRADITIONAL METHODS OF MEASURING PUBLIC EXPRESSION ON DEMOCRACY

Now, if one is of the opinion that support for democracy is important to democratic consolidation, the next logical step would be to attempt to measure the level of this support. There are in fact many empirical research programmes that are monitoring public support of democracy and democratic institutions. Some of the most well-known of these programmes are the AfroBarometer, the European Values Survey, the New Russia Barometer, the LatinoBarometer, and the World Values Survey. Each of these programmes have sections in their surveys that aim at gauging individual support for democracy.

These individual levels of support can then be tallied to show mass support for democracy in a specific area.

It is interesting to note, that according to Inglehart (2003: 51), many of the questions from the different surveys overlap in their use of indicators for democratic support. What is even more relevant to this study is that these surveys contain a high level of internal consistency. What this means is that if a person supports democracy on one of the indicators, they tend to support democracy on the other indicators as well. For this study it would mean that if a person (Tweet in this case) shows support for a specific aspect of democracy, there is a high probability that the person would also support other aspects of democracy.

This study treats support for democracy more broadly than the research programmes mentioned above. For this introductory study focus is only placed on broad levels of support for democracy as expressed through positive or negative sentiments contained in Tweets relating to democracy for the chosen period.

In order to gain a better understanding of exactly what the various surveys measuring expressions and opinions on democracy aim to determine, this study will now inspect a number of global and regional research projects on democracy.

3.3.1 Global Research Projects

Firstly, research projects that examine attitudes and expressions on democracy on a global level will be examined.

3.3.1.1 The World Values Survey

The World Values Survey (WVS) calls itself: “The world’s most comprehensive investigation of political and sociocultural change” and is the most famous global research project examining opinions, values and beliefs of citizens, on democracy. Since 1981 the project has collected data from more than 100 countries (WVS, 2012). The most recent available World Values Survey is the 2005-2008 round. The survey used in this round contains a number of subsections concerning democracy. Of these subsections, the most direct question regarding the support for democracy is in section V162, namely:

1. *V162. - How important is it for you to live in a country that is governed democratically? On this scale where 1 means it is not at all important and 10 means absolutely important - what position would you choose?*

		Total
V162: Importance of democracy	Not at all important	1.50%
	2	0.50%
	3	0.70%
	4	1.10%
	5	5.90%
	6	4.70%
	7	7.40%
	8	14.10%
	9	14.40%
	Absolutely important	49.60%
	Total	73986 (100%)

Source: (WVS, 2012).

From the respondents of the 57 countries in this round, a majority (90.20 percent) answered between 6 and 10, which can broadly be seen as being in support of democracy (WVS, 2012).

Another question relating to the political system of democracy reads:

2. *V151. - I'm going to describe various types of political systems and ask what you think about each as a way of governing this country. For each one, would you say it is a very good, fairly good, fairly bad or very bad way of governing this country? Having a democratic political system.*

		Total
V151: Political system: Having a democratic political system	Very good	51.50%
	Fairly good	40.10%
	Fairly bad	6.30%
	Very bad	2.20%
	Total	73787 (100%)

Source: (WVS, 2012).

These results show that 91.6 percent of the respondents believe that having a democratic political system is a fairly or very good way of governing their specific country. What both the results from question number 162 and 151 show is that it seems that globally people are broadly in support of democracy as a political system and would like to live in a democracy.

It should be noted that there are other sections of the survey that examine characteristics of people in democracies and their opinions on institutions in democracy. Seeing that this thesis focuses on broad opinions of democracy, these aspects will not be examined here.

3.3.1.2 Voice of the People

Gallup International Association ran an annual international survey from 2002-2005 called Voice of the People (2012). This is an example of another global research project, which examined opinions of respondents from more than 50 countries. This survey examined public opinion on various social elements, one of which was opinions on democracy. The two relevant questions, for this study, that this survey asked respondents regarding democracy were:

1. *“In general I am satisfied with democracy?” (Agree/Disagree).*
2. *Democracy may have problems but it is the best system of government?” (Agree/Disagree).*

In the final 2005 round the results for both questions showed that a majority of the respondents were in general satisfied with democracy and that it was the best system of government. For the first question 65 percent of the global respondents were satisfied with democracy. The second question showed that 79 percent of the global respondents believed that democracy was the best system of government. Again, as in the WVS, it seems that a majority of the global population is in broad support of democracy. The results of the 2005 VOTP survey can be seen in the two graphs below (*Figure 7 and Figure 8*).

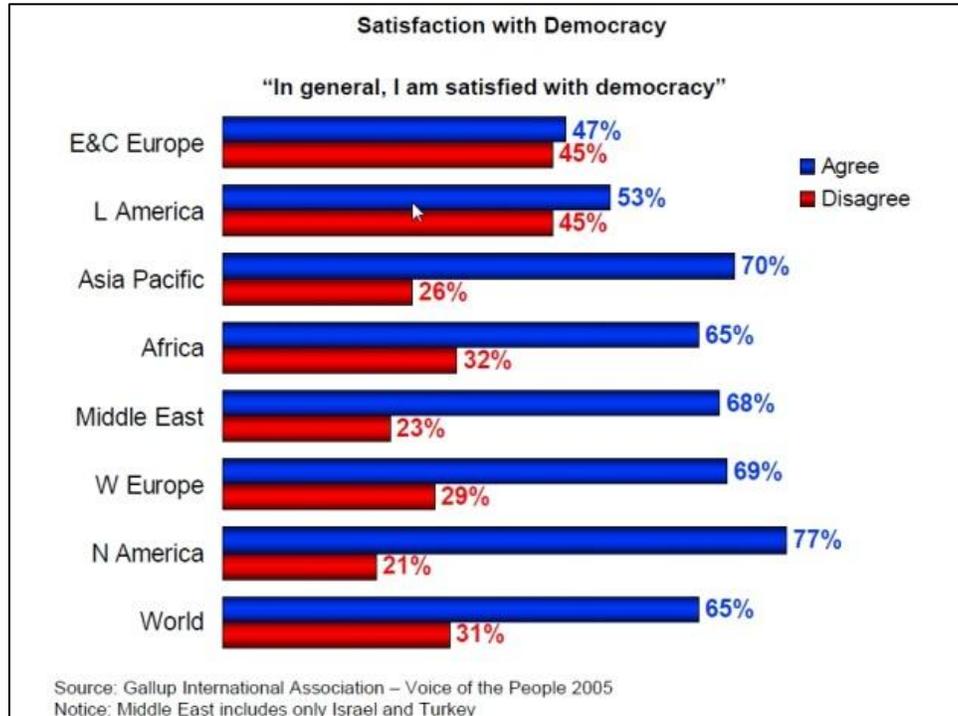


Figure 7: VOTP Global Satisfaction with Democracy - Source: (VOTP, 2005)

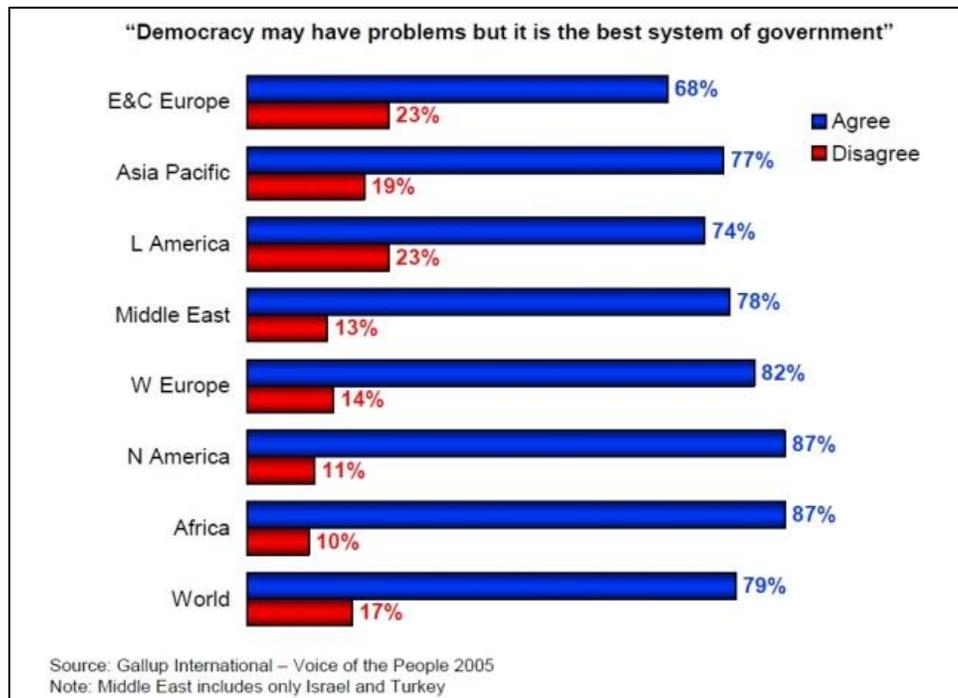


Figure 8: VOTP Global Sentiment on Democracy - Source: (VOTP, 2005)

3.3.2 Regional Research Projects

There are a number of regional projects that gather data similar to that gathered in the WVS, but more specific to a region. In this section questions regarding democracy from a number of these projects will be inspected. One should note that again this researcher only highlights certain general questions about democracy and that these surveys contain much more in-depth questions. These deeper questions could be considered for future studies.

3.3.2.1 AfroBarometer

AfroBarometer, in partnership with the Institute for Democracy in South Africa (IDASA), the Ghana Centre for Democratic Development (CDD), and the Department of Political Science at Michigan State University, measures public attitude on democracy, markets, civil society, and other aspects of development (AfroBarometer, 2008). There have been four rounds of this programme, the most recent of which is the 2008-2009 round. In the 2008-2009 round of the survey, 26 414 respondents from 20 African countries answered questions such as:

1. Q30. - Which of these three statements is closest to your own opinion?

		Total
Q30. Support for democracy	Missing	*
	Statement 3: It doesn't matter	12.6 %
	Statement 2: Sometimes non-democratic preferable	11.8 %
	Statement 1: Democracy is preferable	75.6 %
	Total	22127 (100%)

Source: (AfroBarometer, 2008a)

From this question one can see that 75.6 percent of the respondents from the 20 African countries preferred democracy. As with the results from the global level programmes above, there again seems to be a strong broad support for democracy.

2. Q42a - In your opinion, how much of a democracy is [country] today?

		Total
Q42a: Extent of democracy	Missing	0.10%
	Not a democracy	6.90%
	A democracy, with major problems	27.50%
	A democracy, but with minor problems	31.50%
	A full democracy	29.20%
	Do not understand question/democracy	4.80%
	Total	22474 (100%)

Source: (AfroBarometer, 2008a)

3. Q43 - Overall, how satisfied are you with the way democracy works in [country]?

		Total
Q43: Satisfaction with democracy	Missing	0.10%
	Not a democracy	1.50%
	Not at all satisfied	18.50%
	Not very satisfied	26.50%
	Fairly satisfied	30.20%
	Very satisfied	23.10%
	Total	21963 (100%)

Source: (AfroBarometer, 2008a)

Questions Q42a and Q43 both focus on opinions surrounding democracy. Q42a shows that 65.90 percent of the respondents felt that they were either not living in a democracy, or that there were some problems with the democracy within which they lived. Q43 shows that about 45 percent of the respondents were not at all or not very satisfied with democracy compared to 53.30 percent that were fairly to very satisfied. The results of both these questions highlight the fact that even though the respondents were broadly in support of democracy (as explained by Q30 above), there were strong issues with the functioning of democracy and fairly high levels of dissatisfaction. This researcher believes that the results of this study (found in Chapter Five) relate strongly to these types of question. Many of the Tweets that were analysed in this study also express sentiment on problems within specific democracies and statements as to whether

a country is a democracy at all. It is therefore this type of question that this researcher believes can be supplemented with the methods developed in this study.

3.3.2.2 *LatinoBarometer* (*Latinobarómetro*)

The LatinoBarometer is similar to the AfroBarometer, except that it examines attitudes from respondents from 18 Latin American countries. The 2010 round of this survey includes similar questions to that of the AfroBarometer 2008-2009 round, as can be seen below (LatinoBarometer, 2012). For the purposes of this study, the questions, originally in Spanish, have been translated using Google’s automatic translation function embedded in its web browser, Google Chrome (Chrome, 2012).

1. *A101. - Which of these three statements is closest to your own opinion?*

		Total
A 101: Support for democracy	Democracy is preferable to any other form of government	68.00%
	In some circumstances, an authoritarian government can be preferable	15.30%
	People like me, do not care about being democratic	16.80%
	Total	21117 (100%)

Source: (LatinoBarometer, 2012)

2. *A107. - Democracy may have problems, but it is the best system of government (Churchill’s phrase).*

		Total
A 103: Democracy may have problems, but it is the best system of government	Strongly agree	29.70%
	Agree	53.30%
	Disagree	14.10%
	Strongly disagree	2.90%
	Total	21325 (100%)

Source: (LatinoBarometer, 2012)

In this round of the LatinoBarometer, 68 percent of the respondents stated that democracy is preferable to any other form of government. A further 83 percent of respondents agrees or agrees strongly that even though democracy might have problems, it still is the best system of government. Again, as with the previous studies, one can see the broad support of democracy as a form of government.

Question A102 below, however, again highlights the fact that there are problems in the functioning of democracy, seeing that 53.9 percent of the respondents are not very – or not satisfied at all - with democracy in their country. This again is the type of question that this researcher believes can be supplemented by data gathered with the novel method proposed in this study.

3. A102. - In general would you say you are very satisfied, not satisfied with the way democracy works in [country]?

		Total
A 102: Satisfaction with democracy	Very satisfied	10.00%
	Fairly satisfied	36.20%
	Not very satisfied	38.80%
	Not at all satisfied	15.10%
	Total	21117 (100%)

Source: (LatinoBarometer, 2012)

The LatinoBarometer contains a total of 50 questions relating to democracy. Some of the more relevant questions for this study, besides the ones discussed above, are:

- A108 Democracy is the only system that (country) can be developed
- A109 Democracy is indispensable to become a developed country
- A110 Would you be willing to defend democracy when it is threatened?
- A111 Trust in democracy as a system of government for their country to become a developed country

Due to the scope of this study, all these questions will not be examined in detail.

There are various other research projects looking at opinions on democracy such as the ArabBarometer, the European Value Survey, New Democracies Barometer³⁸, the EuroBarometer³⁹, the New Russia

³⁸<http://www.nccr-democracy.uzh.ch/research/module5/barometer/democracy-barometer-for-established-democracies>

Barometer⁴⁰, the Korea Barometer,⁴¹ the Democracy Barometer⁴², etc. It also falls outside the scope of this study to examine all of these projects. These are nevertheless all examples of projects that use broadly the same methods of survey research in order to gauge broad public expressions.

3.4 TWITTER AS A NEW PUBLIC SPHERE

The question now beckons whether Twitter would be able to provide one with additional insights to the above-mentioned projects. Before this question is answered, the idea of Twitter as a type of Habermasian public sphere needs to be briefly examined. As a delimitation of this study, Twitter is viewed as such a public sphere, but it is important to understand why this is done and what is meant by this assumption.

It was shown in the introductory chapter of this study that there have been various forms of public spheres throughout history. Scholars such as Habermas (1989) and Arendt (1963) believe that the public sphere is constantly changing. Habermas (1989) has looked specifically at the influence of the media on political opinion in the public sphere. In 2000, John Thomson (2000) argued that the media is the major component of the public sphere in industrial society. However, this belief is not held by all academics.

The public sphere is more than just a public space. A healthy public sphere should function independently from the state and/or the economy (Garnham, 1992; Habermas, 1974). Furthermore, it should facilitate rational discourse on public affairs directed towards the common good (Garnham, 1992; Habermas, 1974). Authors such as Carey (1995), Habermas (2004) and Papacharissi (2010) have argued that the commercialised mass media have created a public space rather than a healthy public sphere. In this public space the rhetoric and objectives of the media are colonised by capitalist interests. The question can now be asked whether Twitter would fall under the category of public space, or a public sphere. In the public space social matters might be discussed, but would not necessarily lead to the formation of public opinion that could potentially award legitimacy to specific state institutions such as democracy, as would happen in a Habermasian (1974) public sphere.

Using the ideas from Habermas, a number of scholars such as Papacharissi (2002, 2009, 2010), Castells (2008), Downey and Fenton (2003), Dahlberg (2001), Trenz (2009) and Gruicic (2011) have examined the Internet as a new type of public sphere. This researcher is, however, unaware of any formal studies examining whether Twitter, in fact, can be fully classified as a public sphere. There are informal articles such as those by Sluis (2009), Hauth (2010) Clemens (2012) and Neal (2012) that express a range of views on this question. However, most of these articles conclude that it is unclear whether Twitter at

³⁹ http://ec.europa.eu/public_opinion/index_en.htm

⁴⁰ http://www.cspp.strath.ac.uk/catalog1_0.html

⁴¹ <http://www.koreabarometer.org/>

⁴² <http://www.democracybarometer.org/>

present is a healthy public sphere. The closest analysis of new media and the Internet as a form of public sphere can be found in the works of Zizi Papacharissi (2002, 2010) and Downey and Fenton (2003).

Downey and Fenton (2003: 200) state that public discussion on new media sites might lead to the creation of counter-public spheres, which could challenge the opinions of the current public sphere. This idea of counter-public spheres was developed by Habermas (1992:427) when he recognised that alternative public spheres might challenge the current dominant sphere. This might also be the case for Twitter, but this researcher is of the opinion that expressions on Twitter would rather form part of the dominant public sphere. This however remains to be tested in future studies.

In her book, *A Private Sphere: Democracy in a Digital Age*, Papacharissi (2010) argues that new technologies have in fact eroded the boundaries between the private and the public sphere. She believes that the public sphere is no longer the centre of democratic debate and that private citizens now broadcast gestures of social, civic, political, cultural and economic nature from the mobile enclosures of the private sphere, to mass global audiences. She believes that these spheres are not fully public as there is an unfair distribution of access to information, a lack of reciprocity in communication and an overarching commercialisation that inhibits open discussion. Papacharissi therefore would rather explain expressions on new media as falling within the private sphere.

There is no definitive answer yet as to whether Twitter can be seen as a Habermasian public sphere in which public opinion can be created. In this study it will however be accepted that Twitter does indeed meet the criteria of a public sphere. The expressions measured in this study would therefore form part of the discussion that influences the formation of public opinion.

3.5 EMERGING METHODS OF MEASURING PUBLIC EXPRESSIONS

There are other ways of gauging public opinion than the traditional research projects described previously. With the increasing availability of opinion-rich resources – specifically online sources – there has been an eruption in the area of opinion mining and sentiment analysis in recent years (Pang and Lee, 2008). The development of automated computerised methods of analysing Big Data online has given birth to a new generation of studies on public expression. These studies can be broadly grouped into those that aim to predict future events and those that aim to explain or describe the present using data from online sources. This study falls under the latter category.

Prediction studies have aimed to broadly predict the future, with studies such as Asur and Huberman's (2010) work on predicting the future with social media. Other researchers have tried to predict specific events such as changes in the stock market (Bollen *et al.*, 2011), elections (Lindsay, 2008; Tumasjan *et*

al., 2010; and Gayo-Avello, 2012) and even medical emergencies such as a Swine Flu pandemic (Ritterman *et al.*, 2009). However, these predictive studies are in their infancy and have received strong criticism from writers such as Gayo-Avello (2012), who stated that at present it is still not possible to predict future events from social media data.

This study falls under the second category of studies looking at automated analysis of Big Data from social media sites as a descriptive method. These more descriptive studies have focused on measuring public mood, public opinion and various forms of public expressions from social media sites such as Twitter. A brief look at a number of influential studies can provide one with a better understanding of this emerging field of research.

Studies such as Pang and Lee's (2008) *Opinion Mining and Sentiment Analysis* examine various sources of online information that could provide meaningful commercial and political opinions. The Pang and Lee study specifically examines online advice columns, blogs, consumer ratings on commercial sites, discussion forums and social media sites. Pang and Lee (2008: 5-6) aimed to develop an opinion search application that will "crawl" through the Internet and extract opinionated information. There are, however, a number of obstacles listed in this study regarding this type of automated extraction application. Most of the obstacles relate to the ability of computers to understand natural (human) language (Pang and Lee, 2008).

A big portion of research on measuring public expressions from online sources focuses on data from Twitter. The reason for this, as has been discussed earlier, is that the Twitter information is publicly available, downloadable and contains short concrete sentences often expressing sentiment.

One of the earliest studies looking at Tweet data was conducted by Bollen *et al.* (2009) and attempted to model public mood and emotion using Twitter. This study aimed to use Tweet data in combination with a psychometric mood measurement instrument in order to classify the mood of Tweets. The results of the study showed that real world events such as stock market fluctuations and elections have an immediate effect on various dimensions of public mood as expressed through Twitter. They speculated that a large-scale analysis of public mood could provide a framework for measuring collective social and emotional trends (Bollen *et al.*, 2009).

However, most of these studies were created by computer scientists; but in 2010, Hopkins and King (2010) decided that they wanted to create a method to categorise the large amount of digitally created information of measuring sentiment that would be more useful to the social sciences. In this study they developed a method of automated nonparametric content analysis for social science. This method

examines a collection of digitised texts and approximates unbiased estimates of category proportions. The authors also include a coding package (Hopkins and King, 2009) that can be used to replicate their results. The incorporation of this automated nonparametric content analysis method can be considered for future research expanding on the work in this paper⁴³.

In 2011, a study by Diakopoulos and Shamma (2011) examined the possibility of characterising debate performance from aggregated sentiment expressed on Twitter. They looked specifically at sentiments after televised US presidential debates in 2008. The results of their research showed that sentiments on Twitter favoured Obama over McCain during the debates. They believe that news agencies would benefit from analysing debate results in this manner.

The most significant work with regards to this study however is an article by O'Connor *et al.* (2010). This study, titled “*From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series*”, compared survey data on consumer confidence and political opinion with sentiment word frequencies in Twitter messages. The study found that there is a strong correlation between the survey- and sentiment data - in some cases the correlation was as high as 80 percent. The study therefore showed that it is theoretically possible to supplement, or in some cases even substitute, survey data with sentiment data gathered from Twitter. They argued that examining public expressions from Twitter to supplement or replace survey data could save money, time and allow one to collect information on a much wider variety of topics. This study is used as the backbone of the operationalisation section of this study which can be found in the next chapter (Chapter Four).

3.6 CONCLUSION

In this chapter the relevance of measuring public expressions and opinions on democracy has been examined. It has been shown that opinions on democracy are paramount to the democratic consolidation and broad functioning of democracy in a country. Seeing that opinions on democracy are important, it is understandable that there are many projects attempting to measure these expressions and opinions. In this chapter, traditional, survey based, projects of measuring opinions on democracies have been described. It has further been shown that even though it is still uncertain whether Twitter can be seen as a Habermasian public sphere, this study will treat it as such. Finally a number of novel methods for measuring expressions from online sources have been described. One of these methods, described by O'Connor *et al.* (2010), will be used as a basis for the method of measuring public expressions on democracy from Twitter that will be developed in the next chapter.

⁴³ It was excluded in this study as the Chatterbox Platform was used for analysis.

Chapter 4: Operationalisation and Measurement

4.1 INTRODUCTION

This chapter is divided into three broad sections. In the first section the operationalisation of measuring public expressions from democracy is examined. This section develops the specific research procedures (operations) that will result in empirical observations for measuring public expressions on democracy from Twitter. This is the method that forms the basis of the main research question on the *extent to which one can measure public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software*. In this section the measurement instrument, the Chatterbox Analytics platform, is examined. The rationale for choosing - the functioning of - and the limitations of this platform will also be expressed.

The second section of this chapter contains the actual measurement of *public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software*. This section relates to the first descriptive supplementary research question on measuring *public expressions on democracy for the period of 1 May to 31 July 2012*. This section is aimed at proving the hypothesis that *public expression of democracy between 1 May and 31 July 2012 - according to interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis, and sentiment analysis software – is more negative than positive*.

The third and final section of this study examines the second supplementary research question of this study on the *accuracy of content analysis and sentiment analysis software used in measuring public expressions on democracy between 1 May and 31 July 2012 from Twitter*. This section contains a manual assessment of the Chatterbox classification results by the researcher.

4.2 OPERATIONALISATION

In order to answer the main research question and the two supplementary research questions, a number of the concepts that have been conceptualised in Chapter One need to be operationalised.

4.2.1 Extent

With regards to the main research question, the concept of ‘extent’ needs clarification. The extent to which one can measure public expressions on democracy using the novel methods proposed in this study is measured by looking at the two supplementary research questions. The first descriptive research question illustrates the practicality and possibility of conducting such a measurement of public expressions on democracy. The second supplementary research question aims to demonstrate the accuracy of using such a method. The combination of the results of these two supplementary questions will provide a broad description of the extent *to which one can measure public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software*. What this means is that the two supplementary research questions will be used in answering the main research question.

4.2.2 Public Expressions

For the first supplementary research question the concept of ‘public expressions’ needs to be clarified. The actual measurement method will be discussed in more detail subsequently, but the result of this measurement will provide an aggregate level of positive and negative sentiment for the chosen period as two separate percentage scores. The first percentage will represent the aggregate ratio of negative sentiment within the population of Tweets; the second is similar for the negative Tweets. One will therefore be able to see what percentage of the Tweets is classified as positive and what percentage as negative.

The second supplementary research question measures the accuracy of content analysis and sentiment analysis software used in the first supplementary research question. In order to measure this accuracy, a simple random sample - stratified by positive and negative sentiment of Tweets from the population - will be selected and classified by hand. The researcher will, within the statistical limitations, attempt to provide an estimate of the accuracy of Chatterbox as a measuring instrument by looking at the case study described in the first supplementary research question.

4.2.3 Sentiment Classification

The operationalisation of sentiment classification is based on the O’Connor *et al.* (2010) study. This study used 1 billion Tweets for the period 2008–2009. They made no attempt to identify user location or language. They acknowledge the fact that better stratified data would provide even more relevant results, but did not feel that their exploratory study necessitated further analysis. A similar approach is taken in this study as the results will not be stratified according to location. As stated previously, the results will

only be representative of the 104 037 English Tweets as the Datasift programme has been set to only extract these Tweets.

The Tweets from the O'Connor *et al.* (2010) study were processed in a number of ways. Firstly, topical keywords were extracted manually to represent specific issues such as consumer confidence, elections and presidential approval. For consumer confidence the topical words *economy*, *jobs* and *job* were chosen, for example. This same manual selection method has been implemented in this study. A number of colloquial versions of the word democracy were chosen in order to identify Tweets relating specifically to democracy⁴⁴:

democracy, Democracy, Liberal Democracy, Discursive Democracy, Pluralist Democracy, deliberative democracy, social democracy, demockerycy, deMockracy, democratees, democrabeep, democrackcy, democrack, democrap, democrapsy, democrASS, D3m0cr@cy, plutocracy, demo-crazy.

After extracting topical Tweets the data need to be classified according to sentiment scores. This will indicate whether a Tweet expresses more negative or more positive sentiment. In the O'Connor *et al.* (2010) study a subjectivity lexicon from OpinionFinder was used to train a sentiment analysis programme to extract sentiment from the Tweets. This type of natural language processing is still quite a noisy⁴⁵ instrument of measurement, but O'Connor *et al.* (2010: 4) state that because they are only interested in aggregate sentiment and have so many data points, the errors will mostly cancel each other out.

For this study the Chatterbox Analytics platform was chosen to provide sentiment analysis through its online platform that has been trained with a lexicon library from Mashape (Chatterbox, 2012b). In the next section the rationale of this choice will be examined.

4.2.4 The Measurement Instrument: Chatterbox Analytics Platform

In measuring expressions on democracy from Twitter there are a number of methods that can be used in order to collect and analyse data. With regard to the analysis of the collected data one could create one's own analysis software, as studies such as O'Connor *et al.* (2009, 2010) and Hopkins and King (2010) have done. Another method is accessing an API⁴⁶ that effectively allows one to do the processing online using a programme which a specific company such as Chatterbox Analytics⁴⁷ or ViralHeat⁴⁸ have created. A final method, the chosen method for this study, is to use an online platform that collects analyses and

⁴⁴ For future studies it should be noted that it is unnecessary to add capital letters as the program searches for these automatically. It is also redundant to search for both the term 'democracy' and 'liberal democracy' as the search for 'democracy' will include phrases on liberal democracy.

⁴⁵ "Noisy" in this case refers to the difficulty of computers to understand, analyse and interpret the data.

⁴⁶ As explained in Chapter One

⁴⁷ <http://chatterbox.co/api/>

⁴⁸ https://viralheat.com/developer/sentiment_api

displays the information automatically online. With this type of platform the user needs minimal computer skills as he/she is only required to enter the specific keywords that the platform should use in its data collection and analysis.

There are a multitude of these platforms and programmes⁴⁹ that are capable of collecting and analysing sentiments from social media data. At the MIH Media Lab, as at many other research centres worldwide, computer scientists such as Dirko Coetsee, Stephan Gouws and Peter Hayward are working on machine learning methods in natural language processing that could increase the capabilities and accuracy of the sentiment analysis programmes and platforms (MIH Media Lab, 2012). These programmes and platforms are mainly aimed at market research, allowing brands to see what their customer base is saying about a brand on social media sites. This allows the specific brand to respond to customer complaints and gauge public sentiment on their specific brand instantly. This method is easy to use and is a perfect choice for a user such as this researcher who does not have strong programming capabilities. In this study democracy is treated as such a brand and analysed with one of these platforms, Chatterbox Analytics (Chatterbox, 2012).

The chosen data collection and measurement instrument for this study is therefore the Chatterbox Analytics platform. On 17 August 2011, this company launched its social media analysis website. During October 2011 the company opened beta signup to its analysis platform and invited partners to collaborative case studies using its platform. This platform allows users to gather Tweet data on specific keywords and then uses data and sentiment analysis software to analyse and present the processed data (Chatterbox, 2012b). This user signed up for beta testing on the Chatterbox platform in March 2012.

Below is a screenshot from the Chatterbox platform representing analysis of data for the last week of the chosen period (23/07/2012 – 29/07/2012). The red, grey and green dots in this picture each represent a Tweet. The further to the right the dot is situated, the more positive sentiment classification it has received. The more to the left, the more negative the classification is. On the vertical axis the influence of a Tweet is measured. The higher on the vertical axis a Tweet is, the more times it has been re-Tweeted. It is possible to scroll over each dot and see, in the right hand column, the Tweet, its classification score, the date it was created, the name of the user and a picture of the user, if it is available. At the top of this column one can also see the total sentiment classification scores for the chosen period. For this period one can note that Chatterbox has classified the Tweets as 41.8 percent positive, 19.6 percent negative and 38.5 percent weak sentiment (neutral). The sentiment classification will be discussed

⁴⁹ Such as: www.tweetarchivist.com; sproutsocial.com; www.sentiment140.com; www.tweetfeel.com; smm.streamcrab.com; datasift.com; chatterbox.co.

in more detail, but this means that, according to Chatterbox, more Tweets that contain the word democracy are positive for this week than are negative.



Figure 9: Chatterbox Analytics Platform - Source: (Chatterbox, 2012).



Figure 10: Chatterbox map representation of global tweets on Democracy - Source: (Chatterbox, 2012).

This information is represented in a number of graphic ways, including a world map (Figure 10) showing where the Tweets come from. However, it falls outside the scope of this study to examine the full Chatterbox platform. What is important is to note that the Tweets that have been gathered can be exported into a Microsoft Excel file. This file contains all the collected Tweets for a chosen period, the exact date and time each Tweet was created, a sentiment score of 1 for positive and -1 for negative for each Tweet and a decimal sentiment score ranging from -1 to -0.99999999 etc. up to 0 and to +0.99999999 etc. up to +1. This allows one to see exactly how positive or how negative the platform has classified a specific

Tweet. Below is an example of one such Tweet that has been classified as quite negative (-0.68094) by Chatterbox.

Creation Date	Text	From (User)	Sentiment Label	Sentiment Value
8/12/2012 9:48	America is really bad at democracy. http://t.co/i8nmPjdU	wellslawoffice	-1	-0.68094

A sentiment analysis programme, or “Sentiment Analyser” such as the one contained in the Chatterbox platform, contains a library of positive and negative words which it uses to look for corresponding positive or negative words in each Tweet (Yi and Nasukawa, 2003). The full process of automated sentiment analysis is however much more complex and it falls outside the scope of this study to examine this process⁵⁰. What is important to know is that according to Chatterbox the Tweet above has been classified to be about 68 percent negative.

This researcher has been in correspondence with the co-founder and chief technical officer of Chatterbox Analytics, Dr Stuart Battersby, since the outset of this study. Dr Battersby has been extremely helpful in answering technical questions regarding the Chatterbox platform. This researcher has also sent feedback to Chatterbox in an effort to improve the platform. This study will also be sent to Chatterbox in order to highlight the potential of the political applications of their platform.

On 27 February 2012, the company DataSift launched DataSift - Historics. This platform allows one to gain access to the full Twitter firehose⁵¹ all the way back to January 2010. On top of this social data platform, it can filter the social media information with sentiment analysis software, provide demographic information and online influence, allowing extremely complex filters for gathering information (DataSift, 2012; Rao, 2012). This platform was also considered for this study, but due to budgeting constraints⁵² was not used. Chatterbox was however sufficient for this introductory methodological study. Future studies with larger budgetary allowances might consider using a combination of platforms, including results from DataSift and Chatterbox.

⁵⁰ If interested see: Yi and Nasukawa (2003) or Pang *et al.* (2002).

⁵¹ The Twitter stream API allows one to access ‘streams’ of public data flowing through Twitter. This is done by accessing what Twitter calls the public *firehose*.

⁵² DataSift charges between US\$ 3000 and US\$ 15000 per month. This study would have cost approximately R82 000 if DataSift was used. <http://datasift.com/pricing>

4.3 DATA COLLECTION

As with the data analysis, there are a number of methods in collecting Tweet data. Firstly, an individual can get access to the Twitter firehose stream and examine up to one percent of all created Tweets per processing second. In order to do this one would need a persistent HTTP (hypertext transfer protocol) connection to the low latency connection of Twitter's global stream (Twitter Developers, 2012). This would mean that you would need an uninterrupted connection to the Internet for the chosen timeframe. For this study this was deemed unnecessary as there are dedicated platforms, such as Chatterbox, that can collect the Tweet data.

As stated in the first chapter, the Chatterbox Analytics platform has access to the public Twitter firehose. This means that effectively one percent of all Tweets on Twitter per processing second can be collected and analysed by Chatterbox. As explained above, there are platforms such as DataSift that have access to the full firehose, but in order to answer the research questions posed in this study one percent was fully sufficient.

In order to collect the Tweets for the period of 1 May to 31 July 2012 this researcher had to sign up for a beta testing trial with Chatterbox Analytics. This allowed the researcher full access to the platform and technical assistance. The first step then was to indicate the keywords relating to democracy that Chatterbox would use to search through the Twitter firehouse and collect Tweets containing these words (as described in Chapter One).

The keywords were chosen to represent a broad range of Tweets relating to democracy. The website, Urban Dictionary (2012), was used in an attempt to gather colloquial versions of the word democracy.

After entering the keywords Chatterbox begins its collection and analysis process, constantly searching the Twitter firehose stream for Tweets containing any of the keywords. For the chosen period, Chatterbox collected 104 037 Tweets containing the above-mentioned keywords. This meant that this study now had 104 037 units of analysis relating to democracy.

4.4 REPRESENTING MEASURED EXPRESSIONS ON DEMOCRACY FROM TWITTER

From the Chatterbox Analytics platform 104 037 Tweets were collected for the period 1 May – 31 July. There are a number of ways in which this data can be represented. The onus was on this researcher to find the most appropriate method of representing the analysed data in order to answer the research questions. In this section the process of finding an efficient way of representing the data that has been collected and analysed via Chatterbox is described. From the three methods described below, various conclusions can

be made, but in order to answer the research questions of this study, the simple bar chart representing aggregated sentiment on the population of Tweets was most useful.

4.4.1 Utilising Histograms

The graph below (*Figure 11*) shows a histogram of Chatterbox sentiment classification for the full period of this study. The histogram is broken into 0.025 sentiment increments. On the left of the graph one can see the red negative sentiment and on the right the blue positive sentiment. It is clear to note that Chatterbox has classified more Tweets in the positive range than in the negative range. At first glance it would therefore seem that a majority of the public expression on democracy for the chosen period has a positive sentiment according to Chatterbox. However, this is not the case, as will be explained in the subsequent section on the reliability and the accuracy of Chatterbox as a measuring instrument.

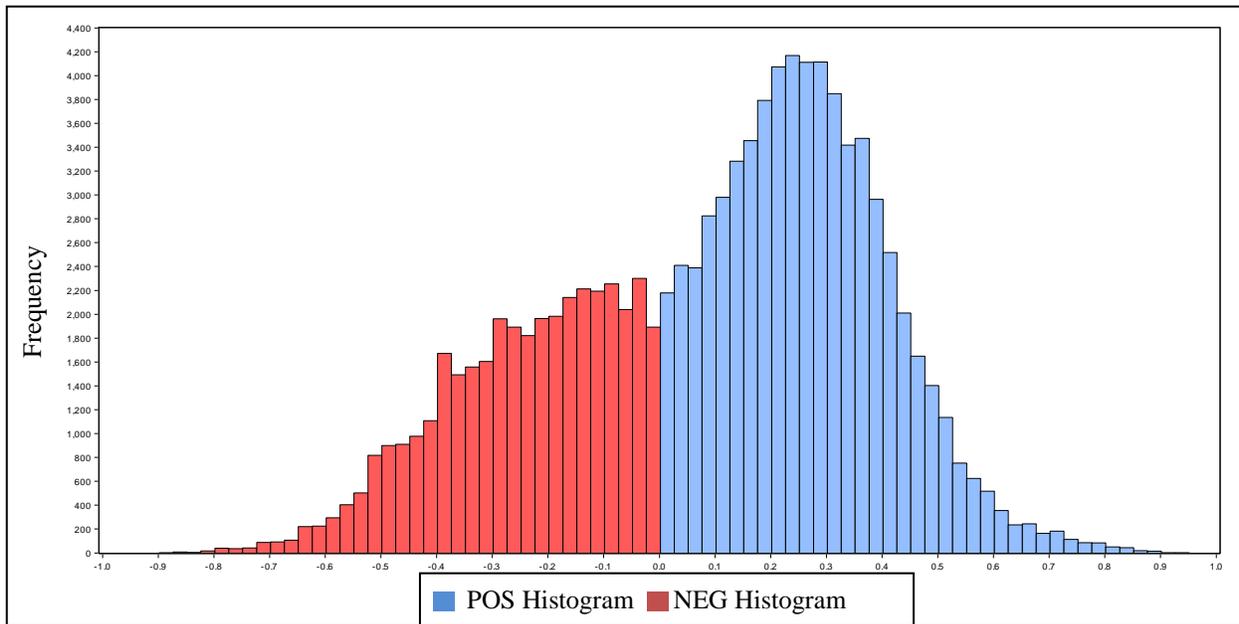


Figure 11: Chatterbox sentiment classification histogram – Adapted from collected data. Source: (Chatterbox, 2012).

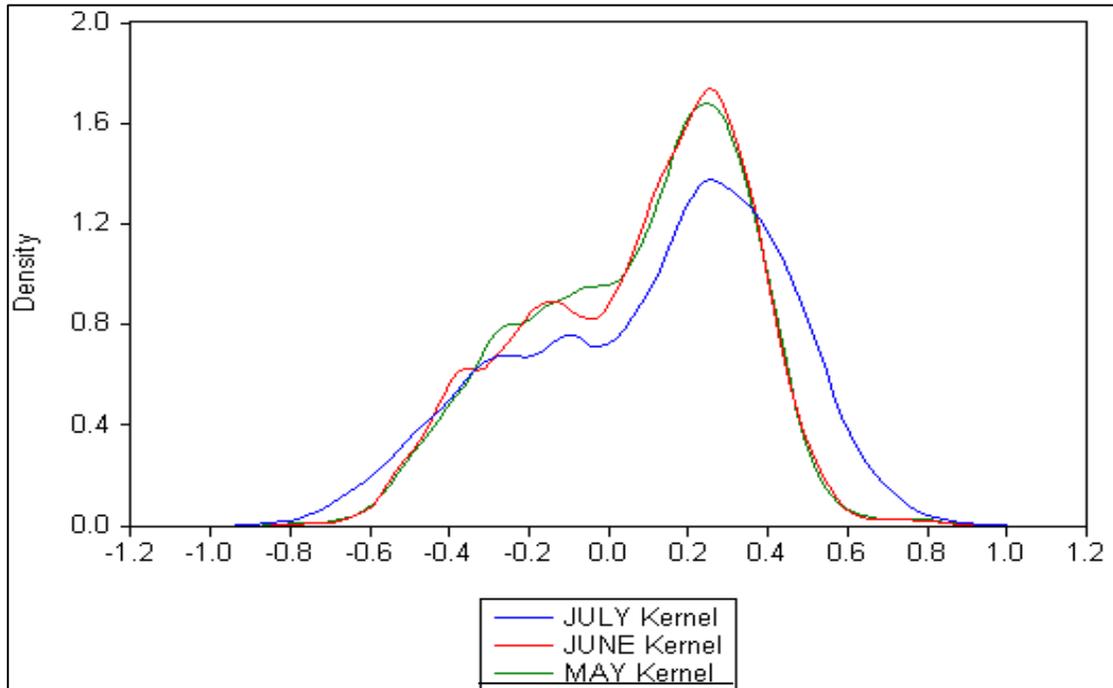


Figure 12: Chatterbox sentiment classification monthly density spread – Adapted from collected data.
 Source: (Chatterbox, 2012).

Figure 12 shows the density spread of the sentiment for each month of the study. One should also note that the shape of each of the three represented months is quite similar. This indicates that the spread of sentiment regarding democracy as represented in the Tweet data for each month is quite similar. One notes two clear peaks in the data. The smaller peak between -0.3 and -0.1 represents the bulk of the negative sentiment Tweets. The bigger peak around 0.2 represents the bulk of the positive sentiments. The spread of the negative Tweets, as seen in both Figure 11 and Figure 12 are broader than the spread of the positive Tweets. This indicates that a wider variety of negative classified Tweets has been classified than positive Tweets. This fact led this researcher to examine the sentiment classification more closely. This researcher wanted to understand whether the shape of these graphs could be explained by the data or whether it was something in the Chatterbox programming that caused this spread. The findings of this closer examination can be found in the subsequent section on the accuracy of the measurement instrument. It should be noted that the shape was in fact explained by the programming and not the data.

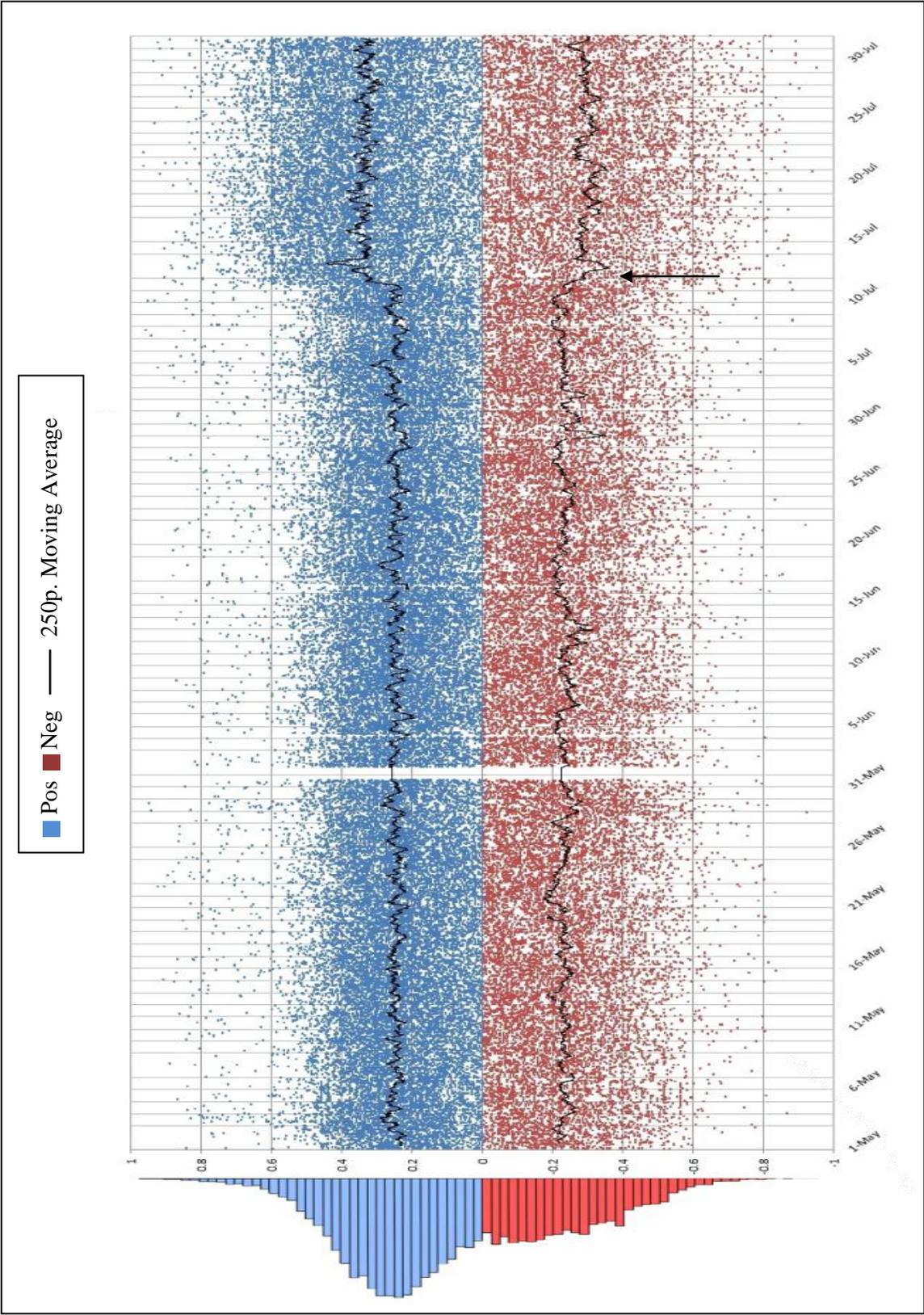


Figure 13: Chatterbox sentiment classification scatter graph – Adapted from collected data. Source: (Chatterbox, 2012).

4.4.2 Utilising Scatter Graphs

Another informative way of representing the sentiment data on the 104 037 Tweets is with a scatter graph. *Figure 13* above shows the sentiment of each Tweet, represented as a blue dot for positive and a red dot for negative, for the full period of this study. A 250 point moving average line has been added to see if there are any interesting movements or trends in the data. There is no clear negative trend in the data, but this was expected as only a short three-month cross section of data has been selected. Future studies might look at a longer period in order to determine whether there is a negative or a positive trend in the data.

There are a number of anomalies in the figure that merit discussion. For example, there are a number of gaps in the data. Gaps exist between 11:52 on 31 May and 14:44 on 1 June; also between 02:51 and 04:54 on 15 June; and between 23:52 on 30 June and 02:49 on 1 July. Another big variation in the figure, which cannot be explained by the data, is the outward shift that the sentiment seems to make around 11 July. These gaps and variations were created when Chatterbox updated its beta programme and will hopefully not exist when the programme is officially launched. This researcher has contacted Dr Battersby from Chatterbox regarding the gaps and the outward shift in the data. Dr Battersby confirmed this researcher's suspicion that these were indeed instances when the Chatterbox platform was updated and changed.

On certain days such as 28 June (indicated by the arrow) there are peaks in the 250 point negative moving average line. After further analysis of a number of these peaks the researcher has found that these tend to indicate Tweets that have been re-Tweeted by a big number of users on a specific day. On 28 June there were two such Tweets that explain the spike in the data. The one reads:

“US has a false democracy, says Mahathir: He says in America they torture people from around the world and that ... <http://t.co/GCXUc0uE>”

This Tweet relates to a Malaysian news story (Yatim, 2012) on how Malay democracy is truer than American democracy. The Tweet was re-Tweeted 32 times within the sample collected for this study.

The next Tweet reads:

“RT @MSignorile: Breaking: Conservatives planning to leave U.S., but can't find wealthy Western democracy without universal health care. #hrc #scotus”

This is a re-Tweet of a Tweet by Mike Signorile, the editor-at-large for the Gay Voices Vertical of The Huffington Post, which comments on the US Supreme Court's decision to uphold the PPACA (Patient Protection and Affordable Care Act), colloquially known as Obama Care, and how this will affect conservatives in America. It was re-Tweeted 174 times within the sample collected for this study.

The days on which a large number of the same Tweet was Tweeted can more clearly be seen on the next three scatter graphs (*Figure 14*). The dots on these three graphs represent all the Tweets that occurred during each month. It is clear to see that some of the dots form straight lines. These lines represent the same Tweet that has been re-Tweeted multiple times over a specific period. One can, for example, see the Signorile Tweet towards the end of the June chart (circled area). Within the Tweet data set, these Tweets were ‘trending’ and were therefore quite topical. ‘Trending’ is a term used on various online sites to indicate a topic that is receiving high levels of attention at a specific time. An example of Twitter trends can be seen on the real time Twitter Trends Map at <http://trendsmap.com/>.

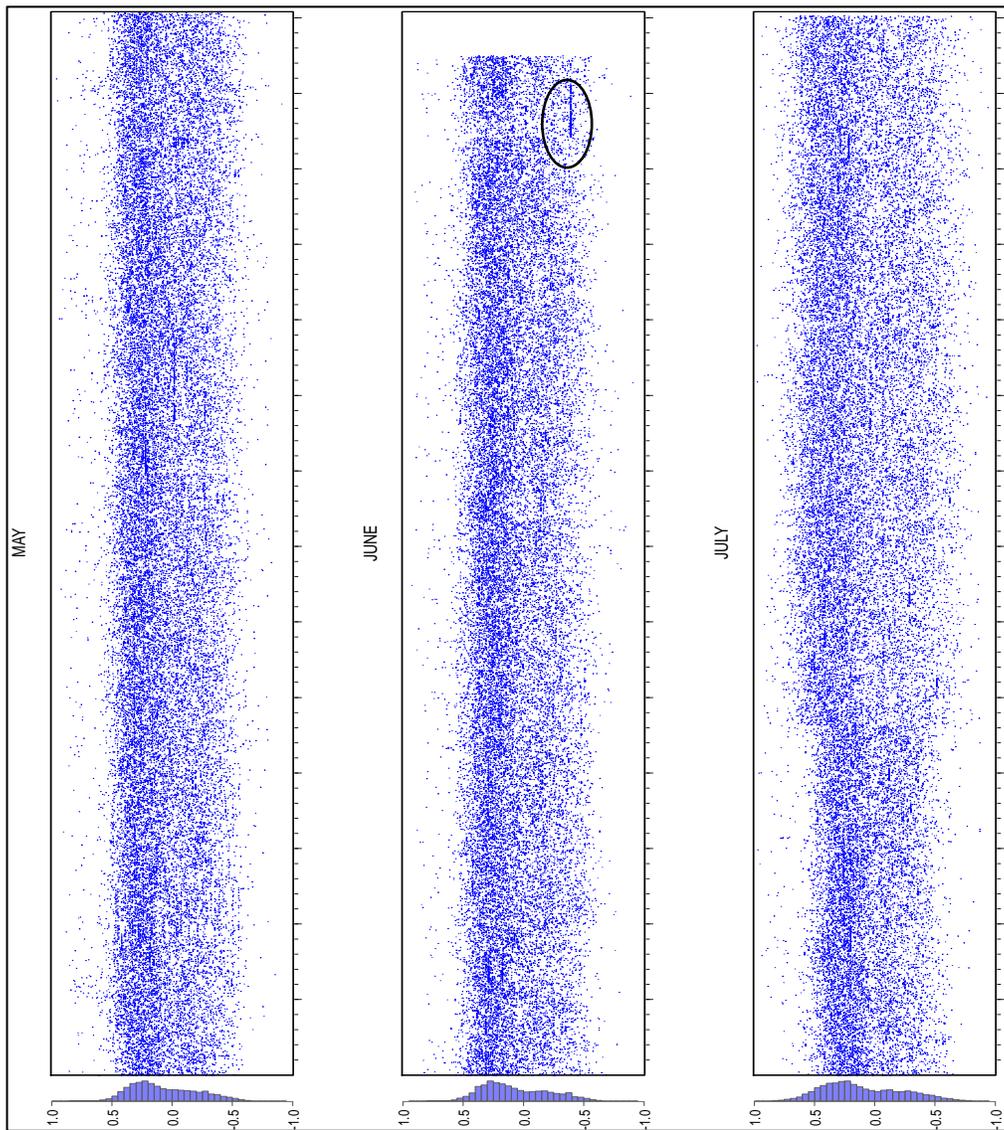
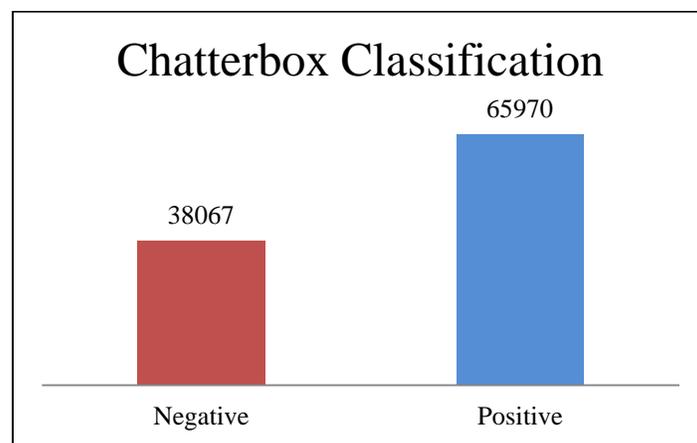


Figure 14: Chatterbox sentiment classification monthly scatter graphs – Adapted from collected data.

Source: (Chatterbox, 2012).

4.4.3 Utilising a Bar Chart

In order to prove or disprove the hypothesis set for the descriptive question of this study, one needs to determine whether the majority of the collected Tweets are positive or negative. In order to do this the small variance between the classified sentiments is not important. What this means is that in order to prove or disprove the hypothesis, a Tweet will only be seen as positive or negative. Any Tweet that is classified between -1 and 0 will therefore be seen as negative and any Tweet classified between 0 and 1 as positive. In the data collected from Chatterbox this broad positive or negative label has already been added. If one aggregates the results by counting the entire set of positive and all the negative Tweets one finds that Chatterbox has classified 36.589867066 percent (30867) of the collected Tweets as negative and 63.410132933% (65970) of the Tweets as positive. These results can be seen in *Figure 15* below.



	May	Jun	Jul	Total
Negative	13239	12357	12471	38067
Positive	21948	21405	22617	65970

Figure 15: Aggregated Chatterbox sentiment classification – Adapted from collected data. Source: (Chatterbox, 2012).

Looking at the figure above it would again seem that public expression on democracy as expressed by the sentiment data is more positive. In fact, the Chatterbox classification shows that in the chosen period there were 27 903 more positive Tweets than negative Tweets about democracy. This means that, using the Chatterbox sentiment classification for the 104 037 collected Tweets, the hypothesis has not been proved. More expressions on democracy found in Tweets for the selected period, classified by Chatterbox are positive than negative.

For the purposes of this research question, public expressions on democracy between 1 May and 31 July 2012, according to interest articulated sentiments gathered from a sample of Tweets, processed with

content analysis and sentiment analysis software, are 36.589867066 percent negative and 63.410132933 percent positive.

This researcher was however sceptical about these results. Knowing that the Chatterbox platform is still in its development stages raised concerns. Furthermore the above-mentioned concerns regarding the strange shape of the density graphs in *Figure 11* and *Figure 12* prompted this researcher to examine the sentiment classification more closely. The next section contains this closer examination.

4.5 ACCURACY OF THE MEASUREMENT INSTRUMENT

4.5.1 Determining the Accuracy of Measurement Instrument

The second supplementary research question of this study aims to examine the accuracy of the measurement instrument used in this study, which in turn can show how accurate the methodology developed in this study is:

How accurate is the content analysis and sentiment analysis software used in measuring public expressions on democracy between 1 May and 31 July 2012 from Twitter?

Even though the measurement instrument, namely the Chatterbox Analytics platform, is broadly treated as a black box in this study, this researcher felt that it was necessary to examine the accuracy and reliability of the instrument a bit more closely. With this in mind the reliability of the Chatterbox Analytics platform was manually tested. This test aimed to determine the level of accuracy that the platform achieves in classifying the collected Tweets about democracy, for the chosen period, as positive or negative.

A random sample of Tweets was collected using a simple random sampling method on the negative and the positive Tweets separately. There were therefore two samples, stratified according to positive and negative classified Tweets. In order to achieve a confidence level of 95 percent and a confidence interval of 4, a random sample of 595 Tweets from the positive population of 65 970 was manually classified. In order to achieve the same level of statistical confidence, a random sample of 591 Tweets of the negative population was also manually classified.⁵³

The manual classification was then conducted by this researcher⁵⁴. Each Tweet in the random sample was examined and then classified to contain a positive, negative or neutral expression on democracy and

⁵³ Required sample size calculated with Creative Research System's Sample Size Calculator (<http://www.surveysystem.com/sscalc.htm>).

⁵⁴ An example of this manual classification can be found in Appendix B. The full classification is also available and can be provided if needed (email bflutz@gmail.com).

assigned a corresponding value 1 for positive, 0 for neutral and -1 for negative. The neutral category was added as the researcher found that many of the Tweets did not contain opinionated data. The researcher is aware that the Chatterbox results did not contain a neutral classification, but as explained earlier treats this as a delimitation of this study.

Elaborating on this delimitations, there are computerised methods of extracting opinionated key phrases from text in existence. An example of such method is examined in Anil Kumar and Suresha (2010). Incorporating such methods would enable future research to more accurately extract only opinionated Tweets, but Chatterbox at present does not include this function. What this means is that some of the Tweets collected by Chatterbox contains opinionated statements. For the purpose of this study, incorporating the neutral classification limits the statistical validity of conclusions that can be drawn from the results. However, the aim of this second supplementary research question is not to prove the hypothesis of the first supplementary question. The aim is to measure the accuracy of the measurement instrument and the most insightful way of doing this is by adding a category for - neutral Tweets - that does not represent opinions on democracy. The fact that the results therefore cannot be extrapolated with full statistical validity - which could prove the hypothesis of supplementary research question - is therefore irrelevant.

4.5.2 Manual Classification of Random Sample

Negative results

Firstly, the 591 randomly selected Tweets from the 38 067 population of negative Tweets were examined. The negative sample showed that 64 percent (378) of the Tweets that were classified as negative by the Chatterbox platform were also negative according to the researcher's subjective opinion. This means that 64 percent of the Tweets were correctly classified according to this researcher. Then seven percent (43) of the Tweets were wrongly classified - which means that they were actually positive according to this researcher, but that Chatterbox classified them as negative. Lastly, 29 percent (170) of the random sample of Tweets did not contain opinionated data and were therefore classified as neutral by this researcher. *Figure 16* illustrates these results in the form of a pie chart.

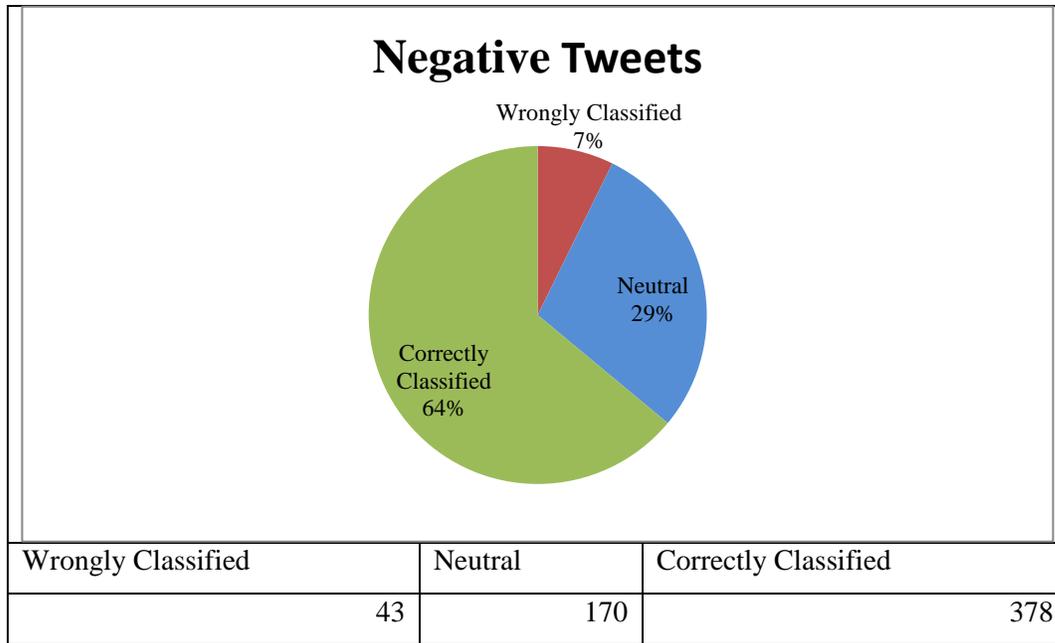


Figure 16: Manual Classification of Negative Chatterbox classified Tweets – Adapted from collected data. Source: (Chatterbox, 2012).

What these results show is that there is a 95 percent chance that all the 38 067 negative Tweets in the population are classified correctly 59 to 68 percent of the time (due to the confidence level and -interval). It is however important to note that this percentage is quite high in comparison to human standards. The seminal study by Gale *et al.* (1992), *Estimating Upper and Lower Bounds on the Performance of Word-Sense Disambiguation Programs*, has shown that when humans classify sentiment statements they tend to agree only about 68 percent on their classifications. Seeing that the Chatterbox platform and this researcher agree between 59 and 68 percent of the time, it means that the programme is between 87 and 100 percent as accurate as using another human to classify the negative Tweet in this study (with a 95 percent confidence level).

Positive results

The 595 Tweets from the 65 970 population of positive Tweets were examined. As expected when looking at *Figure 11* and *12*, the positive sample was not classified nearly as effectively as the negative. The results showed that only 35 percent (209) of the Tweets that were classified as positive by the Chatterbox platform were also positive according to this researcher. Then 30 percent (180) were wrongly classified, meaning that they were actually negative. About 35 percent (206) of the Tweets that were classified as positive by Chatterbox did not contain opinionated data and were therefore classified as neutral by the researcher. *Figure 17* illustrates the above-mentioned results.

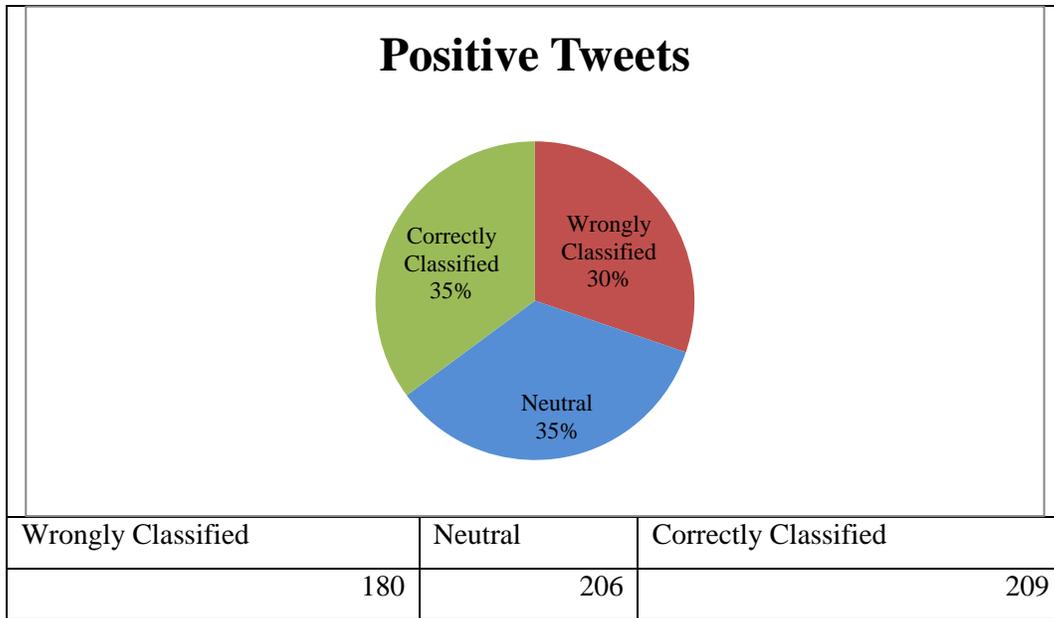


Figure 17: Manual Classification of Positive Chatterbox classified Tweets – Adapted from collected data. Source: (Chatterbox, 2012).

What these results show is that there is a 95 percent chance that all the 65 970 negative Tweets in the population are classified correctly 31 to 39 percent of the time. As explained above, human sentiment classification correlates on average 68 percent. The Chatterbox classification is therefore only between 46 and 57 percent as accurate as using another human to classify positive Tweets (with a 95 percent confidence level).

What does this mean?

The Chatterbox platform therefore preforms well with classifying negative Tweets regarding democracy, but performs poorly with classifying positive Tweets regarding democracy. There are also quite big segments of each selection that are neutral Tweets that do not provide sentiments on democracy or democracy in a specific country.

Further extrapolation

With this new information it is possible to create a closer estimation as to the actual amount of positive and negative Tweets that are contained within the total population of collected Tweets. It is imperative to note that this is however still a subjective estimation and will only represent a range of possibility that the Tweets would have been classified as positive or negative by this researcher. In order to limit this subjectivity, future research might increase the number of participants that classify the Tweet sample.

Furthermore, the results are only representative of the 104 037 collected Tweets and contain limitations to the validity, expressed above, due to the added neutral category.

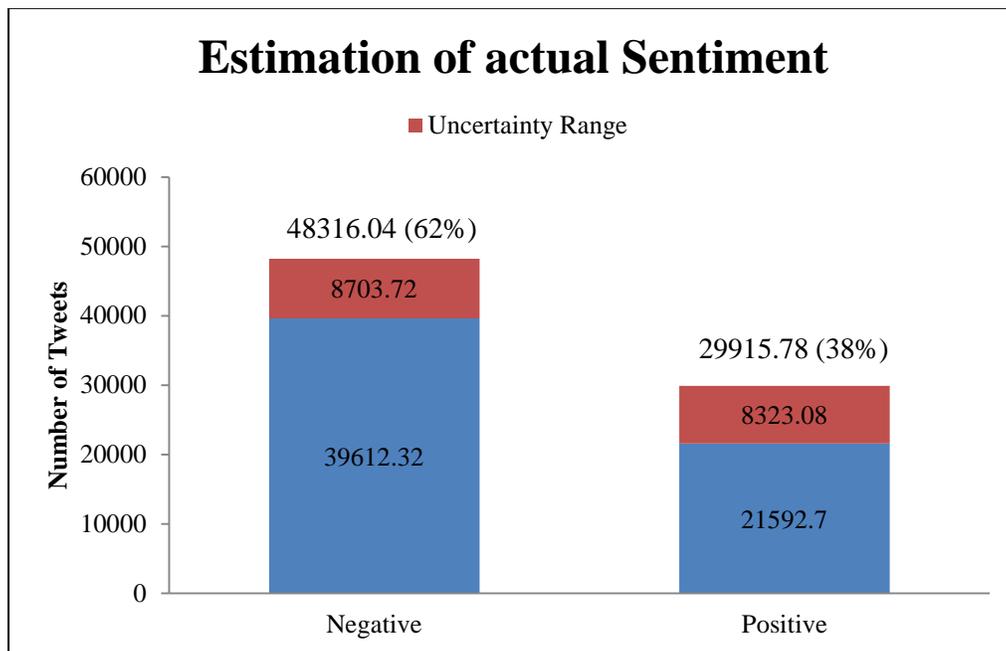


Figure 18: Estimation of Actual Sentiment Levels – Adapted from collected data. Source: (Chatterbox, 2012).

Figure 18 above represents the new estimation on the proportion of the Tweets that are classified as either positive or negative.⁵⁵ What this figure shows is the data calculated above. From this figure one can see that, according to this researcher’s subjective estimations, the actual amount of Tweets for the given period that are negative - are between 39612.32 and 48316.04. The actual amount of positive Tweets is estimated to be between 21592.7 and 29915.78.

Error range in Chatterbox classification

These results show that the Chatterbox sentiment classification, when performed on the selected Tweets regarding democracy, is quite inaccurate. In fact, there is a variance between 3.9 and 21 percent in the negative classification when comparing the Chatterbox results to the manual results of the researcher. The positive classification in turn shows a variance of between 55 and 67 percent when comparing the Chatterbox results to the manual results of the researcher. The calculations for this variance follow below:

⁵⁵ The calculations behind this figure can be found in Appendix A.

Negative variance:

$$100 - (38067/39612.32*100) = 3.90110955379538$$

$$100 - (38067/48316.04*100) = 21.21250003104559$$

Positive variance:

$$100 - (21592.7/65970*100) = 67.26891011065636$$

$$100 - (29915.78/65970*100) = 54.65244808246173$$

Chatterbox was however still in its beta testing stages when this research was conducted and therefore the larger error range is understandable. Furthermore, the Chatterbox platform has not been created specifically for measuring sentiment on democracy. As explained in the delimitations section in Chapter One, the Chatterbox sentiment classification method has not been trained specifically for analysing data on democracy. In future research the researcher might look at training a sentiment analysis programme specifically for politically related data, but this fell outside the scope of this research.

4.6 CONCLUSION

In the first section of this chapter the method of *measuring public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software* was developed and operationalised. The Chatterbox Analytics platform has been utilised in conjunction with methods from the O’Conner *et al.* (2010) study in order to gather and analyse the Tweets needed to answer the research questions posed by this thesis.

This section was followed by a section on the actual measurement and analysis of the chosen population of Tweets using the method developed in the previous section. Various ways of representing the analysed data were highlighted. Finally, a bar chart showing the aggregated ratios of positive and negative sentiment for the chosen period was selected as the most efficient method of representing the data and answering the second supplementary research question. The bar chart showed that 36.589867066 percent of the collected Tweets were classified as negative by Chatterbox and 63.410132933 percent as positive.

There were, however, concerns as to the accuracy of the measurement instrument (Chatterbox). A section on determining the accuracy of the measurement instrument therefore followed. In this section the researcher showed that Chatterbox’s classification of the Tweet data was not perfect. It has been shown that the positive classification by Chatterbox was specifically faulty. With these faults in mind an estimation of the actual sentiment that can be found in the collected Tweet data set was created. This

showed that - if the sentiment classification method was more accurate, the collected Tweets would in fact contain more negative (62 percent) than positive (38 percent) sentiment Tweets on democracy.

Chapter 5: Findings and Recommendations

5.1 INTRODUCTION

In this chapter the findings of this methodological and descriptive study will be given. Firstly, an brief overview of the argument in this thesis will be given. Next the results relating to each of the three research questions will be discussed. Following these three sections, a section on the broad relevance of these findings can be found. There are a number of other interesting findings that have emerged out of this study which will also be briefly examined. The researcher holds the belief that this study has the potential to open up the field of automated measurement of public expressions from online textual data. The next section therefore examines potential routes that future research within this novel field might take. Finally, a short conclusion section can be found.

5.2 BRIEF OVERVIEW OF ARGUMENT

It has been shown in Chapter Two that democracy has followed a route of turbulent development throughout its more than 2500 years of existence. During these years the concept of democracy has evolved and adapted in order to continue to function. If one, for example, considers Huntington's (1991) three waves of democracy it is clear that one of the factors that affects the development of democracy is the public's sentiment or opinion on democracy. This notion that the public's sentiment on democracy influences the development and the functioning of democracies has been discussed in Chapter Two and Chapter Three of this study, but also by authors such as Bratton and Mattes (2001); Diamond (1999); Fails and Pierce (2010); Inglehart (2003), Inglehart and Welzel (2005); Linz (1990); Rose (2001); Rose *et al.* (1998); Shin (2006); Welzel (2002) and countless others.

In this thesis it is argued that the development of democracy follows an adaptive and evolving route that is strongly influenced by sentiment held by the citizenry. It has also been shown in Chapter Two that a critical view on democracy might be on the rise. Van Beek (2012:12) for example speaks of the global erosion in the confidence of democracy. Various arguments for the potential erosion of democracy have been put forward. For example, Diamond (2011) and van Beek (2012) examined the 2008-2009 financial crises as a potential source for negative sentiment on democracy. If this is the case, and one accepts the idea of an adaptive democracy, democracy might have to evolve again in order to adapt to overly negative global sentiments on democracy.

With this potential erosion of confidence in democracy in mind, this researcher decided to examine how public expressions on democracy are being measured (Chapter Three). Seeing that the opinion of the public is paramount to the functioning of democracies, the accuracy and effectiveness/efficacy of measuring expressions on democracy is also vital. It has been shown that traditional methods of measuring public opinions on democracy mostly take the form of research projects, such as the World Values Survey, which uses standardised surveys to ask respondents a broad list of questions. Some of these questions directly relate to democracy and some examine support for the institutions of democracy. These surveys, however, are quite expensive and time consuming to conduct. This is where automated content and sentiment analysis software might offer some help.

All of the traditional surveys on democracy seem to indicate that there is broad support for democracy, that people tend to want to live in democracies and that people feel that even though democracy has certain problems, it is still the preferred system of government. Questions relating specifically to the functioning of democracy in specific countries do show that there are concerns with the functioning of democracy in their countries. The AfroBarometer (2008) for example shows that 65.90 percent of their respondents in 2008-2009 felt that they were either not living in a democracy, or that there were some problems with the democracy within which they lived. Also, about 45 percent of the respondents were not at all or not very satisfied with democracy compared to 53.30 percent that were fairly to very satisfied. These results may indicate a deep-rooted, growing negative sentiment on democracy. It is these types of questions that the method developed in this thesis aims to supplement

Seeing that survey data gathering used in the above-mentioned programmes are difficult and expensive to conduct, this researcher decided to examine alternative methods of measuring expressions on democracy. The goal was to find methods of measuring expressions on democracy that could supplement the traditional survey gathered data. In Chapter Three a number of emerging methods of measuring public expressions using automated computerised methods were examined. Many of these methods utilise Twitter in order to gather expressions on various topics (Bollen *et al.*, 2009; Hopkins and King, 2010; Diakopoulos and Shamma, 2011; O'Connor *et al.*, 2010). This study argues that Twitter can be seen as a form of Habermasian public sphere. From such a public sphere one would be able to extract expressions that might give an indication as to the public opinion of a certain segment of the population. A study by O'Connor *et al.* (2010) was found that linked text sentiment to public opinion. This study formed the backbone for the method developed for measuring expressions on democracy from Twitter (Chapter Four).

In Chapter Four, the main concepts that are used to answer the main research question and the two supplementary research questions were operationalised. Chapter Four also contains a description of the functioning of the data collection and analysis instrument - namely the Chatterbox Analytics platform.

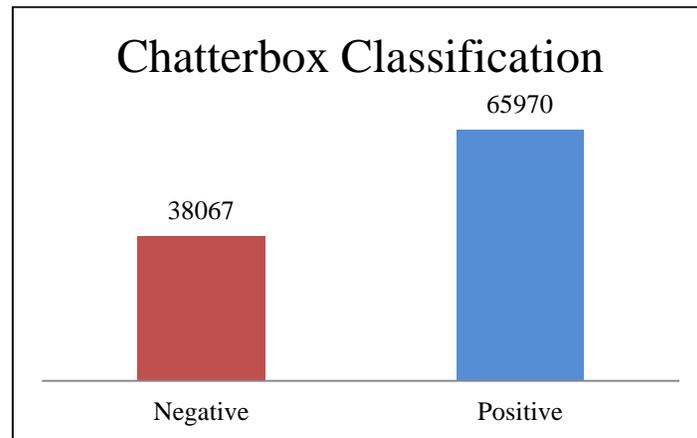
This study aims to answer the main research question with regards to *the extent to which one can measure public expressions on democracy from interest articulated sentiments gathered from a sample of "Tweets", processed with content analysis and sentiment analysis software*. In order to answer this main research question, two supplementary research questions have been asked. The first asks the descriptive question: *What are public expressions on democracy between 1 May and 31 July 2012, according to interest articulated sentiments gathered from a sample of "Tweets", processed with content analysis and sentiment analysis software?* The aim of this question is to provide a sample case for testing the methods of measuring public expressions on democracy from Twitter that is developed in this study. This case illustrates the practicality and feasibility of using the developed method. The second supplementary research question strives to determine the *accuracy of the content analysis and sentiment analysis software used in measuring public expressions on democracy between 1 May and 31 July 2012 from Twitter*. The aim of this question is to determine how accurate the results from the case study, using the methods developed in this thesis, are. In the following sections the results of these three research questions will be discussed.

5.3 CASE STUDY: PUBLIC EXPRESSIONS ON DEMOCRACY FROM TWITTER

Using the Chatterbox Analytics platform, this study measured public expressions on democracy between 1 May and 31 July 2012, according to interest articulated sentiments gathered from a sample of "Tweets". Chatterbox managed to collect 104 037 units of analysis (Tweets) for the given period that related to democracy⁵⁶. The automatic sentiment classification programme embedded in the Chatterbox platform determined that 36.589867066 percent (30867) of the collected Tweets contain negative sentiment and 63.410132933 percent (65970) of the Tweets contain positive sentiment. These results are represented by *figure 18*, which is a replication of *figure 15*. According to this study's conceptualisation of public expressions on democracy, the aggregated negatively classified sentiment is seen as a broad negative expression on democracy, and the aggregated positively classified sentiment is seen as a broad positive expression on democracy.

⁵⁶ The full dataset is available and can be obtained by sending the researcher an email requesting a copy. (Email: bflutz@gmail.com)

The findings of this study therefore show that according to the described classification method, public expression on democracy is positive. This does not prove the hypothesis that *public expression of democracy between 1 May and 31 July 2012, according to interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis, and sentiment analysis software, is more negative than positive.*



	May	Jun	Jul	Total
Negative	13239	12357	12471	38067
Positive	21948	21405	22617	65970

Figure 19: Aggregated Chatterbox sentiment classification – Adapted from collected data. Source: (Chatterbox, 2012).

The relevance and the significance of these findings are multifold. Firstly, even though these results did not prove the hypothesis, it illustrates the possibility and practicality of measuring public expressions on democracy from Twitter using automated computerised methods. The study was indeed able to aggregate public expressions on democracy for the chosen timeframe by collecting and analysing Tweets. Secondly, this method, as explained in Chapter Four, was uncomplicated as it did not require the researcher to possess programming skills. It allowed the researcher to examine a constant⁵⁷ flow of Tweets regarding democracy that represents one percent of all the Tweets from the more than 500 million users on Twitter. It should be noted, as stated earlier, this study did not aim to extrapolate to all 500 million users and thus the analysed data are only statistically representative of the 104 037 Tweets.

As discussed in Chapter Four, this researcher did however have a number of concerns with the results. These concerns related to the accuracy of the sentiment classification by the Chatterbox measurement

⁵⁷ Except for small gaps in the data as explained in Chapter One.

instrument. These concerns were addressed by evaluating the accuracy of the results of this study and are discussed in the next section of this chapter.

5.4 ACCURACY OF MEASUREMENT INSTRUMENT

In order to evaluate the accuracy of the Chatterbox platform as the measurement instrument used in this study, this researcher manually examined the results of the sentiment classification described above. The aim of this evaluation was to answer the second supplementary research question on the accuracy of *the content analysis and sentiment analysis software used in measuring public expressions on democracy between 1 May and 31 July 2012 from Twitter*. This researcher classified a representative sample of Tweets that were collected using simple random sampling and that were stratified according to positive and negative sentiment. This sample of 591 negative Tweets and 595 positive Tweets allowed the researcher to extrapolate, within a confidence interval of 4 with a 95 percent confidence level, on the 104 037 population of Tweets. In order to provide a more exact evaluation of the accuracy of the measurement instrument, the researcher decided to add a neutral category of Tweets. This neutral category contained Tweets that did not represent opinionated data on democracy.

As explained in Chapter Four, the aim was to measure the accuracy of the measurement instrument and the most insightful way to do this was by adding a category for neutral Tweets that do not represent opinions on democracy. This meant that the results therefore could not be extrapolated with full statistical validity and compared to the Chatterbox results. If this could be done it would have been possible to prove the hypothesis of the supplementary research question, described above. This was however not the goal of this research question; which rather focuses on the accuracy of the measurement instrument.

The results of this evaluation showed that Chatterbox was capable of quite accurately classifying the negative Tweets relating to democracy. The Chatterbox classification of negative Tweets corresponded between 59 and 68 percent with this researcher (with a 95 percent confidence level). Seeing that, according to Gale *et al.* (1992), humans normally only correspond about 68 percent on classification of sentiment, this meant that the Chatterbox platform was between 87 and 100 percent as accurate as using another human to classify negative Tweets.

The evaluation of the positively classified Chatterbox Tweets, however, painted a different picture. The results showed that only between 31 and 39 percent of the positive Tweets - as classified by Chatterbox - corresponded with the researcher's classification (with a 95 percent confidence level). The Chatterbox classification was therefore only between 46 and 57 percent as accurate as using another human in classifying positive Tweets.

The weak results of the positive classification by Chatterbox compelled the researcher to attempt to estimate an actual level of positive and negative sentiment that is contained in the collected Tweet data on democracy. In order to accomplish this, the results were tallied (fully explained in Chapter Four) and then extrapolations on the 104 037 Tweets were made according to the above-mentioned levels of statistical validity. The results of this estimation showed that, if the Chatterbox platform was more accurate in classifying sentiment on democracy, the results would have shown that the actual amount of Tweets containing negative sentiment was between 3961.32 and 48316.04 (about 62 percent) and that the positive was between 21592.7 and 22915.78 (about 38 percent) (with a 95 percent confidence level). These results can be seen in *Figure 19* below.

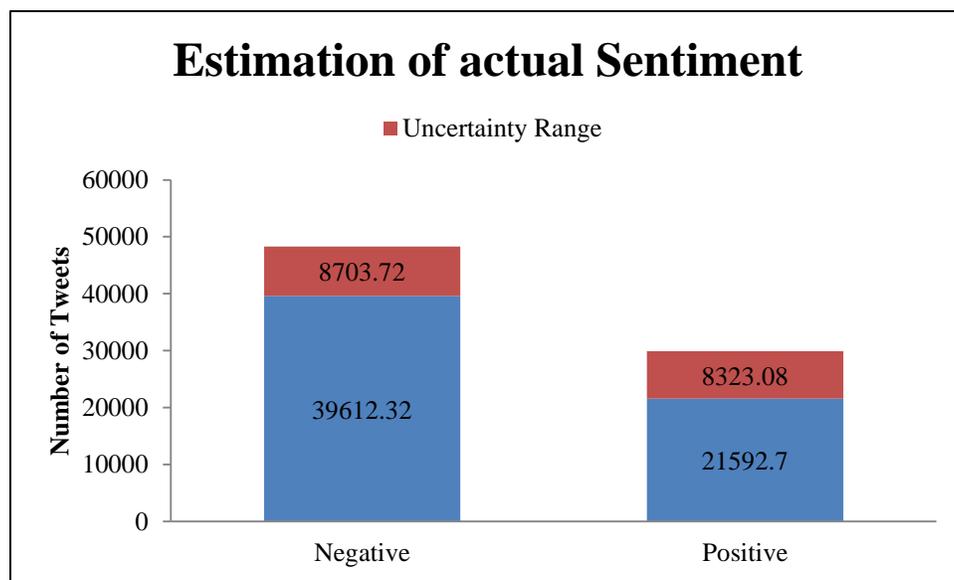


Figure 20: Estimation of Actual Sentiment Levels – Adapted from collected data. Source: (Chatterbox, 2012).

In answering the second supplementary research question, the accuracy of measuring expressions on democracy with the described methods in this thesis, for the chosen period, is therefore limited. The results showed that there existed a variance of between 3.9 and 21 percent in the negative classification, when comparing the Chatterbox results to the manually classified results of the researcher. The positive classification in turn shows a variance of between 55 and 67 percent when comparing the Chatterbox results to the manual results of the researcher. The relevance of these findings was that it highlighted the inaccuracy in the measurement instrument. Seeing that Chatterbox is still in its beta testing phase and the fact that the methods of automated sentiment analysis are still novel, these results are understandable.

With the present level of technology (Chatterbox platform) the accuracy of measuring expressions on democracy for the chosen period is therefore limited. In the prospects for future research section of this chapter, methods for increasing this level of accuracy are discussed.

5.5 FEASIBILITY OF MEASURING PUBLIC EXPRESSIONS ON DEMOCRACY FROM TWITTER

The main research question of this thesis asked the question as to the extent to which it would be possible *to measure public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software?* The conceptualisation section in Chapter One and the operationalisation section in Chapter Four explained how the ‘extent’ is measured. This study determines the extent by looking at the two supplementary research questions.

What this thesis therefore finds, as shown in section 5.3, is that it is possible to measure public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software. The method of using the Chatterbox platform to gather and analyse Tweets for the chosen period was therefore successful.

The evaluation of the measurement instrument’s accuracy, as explained in section 5.3, was however limited. It was shown that the automated classification of negative Tweets was almost as effective as using a human. However, this method used was much faster than using a human as these results are calculated almost instantly by the Chatterbox platform. However, there are concerns with the level of accuracy obtained in the classification of the positive Tweets. The Chatterbox platform seems to have problems with classifying positive sentiment in the collected Tweets relating to democracy.

What this means is that - if one views the Chatterbox platform as the current level of technology for automated sentiment analysis of Tweets relating to democracy - the level of accuracy is limited. These methods are being improved constantly as Chatterbox is still in its infancy. There are also other platforms and programmes that could be used in future studies which will subsequently be examined.

The concept of extent is, however, somewhat subjective and in this case will be examined as a qualitative expression by the researcher. Seeing that it is possible to measure public expressions on democracy with the above-mentioned methods - but that there are problems with the accuracy of the measurement instrument - this study finds that: the extent to which one can measure public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software, is limited.

The closest quantitative explanation that this study can give as to this limitation, as expressed above, is that the results showed that there existed a variance of between 3.9 and 21 percent in the negative classification when comparing the Chatterbox results to the manual results of the researcher. The positive classification in turn shows a variance of between 55 and 67 percent when comparing the Chatterbox results to the manual results of the researcher.

What this means is that, using Chatterbox to measure negative public expressions on democracy, would give between 79 and 96.1 percent the same results as if this researcher would manually read through all 104 037 Tweets. For the positive public expression classifications, Chatterbox would give between 33 and 45 percent the same results as this researcher. The ‘extent’ of the method developed in this study is therefore limited to the level of technology that is available to the researcher. If these methods are improved the extent will also improve. As expressed throughout this study, these methods are indeed being developed and improved.

This researcher does however believe that even if the extent of measuring public expressions on democracy is limited, it can still be a helpful tool in supplementing survey data. The relatively high level of accuracy in measuring negative sentiment on democracy from the Tweet data can, for example, give a researcher insights as to what the main issues with democracy within a specific section of the population are. This is discussed in more detail in the subsequent section on future research.

5.6 OTHER POTENTIAL (INTERESTING) FINDINGS

Apart from answering the research questions, this study has also found a number of other interesting findings. These findings will however be limited due to the length restrictions of this study and will therefore only be discussed briefly.

Firstly, as shown in Chapter Four, one is able to see the effect of certain real-world events on the collected Tweet data. It has been shown how certain Tweets are re-Tweeted and become trending topics that can be clearly seen on a scatter graph representing Tweet sentiment. These trending Tweets might be a good starting point for future researchers to examine topical issues within a set of Tweets.

Secondly, it is interesting to note that there is no clear trend apparent when one views the sentiment data for Tweets on democracy over the period of three months. Prior to this study the researcher believed that a trend would emerge even if such a short cross-section was chosen. If future researchers were to look for a trend in the sentiment data, this researcher would recommend that a longer period be chosen.

Thirdly, a number of anomalies in the data emerged when the data was plotted on a scatter graph. It was clear to see that there were gaps in the data and outward shifts in the classification of sentiment. This led

the researcher to contact Chatterbox in order to confirm that these were changes in the platform and not in the data. These anomalies affected this introductory study minimally, but it is definitely something that future studies should look out for if using a platform such as Chatterbox.

From the literature on the development of democracy (Chapter Two), it was also interesting to note the apparent ‘evolution of democracy’ seems to be strongly influenced by public opinion on democracy. This opinion, according to Habermas (1989), is to an extent created in the public sphere. If one views Twitter as this type of public sphere, as this study does, it becomes possible for the functioning of democracy to be effected by what people are saying on Twitter. Even though the researcher knew that people were talking about democracy and that some of these opinions were negative, it was fascinating to see the complexity of some of the expressions and the dialogue on democracy that is found on Twitter. It was also interesting to see how much information one could gather from these Tweets.

An example is the Tweet by Michelangelo Signorile on 28 June that was re-Tweeted more than 176 times within the collected data. Here is an example of a re-Tweet of the original by RH Reality Check:

“RT @rhrealitycheck: RT @MSignorile: Breaking: Conservatives plan 2 leave U.S., but can't find wealthy Western democracy w/ouniversal health care. #hrc #scotus”

RH Reality Check “is an online community and publication serving individuals and organizations committed to advancing sexual and reproductive health and rights” (RHRealityCheck, 2012). They have more than 12 000 followers which means that this re-Tweet was potentially seen on 12 000 accounts. On 28 June @rhrealitycheck Tweeted another 49 times about the US federal court ruling relating to the PPACA. All of this information and other potential information such as the location of the person Tweeting can be gained from looking at the Tweet data.

The next additional finding of this study relates to the colloquial versions of the term ‘democracy’ that were used in this study. After the analysis of the Tweet data, it has been found that the words “demockerycy, democratees, democrabeep, democrackey, democrack, democrapsy, democross D3m0cr@cy and demo-crazy were not used in any of the collected Tweets. Also, the Chatterbox platform does not look at capital letters and therefore searching for ‘Democracy’ and ‘democracy’ is redundant. Lastly, it was also redundant to search for both terms such as ‘democracy’ and ‘liberal democracy’, seeing that the platform would find all versions of ‘democracy’, including ‘liberal democracy’ from the keyword ‘democracy’.

An observation that this researcher has made regarding expressions on Twitter should also be mentioned. It would seem that as with news media, where negative news receives priority and more coverage

(Gieber, 1955), Twitter also contains more negative news than positive. The ramifications of this observation have however not been tested in this study and the researcher is unaware of studies that have explicitly shown that people speak more negatively on Twitter. If this however is the case, it might in part explain why the final estimation of expressions on democracy in this study show that more people are expressing negative sentiment than positive.

Similar to the findings in these studies, one can, for example, look at The Twitter Political Index. This is a website created by Twitter that shows the percentage of positive sentiment Tweets that are being Tweeted regarding the 2012 US presidential candidates, Barack Obama and Mitt Romney. The results of this website show that both Obama and Romney, on average, score well below 50 percent. What this means is that Tweets concerning both Obama and Romney are more negative than positive. For example, on 17 September 2012 Obama only received 20 percent positive Tweets and Romney only 9 percent (Elections.Twitter, 2012). If one compare sentiment to approval and look at traditional opinion polls one generally note a much higher approval rating (around 50 percent for both candidates) than seen in the sentiment data (Real Clear Politics, 2012; Zeleny, 2012; CNN, 2012). This might indicate that people express more negative sentiment on Twitter than they do in traditional opinion polls. This researcher believes that future studies on the general level of negative sentiment on social media sites such as Twitter is warranted. These types of studies would show if sentiment on Twitter in general is more negative and if so why this is the case. It would also benefit the results of this study as it would give one a better understanding on sentiment statements collected from Twitter. In the next section more ideas for future research is expressed.

5.7 FUTURE RESEARCH

As stated throughout this thesis, this is an introductory study developing and testing the methodology of measuring public expressions on democracy from Twitter. This research therefore aims to open up the field for future research on measuring public expressions from social media sites. Due to the introductory nature of this research and the limitations of a master's thesis, it is impossible to look at all the potential uses of measuring public expressions from social media sites. In the next section further potential analyses using variations of this study are examined. The goal of this is to provide an overview of what this researcher has learnt writing this thesis - which might be of benefit to future studies. Throughout this study reference has been made to potential future studies that can flow out of this study or that might benefit this study. The next section will reiterate and expand on some of these ideas.

Democracy as a topic

Firstly, this study focused on the topic of democracy, but the study could be replicated with any topic of a researcher's choosing. If people are discussing the topic on Twitter, the methodology described in this thesis could be used in order to extract and analyse expressions on that topic as well.

Data from Twitter

Next, Twitter was chosen as the online social media network in focus for this study. The reason for this is that Tweets are relatively freely available in the public sphere and a large section of the global population is represented by these Tweets. Furthermore, the Tweets are short which made it easier to analyse than say looking at full blogs. It would however be possible to extract information from other social media sites such as Facebook, look at blogs and even extract information from mobile social networks such as Mxit, if one can gain access to the data. DataSift for example can provide one with access to a much larger range of social media sites, but this option is quite costly. The automated sentiment analysis method used in this study is however only limited to textual data at present. This means that any textual data that a potential future researcher can gather could be analysed using sentiment analysis software.

Deeper analysis of the data

There are a number of ways in which a future researcher can get an even deeper understanding of the collected data. This study focused only on the sentiment as classified by Chatterbox and the researcher's subjective analysis. Tweets, however, contain much more than just sentiment, as has been explained in Chapter One.

One could, for example, look at the location identification tags that are attached to some of the Tweets. This would make it possible for future research to examine expressions from a specific area of the globe. It should however be stated that at present a user has to activate his/her location setting personally in order for it to be seen in the public realm.

The Tweet data of course also contains the Tweets themselves. The researcher could extract quite a lot of information from the Tweet data with further quantitative or qualitative methods. On the quantitative side for example, the researcher might count the number of words used in all the Tweets with a simple programme. As an example of this, the researcher has extracted some of the most used terms and some interesting words relating to 'democracy' and automatically counted the number of times that it was mentioned in the collected Tweets. With this method one can see what the central concepts that are being discussed in conjunction with the chosen topic are.

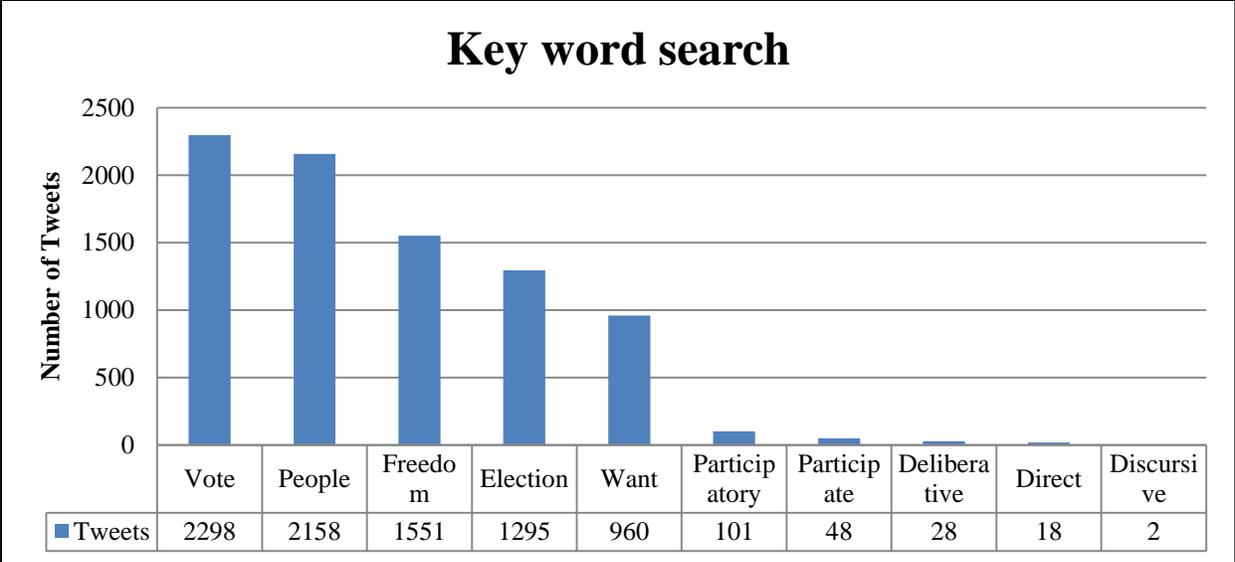


Figure 21: Example Graph: Key Word Search – Adapted from collected data. Source: (Chatterbox, 2012).

These most common concepts can then be used in order to guide the researcher in his/her qualitative analysis. This qualitative analysis would for example look at what is being said within the Tweets themselves. To aid the researcher in this, an automated topic extraction programme might be used. These programmes, as discussed in Chang *et al.* (2002) and Tiun, Abdullah and Kong (2001), attempt to extract the most relevant topics from a set of data. This researcher has attempted a rough topic extraction by hand. This can be found in appendix D. What I did was to look at the most common themes that I thought were relevant to the topic of democracy. From the collected Tweet data it seems that, for example, a lot of Tweets look at how democracy is being devalued by corporate interests and money politics. The next Tweet is an example of this type:

“RT @sunshineejc: Why do Americans Allow corporations to co-opt our democracy by granting them the rights of legal personhood and defining money as speech #p2”.

Another theme that I extracted was problems with the functioning of democracy with Tweets such as:

“RT @DianeRavitch: When we lose neighborhood schools, we lose community. We lose a building block of democracy where people learn to take action. #soschat”.

Lastly, I extracted a topic that contains Tweets that are simply blatant negative sentences regarding democracy, containing Tweets such as:

“RT @Drudge_Report: Sobbing man: 'Democracy died tonight'... <http://t.co/jdFY5XKj>”.

Using this type of qualitative analysis the researcher might be able to extract the issues that are most prominent with regards to the specific topic, in this case ‘democracy’.

Using more advanced data collections and analysis methods

Future studies might also benefit from including more advanced data collection and analysis methods. The use of natural language processing (NLP) methods would ensure that the sentiment analysis software can more accurately analyse the collected data. NLP software would convert the Tweet data into natural language that is more easily understandable to the sentiment analysis software. NLP is a technique, harnessing the computational capabilities of computers, which analyzes naturally occurring texts and process these texts, using one or more levels of linguistic analysis, in order to produce usable and congruent (proper) text. In future studies NLP could be used in order to turn tweets such as:

*“Wn We Spk of "Democracy" in Wch Rt.2 Vote is included then Obviously RT.2 REJECT shud b
CONSTRUED As A Concurrent Right !!! “ (@advocatejaya)*

Into this:



When we speak of democracy, in which the right to vote is included, then obviously the right to reject should be construed as a concurrent right!

This would allow the sentiment classification software to much more accurately classify the Tweets as it would be able to ‘understand’ the sentence better.

Finally, as mentioned throughout this study, the Chatterbox platform is in its beta stages of development. More advanced sentiment analysis software, especially software that has been specifically trained to examine political data, would vastly improve the results of this study. Another recommendation for researchers that would like to conduct this type of study on a larger scale would be to get access to the full Twitter stream. This would allow one to make more direct conclusions regarding the full population of Twitter users than found in this study. Furthermore, researchers with advanced programming knowledge might consider the automated non parametric content analysis method designed specifically for the social sciences by Hopkins and King (2010).

5.8 CONCLUSION

The intention of this concluding chapter has been, primarily, to answer the research questions posed at the outset of this study. This final chapter builds on the argument that has been developed throughout this thesis. The findings of this study shows that with the current level of technology, measuring public expressions on democracy from interest articulated sentiments gathered from a sample of “Tweets”, processed with content analysis and sentiment analysis software, is limited. As technology increases, these limitations are minimised. Although some limitations exist, there are however still interesting uses for evaluating public expressions from social media data. It has, for example, been shown that the analysis of negative sentiment Tweets for the chosen period was almost as accurate as human classification. The method for measuring public expressions on democracy from Twitter developed in this study therefore does show potential and will improve as technology develops. This chapter has additionally examined how this introductory study on novel methods for measuring public expressions on democracy can be improved and developed in future research.

Bibliography

- Adamic, L. A., and Glance, N. 2005. *The Political Blogosphere and the 2004 U.S. Election: Divided They Blog*. Proceedings of the 3rd International Workshop on Link Discovery: 520-522.
- AfroBarometer, 2008. The quality of democracy and governance in Africa: new results from Afrobarometer round 4. *Afrobarometer Network*. Working Paper No. 108. Online. Available: http://www.afrobarometer.org/files/documents/working_papers/AfropaperNo108.pdf. [16/08/2012].
- AfroBarometer, 2008a. *AfroBarometer Online Data Analysis*. Online. Available: <http://www.jdsurvey.net/afro/AnalyzeSample.jsp>. [30/05/2012].
- Almond, G.A. and Verba, S. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. London: SAGE Publications Ltd.
- Arendt, H. 1963. *On Revolution*. London, Macmillan.
- Aristotle. 1981. *The Politics*. Harmondsworth: Penguin.
- Arrow, K. 1951. *Social Choice and Individual Values*. New York: Wiley.
- Asur, S. and Huberman, B.A. 2010. Predicting the Future with Social Media. *Computers and Society*. Cornell University Computer Science Paper.
- Babbie, E. 2010. *The Practice of Social Research*. Belmont: Wadsworth.
- Babbie, E.R. and Mouton, J. 2001. *The Practice of Social Research*. Cape Town: Oxford University Press.
- Barber, B. 1996. *Foundationalism and Democracy*. In S. Benhabib, ed., *Democracy and Difference: Contesting the Boundaries of the Political*. Princeton, Princeton University Press.
- Berg-Schlosser, D., 2012. *The Impact of the Great Depression on Democracy*. In: van Beek, U. and Wnuk-Lipinski, E., (eds.) 2012. *Democracy Under Stress: The Global Crisis and Beyond*. Berlin: Barbara Budrich Publishers.
- Berinsky, A.J. 1999. The Two Faces of Public Opinion. *American Journal of Political Science*. 43(4): 1209-1230.
- Bollen, J., Pepe, A. and Mao, H. 2009. *Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena*. In Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM 2011), 17-21 July 2011, Barcelona, Spain.

- Bollen, J., Mao, H. and Zeng, X. 2011. Twitter mood predicts the stock market. *Journal of Computational Science*. 2(1): 1-8.
- Bratton, M. and Mattes, R. 2001. Support for Democracy in Africa: Intrinsic or Instrumental? *British Journal of Political Science*. 31(3): 447-474.
- Carey, J.W. 1995. *The Press, Public Opinion, and Public Discourse*. In: Glasser, T.L. and Salmon, C.T. (eds.) *Public Opinion and the Communication of Consent*. New York: The Guilford Press.
- Casper, G. and Taylor, M.M. 1996. *Negotiating Democracy: Transitions from Authoritarian Rule*. Pittsburgh: University of Pittsburgh Press.
- Castells, M. 2008. The New Public Sphere: Global Civil Society, Communication Networks, and Global Governance. *The Annals of the American Academy of Political and Social Science*. 616: 78-94.
- Census, 2012. World PopClock Project. *United States Census*. Online. Available: <http://www.census.gov/population/popclockworld.html>. [09/03/2012].
- Chang, J., Lee, J.W., Kim, Y. and Zhang, B. 2002. *Topic Extraction from Text Documents Using Multiple-Cause Networks*. School of Computer Science and Engineering, Seoul National University.
- Chatterbox, 2012. *Home Page*. Online. Available: <http://chatterbox.co/>. [10/03/2012].
- Chatterbox, 2012a. *Chatterbox Sentiment Detection API*. Online. Available: <http://chatterbox.co/api/>. [10/08/2012].
- Chatterbox, 2012b. *Blog Page*. Online. Available: <http://chatterbox.co/blog/?page=2>. [02/8/2012].
- Chrome, 2012. Translate a page to your language. *Chrome home help*. Online. Available: <http://support.google.com/chrome/bin/answer.py?hl=en&answer=173424>. [09/12/2012].
- Clemens, R. 2012. Twitter as a Public Sphere? *21st Century Scholar*. Online. Available: <http://21stcenturyscholar.org/2012/06/05/twitter-as-a-public-sphere/>. [20/08/2012].
- CNN, 2012. *CNN ORC Poll*. Online. Available: <http://i2.cdn.turner.com/cnn/2012/images/09/10/re110a.pdf>. [18/09/2012].
- CommGAP, 2012. *The Public Sphere*. Communication for Governance & Accountability Program, The World Bank. Online. Available: <http://siteresources.worldbank.org/EXTGOVACC/Resources/PubSphereweb.pdf>. [20/08/2012].
- Crozier, M., Huntington, S.P. and Watunaki, J. 1975. *The Crisis of Democracy: Report on the Governability of Democracies to the Trilateral Commission*. New York: New York University Press.

- Dahl, R.A. 1989. *Democracy and its Critics*. New Haven. Yale University Press.
- Dahl, R.A. 1998. *On Democracy*. New Haven: Yale University Press.
- Dahl, R.A. 1972. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.
- Dahlberg, L. 2001. The internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, Communication & Society*. 4(4): 615-633.
- Dalton, R. 2004. *Democratic Challenges, Democratic Choices: Political Support in Advanced Industrial Societies*. Oxford: Oxford University Press.
- Dalton, R. and Shin, D. (eds.). 2006. *Citizens, Democracy and Markets around the Pacific Rim*. Oxford: Oxford University Press.
- DataSift, 2012. *Unlock Trends from Public Tweets*. Online. Available: <http://datasift.com/historics/>. [31/07/2012].
- Destutt De Tracy, A.L.C. 1811. *A Commentary and Review of Montesquieu's Spirit of Laws: Prepared for Press*. New Jersey: The Lawbook Exchange, Ltd.
- Diakopoulos, N.A. and Shamma, D.A. 2010. *Characterizing Debate Performance via Aggregated Twitter Sentiment*. CHI 2010: Proceedings of the 28th international conference on Human factors in computing systems: 1195-1198.
- Diamond, L. 1993. *Political Culture and Democracy in Developing Countries*. Boulder: Lynne Rienner Publishers.
- Diamond, L. 1999. *Developing Democracy: Toward Consolidation*. Maryland: John Hopkins University Press.
- Diamond, L. 2011. Why Democracies Survive. *Journal of Democracy*. 22(1): 17-30.
- Diamond, L., Linz, J.J., and Lipset, S.M., 1995. *Politics in developing countries: Comparing experiences with democracy*. 2nd ed. Colorado: Lynne Reiner.
- Diamond, L., Plattner, M.F. and Chu, Y.T.H. 1997. *Consolidating the Third Wave Democracies*. Baltimore: Johns Hopkins University Press.
- Downey, J. and Fenton, N. 2003. New Media, Counter Publicity and the Public Sphere, *New Media & Society*. 5(2): 185-202.

- Dryzek, J. S. and Dunleavy, P. 2009. *Theories of the Democratic State*. Great Britain: Palgrave Macmillian.
- Dugan, L. 2012. Twitter To Surpass 500 Million Registered Users on Wednesday. *Media Bistro*. Online. Available: http://www.mediabistro.com/alltwitter/twitter-active-total-users_b17655. [08/6/2012].
- Du Plessis, S. 2012. *Collapse. The Story of the international financial crisis, its causes and policy consequences*. In: van Beek, U. and Wnuk-Lipinski, E., (eds.) 2012. *Democracy under stress: The Global crisis and beyond*. Berlin: Barbara Budrich Publishers.
- Eckstein, H. 1966. *Division and cohesion in democracy: A study of Norway*. Princeton: Princeton University Press.
- Economist, 2010. Data, data everywhere. *The Economist*. Online. Available: <http://www.economist.com/node/15557443>. [12/02/2012].
- Elections.Twitter, 2012. *The Twitter Political Index*. Online. Available: <https://election.twitter.com/>. [18/09/2012].
- Fails, M. D. and Pierce. H.N. 2010. Changing Mass Attitudes and Democratic Deepening. *Political Research Quarterly*. 63(1): 174-187.
- Finer, S.E. 1997. *The History of Government: From the Earliest Times. Vol. 1*. Oxford: Oxford University Press.
- Finer, S.E. 1999. *The History of Government. : From the Earliest Times. Vol. 1*. Oxford: Oxford University Press.
- Fukuyama, F. 1989, The end of history? *The National Interest*. 16(Summer): 4-18.
- Fukuyama, F. 1992. *The End of History and the Last Man*. New York: Free Press.
- Garnham, N. 1992. *The Media and the Public Sphere*. In Calhoun, C. (ed.). *Habermas and the Public Sphere*. Cambridge, MA: MIT Press.
- Gayo-Avello, D. 2012. "I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper": *A Balanced Survey on Election Prediction Using Twitter Data*. Computers and Society. Cornell University Computer Science Paper.
- Gieber, W. 1955 .Do Newspapers Overplay 'Negative' News? *Journalism & Mass Communication Quarterly*. 32(September): 311-318.

- Grantz and Reinsel, 2011. Extracting Value from Chaos. *IDC Iview paper*. Online. Available: <http://idcdocserv.com/1142>. [07/30/2012].
- Gruicic, M.D. 2011. *Internet and the Structural Transformation of Public Debate – A Comparison of the Online and Offline Public Spheres in Croatia*. Master's Thesis. Central European University.
- Habermas, J. 1976. *Legitimation Crisis*. London: Heinemann.
- Habermas, J. 1989. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Cambridge: Polity.
- Habermas, J. 1992. *Between Fact and Norms*. Cambridge, MA: MIT Press.
- Habermas, J. 1996. *Between Fact and Norms*. Cambridge, MA: MIT Press.
- Hadenius, A. and Teorell, J. 2006, Cultural and economic prerequisites of democracy: Reassessing recent evidence. *Studies in Comparative International Development*. 39(4):87-106.
- Hansen, M.H. 1992. The Tradition of the Athenian Democracy A.D. 1750–1990. *Greece & Rome*. 39(Second Series): 14-30.
- Harrison, R. 1993. *Democracy*. London: Routledge.
- Hauth, A. 2010. Twitter as a Public Sphere. *New media and Democracy*. Online. Available: <http://www.personal.psu.edu/alh5147/blogs/cas497a/2010/02/twitter-as-a-public-sphere.html>. [20/08/2012].
- Held, D. 1987. *Models of Democracy* (1st edition). Malden, USA: Polity Press.
- Held, D. 1996. *Models of Democracy* (2nd edition). Malden, USA: Polity Press.
- Held, D. 2006. *Models of Democracy* (3rd edition). Malden, USA: Polity Press.
- Higley, J. and Gunther, R. eds. 1992. *Elites and Democratic Consolidation in Latin America and Southern Europe*. New York: Cambridge University Press.
- Hobbes, T. 2008. *Leviathan*. (First published 1651). USA: Forgotten Books.
- Hopkins, D. and King, G. 2009. Replication Data for: A Method of Automated Nonparametric Content Analysis for Social Science. Online: UNF:3:xIE5stLgKvpeMvxzILxzEQ==hdl:1902.1/12898. [22/02/2012].
- Hopkins, D.J. and King, G. 2010. A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science*. 54(1): 229-247.

- Horowitz, I.L. 1991. Anti-Modernization, National Character and Social Structure. *Journal of Contemporary History*. 26(3/4): 355-367.
- Huntington, S.P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. USA: University of Oklahoma Press.
- Huntington, S.P. 1993. The Clash of Civilizations? *Foreign Affairs*. 72(3): 22-49.
- Inglehart, R. 2003. How Solid Is Mass Support for Democracy: And How Can We Measure It. *Asian Barometer*. Online. Available: http://www.asianbarometer.org/newenglish/publications/ConferencePapers/2003conference/T_03_no.11.pdf . [30/05/2012].
- Inglehart, R. and Welzel, C. 2005. *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.
- Krippendorff, D. K. 2004. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage.
- Knutsen, C. H. 2010. Measuring Effective Democracy. *International Political Science Review*. 31(2): 109-128.
- Kruger, J. and Dunning, D. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*. 77(6): 1121-1134.
- Kumar, A.K.M. and Suresha, A. 2010. An Empirical Approach for Opinion Detection Using Significant Sentences. *International Journal of Advanced Engineering & Application*. 6335(2010): 486-497.
- LatinoBarometer, 2012. *Online Results Analysis*. Online. Available: <http://www.latinobarometro.org/latino/LATANalyzeQuestion.jsp>. [16/08/2012].
- Lerner, D. 1958. *The passing of traditional society: Modernizing the Middle East*. New York, US: Free Press.
- Lijphart, A. 1977. *Democracy in plural societies: A comparative exploration*. New Haven: Yale University Press.
- Lindsay, R. 2008. *Predicting polls with Lexicon*. Online. Available: <http://languagewrong.tumblr.com/post/55722687/predicting-polls-with-lexicon>. [25/07/2012].

- Linz, J. 1990. Transitions to Democracy. *Washington Quarterly*. 13: 143-62.
- Linz, J.J. and Stepan, A. 1996. *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe*. Baltimore: Johns Hopkins University Press.
- Lijphart, A. 1985. *Democracies*. New Haven: Yale University Press.
- Lijphart, A. 1990. *Democracies: Patterns of Majoritarian and Consensus Government in 22 Countries*. New Haven: Yale University Press.
- Lipset, S.M. 1959. Some Social Requisites of Democracy: Economic Development and Political Legitimacy. *The American Political Science Review*. 53(1): 69-105.
- Lipset, S.M. 1994. The Social Requisites of Democracy Revisited: 1993 Presidential Address. *American Sociological Review*. 59(1): 1-22.
- Locke, J. 2005. *Two Treatises of Government and a Letter Concerning Toleration*. Sitwell: Digireads.com Publishing.
- Lohr, S. 2012. The Age of Big Data. *The New York Times*. Online. Available: http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?_r=4&hpw=&pagewanted=all. [20/03/2012].
- Lyman, P, and Varian, H.R. 2003. How Much Information 2003. *Technical report, University of California*. Online. Available: <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>. [15/03/2012].
- Machiavelli, N. 1975. *The Prince*. Harmondsworth: Penguin.
- Machiavelli, N. 1983. *The Discourses*. Harmondsworth: Penguin.
- Madison, J. 1966. *The Federalist Papers*. New York: Doubleday.
- Marsilius of Padua. 1980. *Defensor Pacis (The Defender of Peace)*. trans. and ed. Gewirth. A. Toronto: Toronto University Press.
- Meshak. 2012. *Map of a Twitter status object*. Online. Available: <http://mehack.com/map-of-a-twitter-status-object>. [15/03/2012].
- MIH Media Lab. 2012. Next Generation Internet. *MIH Media Lab Website*. Online. Available: <http://ml.sun.ac.za/next-generation-internet/>. [10/09/2012].
- Mill, J. 2008. *Representative Government*. (First published in 1861) USA: Forgotten Books.

- Miller, D., Coleman, J., Connolly, W. and Ryan, A. (eds). 1991. *The Blackwell Encyclopaedia of Political Thought*. Oxford: Blackwell Publishers Ltd.
- Montesquieu. 1952. *The Spirit of Laws*. Chicago: William Benton.
- Nagal, M. 2010. A Mathematical Model of Democratic Elections. *Current Research Journal of Social Sciences*. 2(4): 255-261.
- Neal, C. 2012. The Public Sphere and the New Media. *Social Media Today*. Online. Available: http://socialmediatoday.com/node/495708&utm_source=feedburner_twitter&utm_medium=twitter&utm_campaign=autotweets. [20/08/2012].
- Neuendorf, K. A. 2002. *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage.
- Nie, N.H. 2011. Revolution Analytics Executive Briefing: The Rise of Big Data Spurs a Revolution in Big Analytics. *Revolution Analytics*. Online. Available: <http://www.revolutionanalytics.com/why-revolution-r/whitepapers/The-Rise-of-Big-Data-Executive-Brief.pdf>. [15/03/2012].
- O'Connor, B., Balasubramanyan, R., Routledge, B., and Smith, N. 2010. *From tweets to polls: Linking text sentiment to public opinion time series*. In Proceedings of ICWSM: Carnegie Mellon University.
- O'Donnell, G. 1993. On the State, Democratization and Some Conceptual Problems. *World Development*. 21(8):1355-1369.
- O'Donnell, G. and Schmitter, P.C. 1986. *Tentative Conclusions about Uncertain Democracies*. In: O'Donnell, G., Schmitter, P.C. and Whitehead, L. *Transitions from Authoritarian Rule*, vol. 4. Baltimore: Johns Hopkins University Press.
- Oxford Dictionaries, 2012. *Democracy*. Online. Available: <http://oxforddictionaries.com/definition/democracy?q=democracy>. [24/04/2012].
- Paine, T. 2008. *The Rights of Man*. US: Read How You Want.
- Palmer, M. 1992. *Love of Glory and the Common Good: Aspects of the Political Thought of Thucydides*. USA: Rowman & Littlefield Publishers.
- Pang, B. and Lee, L. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*. 2(1-2): 1-135.
- Pang, B., Lee, B. and Vaithyanathan, S. 2002. *Thumbs up?* Sentiment Classification using Machine Learning Techniques. Proceedings of EMNLP: 79-86.

- Papacharissi, Z. A. 2002. The virtual sphere: The internet as a public sphere. *New Media & Society*. 4(1): 9-27.
- Papacharissi, Z. A. 2009. *The virtual sphere 2.0: The internet, the public sphere and beyond*. In: Chadwick, A. and Howard, P. (eds). *Handbook of Internet Politics*. Oxon: Routledge.
- Papacharissi, Z. A. 2010. *A Private Sphere: Democracy in a Digital Age*. Cambridge: Polity Press.
- Plato. 1974. *The Republic*. Harmondsworth: Penguin.
- Pocock, J. G. A. 1975. *The Machiavellian Moment: Florentine Political Thought and the Atlantic Republican Tradition*. Princeton: Princeton University Press.
- Poggi, G. 1978. *The Development of the Modern State*. London: Hutchinson.
- Polity IV Project, 2012. Polity IV Project: Political Regime Characteristics and Transitions, 1800-2010. *The Center for Systemic Peace*. Online. Available: <http://www.systemicpeace.org/polity/polity4.htm>. [11/07/2012].
- Przeworski, A. and Limongi, F. 1997. Modernization: Theories and Facts. *World Politics*. 49(January): 155-83.
- Pye, L.W. 1990. Political Science and the Crisis of Authoritarianism. *The American Political Science Review*. 84(1): 3-19.
- Raaflaub, K.A. 2008. *Introduction*. In: Raaflaub, K.A., Ober, J. and Wallace, J.W. (eds.) *Origins of Democracy in Ancient Greece*. California: University of California Press.
- Rao, L. 2012. DataSift Unlocks Access To Historical Twitter Data Dating Back To January 2010. *Tech Crunch*. Online. Available: <http://techcrunch.com/2012/02/27/datasift-unlocks-access-to-historical-twitter-data-dating-back-to-january-2010/>. [02/08/2012].
- Real Clear Politics. 2012. *General Election: Romney vs. Obama*. Online. Available: http://www.realclearpolitics.com/epolls/2012/president/us/general_election_romney_vs_obama-1171.html. [18/09/2012].
- RHRealityCheck, 2012. About Us. *RH Reality Check*. Online. Available; <http://www.rhrealitycheck.org/about-us>. [09/13/2012].
- Rose, R., Mishler, W. and Haerpfer, C. 1998. *Democracy and its Alternatives*. Baltimore: The Johns Hopkins University Press.
- Rose, R. 2001. A Divergent Europe. *Journal of Democracy*. 12(1): 93-106.

- Rousseau, J.J. 2008. *The Social Contract*. (First Published in 1762). New York: Cosmino Inc.
- Ritterman J., Osborne, M. and Klein, E. 2009. Using Prediction Markets and Twitter to Predict a Swine Flu Pandemic. *Publication by University of Edinburgh*. Online. Available: <http://homepages.inf.ed.ac.uk/miles/papers/swine09.pdf>. [09/08/2012].
- Rustow, . 1970. Transitions to democracy: Towards a Dynamic Model. *Comparative Politics*. 2(3): 337-363.
- Sassen, S. 2006. *Territory, authority, rights: From medieval to global assemblages*. Princeton: Princeton University Press.
- SemioCast, 2012. Twitter reaches half a billion accounts. *SemioCast*. Online. Available: http://semioCast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US. [03/09/ 2012].
- Schumpeter, J.A. 1976. *Capitalism, Socialism and Democracy*. New York: Harper Perennial.
- Shin, D.C. 2006. *Democratization: Perspectives From Global Citizenries*. CSD Working Papers. Centre for the Study of Democracy, UC Irvine.
- Sigmund, P.E. 1962. The Influence of Marsilius of Padua on XVth-Century Conciliarism. *Journal of the History of Ideas*. 23(3): 392-402.
- Sisk, T. D. 1994. Perspectives of South Africa's transition: Implications for democratic consolidation. *Politikon: South African Journal of Political Studies*. 21(1):66-75.
- Skelton, A. 2012. Social Demographics: Who's Using Today's Biggest Networks (Infographic). *Mashable*. Online. Available: <http://mashable.com/2012/03/09/social-media-demographics/>. [09/05/2012].
- Skinner, Q. 1992. *The Italian City-republics*. In: Dunn. J. (ed.), *Democracy: The Unfinished Journey, 508 BC to AD 1993*. Oxford: Oxford University Press.
- Sluis, E. 2009, Twitter: Public Space or Public Sphere? *Masters of Media*. Online. Available: <http://mastersofmedia.hum.uva.nl/2009/10/06/twitter-public-space-or-public-sphere/>. (20/08/2012).
- Starr, C.G. Jr. 1952. The Perfect Democracy of the Roman Empire. *The American Historical Review*. 58(1): 1-16.
- Tiun, S., Abdullah, R. and Kong, T.E. 2001. Automatic Topic Identification Using Ontology Hierarchy. *Computational Linguistics and Intelligent Text Processing*. 2004(2001): 444-453.

- Tocqueville, A. D., Bradley, P., Reeve, H., and Bowen, F. 1945. *Democracy in America*. New York: A. A. Knopf.
- Trenz, H. 2009. *Digital Media and the Return of the Representative Public Sphere*. Working Paper. Centre for European Studies.
- Tumasjan, A., Sprenger, T.O., Sandner, G. and Welpe, I.M. 2010. *Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment*. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media. Technische Universität München.
- Thompson, J. B. 2000. *Political scandal: Power and visibility in the media age*. Cambridge: Polity.
- Twitter Developers, 2012. *The Streaming APIs*. Online. Available: <https://dev.twitter.com/docs/streaming-apis>. [31/07/2012].
- Urban Dictionary, 2012. *Democracy*. Online. Available: <http://www.urbandictionary.com/define.php?term=democracy>. [15/02/2012].
- Van Beek, U., 2012. *The Crisis that Shook the World*. In: van Beek, U. and Wnuk-Lipinski, E., (eds.) 2012. *Democracy under stress: The Global crisis and beyond*. Berlin: Barbara Budrich Publishers.
- VentureBeat. 2012. Jack Dorsey: Twitter seeing 3 to 5 percent engagement on Promoted Tweets and Trends. *Venture Beat News*. Online. Available: <http://venturebeat.com/2012/01/22/jack-dorsey-promoted-products/#vGidIvFfak4DDYXQ.99> . [30/01/2012].
- Verba, S. 1996. The Citizen as Respondent: Sample Surveys and American Democracy Presidential Address, American Political Science Association, 1995. *American Political Science Review*. 90(1): 1-7.
- Voice of the People, 2012. *An annual worldwide Survey - The Voice of the People*. Online. Available: <http://www.voice-of-the-people.net/>. [16/08/2012].
- VOTP, 2005. Voice of the People 2005, Trends in Democracy, Global Summary. *Gallup International*. Online. Available: http://www.voice-of-the-people.net/ContentFiles/files/VoP2005/VOP2005_Democracy%20FINAL.pdf. [16/08/2012].
- WEF, 2012. Big Data, Big Impact: New Possibilities for International Development. *World Economic Forum: Vital Wave Consulting*. Online. Available: http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf. [12/03/2012].

- Welzel, C. 2002. Effective Democracy, Mass Culture, and the Quality of Elites: The Human Development Perspective. *International Journal of Comparative Sociology*. 43(3-5):269-298.
- Inglehart, R. and Welzel, C. 2005. *Modernization, Cultural Change, and Democracy*. New York: Cambridge University Press.
- Welzel, C. and Inglehart, R. 2008. The Role of Ordinary People in Democratization. *Journal of Democracy*. 19(1): 126-140.
- Wnuk-Lipinski, E., 2012. *The consequences of the Great Recession: hypothesis and scenarios*. In: van Beek, U. and Wnuk-Lipinski, E., (eds.) 2012. *Democracy under stress: The Global crisis and beyond*. Berlin: Barbara Budrich Publishers.
- Wollstonecraft, M. 1792. *Vindication of the rights of Woman*. Harmondsworth: Penguin.
- WVS, 2012. *World Values Survey: The World's most Comprehensive Investigation of Political And Sociocultural Change*. Online. Available: <http://www.worldvaluessurvey.org/>. [16/08/2012].
- WVS, 2012. *Online Data Analysis (Step-4): Question data*. Online: <http://www.wvsevsdb.com/wvs/WVSanalyzeQuestion.jsp>. [18.06.2012].
- Yatim, H. 2012. Our democracy less false than US. *Malaysiakini*. Online. Available: http://www.malaysiakini.com/news/202147?utm_source=twitterfeed&utm_medium=twitter. [09/17/2012].
- Yi, J., Nasukawa, T., Bunescu, R. and Niblack, W. 2003. *Sentiment Analyzer: Extracting Sentiments about a Given Topic using Natural Language Processing Techniques*. IEEE International Conference on Data Mining (ICDM): 427-434.
- Zeleny, J., 2012. Poll: Obama Holds Narrow Edge Over Romney. *The New York Times*. Online. Available: <http://thecaucus.blogs.nytimes.com/2012/09/14/poll-obama-holds-narrow-edge-over-romney/>. [18/09/2012].
- Zhao, F., Hu, Q, Xu, X., Zeng, R. and Lin, Y. 2011. Orientation Mining-Driven Approach to Analyze Web Public Sentiment. *Journal of Software*. 6(8): 1417-1428.

APPENDIX A: CALCULATIONS

Estimation of actual amount of Tweets:

$$n \times \frac{x \text{ or } z}{100} = y$$

n = number of tweets as classified by Chatterbox

x = lowest range of accuracy of Chatterbox classification as determined by the researcher

z = highest range of accuracy of Chatterbox classification as determined by the researcher

y = new estimation of number of tweets

Calculations on estimation of actual sentiment

Estimation of actual **lowest** range amount of negative Tweets: $38068 \times \frac{59}{100} = 22460.12 = A$

Estimation of actual **highest** range amount of negative Tweets: $38068 \times \frac{68}{100} = 25886.24$

= B

Estimation of actual **lowest** amount of positive Tweets: $65970 \times \frac{31}{100} = 20450.70$

= C

Estimation of actual **highest** amount of positive Tweets: $65970 \times \frac{39}{100} = 25728.30$

= D

Estimation of **lowest** wrongly classified negative Tweets: $38068 \times \frac{3}{100} = 1142.04$

= E

Estimation of **highest** wrongly classified negative Tweets: $38068 \times \frac{11}{100} = 4187.48$

= F

Estimation of **lowest** wrongly classified positive Tweets: $65970 \times \frac{26}{100} = 17152.20$

= G

Estimation of **highest** wrongly classified positive Tweets: $65970 \times \frac{34}{100} = 22429.80$

=H

Estimation of Actual **lowest** negative sentiment = $A + G = 22460.12 + 17152.20 = 39612.32$

Estimation of Actual **highest** negative sentiment = $B + H = 25886.24 + 22429.80 = 48316.04$

Estimation of Actual **lowest** positive sentiment = $C + E = 20450.70 + 1142.04 = 21592.74$

Estimation of Actual **highest** positive sentiment = $D + F = 25728.30 + 4187.48 = 29915.78$

APPENDIX B: EXAMPLE OF MANUAL CLASSIFICATION

Text - Random stratified Negative Tweets	My Classification	Chatterbox Classification
Amazing how today, many northerners chanting for freedom and democracy in Sanaa will not even see a problem in that (previous tweet)	-1	-0.058795697
RT @ksmith4138: democracy won; you lost. dont' cry about it Video 'End of Democracy' http://t.co/Z1ovSgXX	-1	-0.221566905
@Labaktee2012nativity,dis shows narrow minded talk pitiful:(Economy @ its lowest social freedom zero &there u say if not 4 PDP no Democracy.	0	-0.695430036
I'm totally anti-JI bit this part is true! RT@NomansMc: #JI is the only party with true intra-party democracy.â€•	0	-0.104092808
I IMPLORE ALL UNIONS TO BAND TOGETHER WITH THE MAJORITY OF THE PEOPLE OF OUR COUNTRY TO TAKE OUR DEMOCRACY BACK! THIS IS NOT A DICTATORSHIP!	1	-0.281952164
RT @MubasherLucman: @IzharulHassan local bodies are true democracy not this sham assembly	-1	-0.119963078
@Smooth_Orator @Rik_Flair hold the fuck up this is not a democracy muthafuckas!!!	-1	-0.428858185
Uncovering the Other ALECs http://t.co/Sm0c0jTh Beware of ALEC-like groups hiding in the bushes! Protect our democracy!	0	-0.11633628
@SteveInCM @bangkokpundit @Bkk_R I suppose it depends on what the priorities are: democracy, justice, getting Thaksin's money back, or what?	0	-0.268446366
@boogiechild5820 People don't want that chance. If they did it would b brought up n a meeting. That is a organization with democracy.	0	-0.308697926
RT @AstroSHayden: Mitt Romney and the whole GOP party is dangerous for America.When Koch Brothers can destroy democracy by buying GOP, DANGEROUS	-1	-0.122551748
Voted @ 4pm in local congressional race. Was 72nd person in my district to vote. Democracy at work #depressing	-1	-0.20049988
Bahraini forces attack pro-democracy protesters http://t.co/pCz9MoXS #Bahrain	-1	-0.347379654
RT @Dputamadre: 21st Century Democracy Empirism: _Destroyed? _ No! DEMOCRATIZED! #Libya #Perception #Empirism #Philosophy http://t.co/X31XgeNS	-1	-0.216932958
RT @haripdbhusal: @waglenikhil @ashutoshibn7 The team Anna is a bundle of selfish people who do not have courage to face Ballots. They are against Democracy.	-1	-0.193012917
RT @yaamyn: Angry Bird on DhiTV says that the pro-democracy protests doesn't have "substance". By that, I presume he means anti-Semitic rants?	0	-0.374997683
RT @sunshinejc: Why do Americans Allow corporations to co-opt our democracy by granting them the rights of legal personhood and defining money as speech #p2	-1	-0.482398066
RT @DirkZ1: The @GOP can't get votes honestly. They need to rig the system w/ Citizens United, Voter ID laws, busting unions and lies. #democracy #p2b	-1	-0.254472172

RT @rachelrosh: I didn't think Olympics was left wing. Proud of suffragettes and NHS and our struggles for democracy. I'm a Tory and I thought it was cool.	0	-0.065018862
With democracy, we witnessed rapid rise of crimes, sagging of economy, and to make it worse, politicians started boycotting tourism!	-1	-0.10114485
@ArmchairArab plus, you can't copy paste democracy. It needs to be custom developed.	0	-0.100748842
someone takes a baby into the area and @neilsinclair_ asks "so....what do you think of democracy?". absolute idiot.	-1	-0.06307624
RT @guptar: Are the mass graves of #Kashmir less heinous because they are the handiwork of a democracy? http://t.co/yyYYqGEn #India	-1	-0.050924795
Buying SR22 Auto Insurance Online The Web Democracy: Are you looking cheap SR22 insurance estimator due to sus... http://t.co/En7gAENC	0	-0.151842642
MEDIA MYOPIA: Oblivious to the destruction of democracy under their noses.#ConHome,#politicshomeuk,#bbcbreakfast	-1	-0.161543613
Our own democracy took centuries to get to where it is now, why expect Egypt to do it in a year? http://t.co/CCztnfAn #Kindle	0	-0.111738822
RT @ForeignAffairs Bad judges! #Pakistan's supreme court is undermining democracy: http://t.co/TcDNYaog	-1	-0.314920199
We cannot forget the lives lost in the fight for democracy and freedom http://t.co/FNGhQZPx #mchuzimix #Kenya #Africa	0	-0.050925027
RT @MIDKNIGHT6: Judge Sides with Monsanto, Ridicules Farmers' Rights http://t.co/A5Bdi06b Pls RT @food_democracy #stopMonsanto #GMOs	-1	-0.086993687
RT @Moawen: What @cherifbassiouni didn't mention in his report, is that #Bahrain govt stopped promoting pro-democracy employees http://t.co/yw9gKKI4	-1	-0.026190131
RT @rhrealitycheck: RT @MSignorile: Breaking: Conservatives plan 2 leave U.S., but can't find wealthy Western democracy w/ouniversal health care. #hrc #scotus	-1	-0.290353241
RT @TolpuddleTim: Today's market news: rich people worried that poor people might take some of their wealth off them as democracy intervenes in profit-making	-1	-0.265868456
@Bill_Cimbrello @BuddyRoemer Hopefully it's a row boat since #Obama will be shutting off ur oil supply after he destroys bridges to Democracy	-1	-0.170983588
Everytime someone says #Kenya is a grown democracy I cringe	-1	-0.614769376
RT @sabokitty: #LibertyPlaza was not a protest. Protests emitted from it every day, but it was an escalation. A radical experiment in direct democracy #ows	-1	-0.346329347
@GetUpStandUp2 pub Ed testing will build a democracy of engineers who are good choosing the best of 4 choices. sad for #science #SOSchat	-1	-0.119605705
Democracy... just a dream?	-1	-0.165319761
@Kirkabout @jonsnowC4 People don't want less democracy, but do want less policians. RT: My big brother, talking sense!!!	-1	-0.208445717
Give Me Liberty: A Handbook for American Revolutionaries (Paperback): As the practice of democracy becomes a los... http://t.co/uMDhCk9D	-1	-0.186360597
Democracy, what democracy? #Australia #Uranium #USbases #Darwin #CSG #Mining #CarbonTax #NoThanks	-1	-0.049826982
http://t.co/Mf5xu9Jh Democracy, Politics and Terrorism - America's Quest for Security in the Age of Insecurity	0	-0.155432708
@ANIALLATOR oh my god! I know, same here! I feel yah!	0	-0.497058821
RT @BenHowe: Liberals think democracy is failing if they can't take money out of worker paychecks against their will and use it to finance campaigns.	-1	-0.248251
Hillary touts democracy with Morsi (@times247) http://t.co/Eu26tBb9 via @sharethis	0	-0.022361796

FM Khar: u don't want Pakistani ppl to think that dictatorship earns Pakistan more US support than democracy #usislam12	-1	-0.514371736
RT @jaskline2007: @thinkprogress Even though we aren't a true democracy but rather a republic.	0	-0.306314459
RT @DianeRavitch: When we lose neighborhood schools, we lose community. We lose a building block of democracy where people learn to take action. #soschat	-1	-0.391050614
The Moment of #Youth. Our Kids are Failing #Democracy. @DrPriceMitchell describes how to can engage youth in #service. http://t.co/xCcQ4RT0	-1	-0.420739909
Washington to Offer Voter Registration Via Facebook: Ah Democracy! Isn't it grand that we live in a nation wher... http://t.co/J4dMniLK	1	-0.189724016
America is a republic, not a full-blown democracy. However, the democratic tradition in America has always been a... http://t.co/eC6RzqL8	0	-0.086065452
RT @Drudge_Report: Sobbing man: 'Democracy died tonight'... http://t.co/jdFY5XKj	-1	-0.486390975
Democracy at work is doing things to appease your voters. Lrn2politics GOP. 4 more years.	0	-0.308802555
Text Random Stratified Positive tweets	My Classification	Chatterbox Classification
The Digital Origins of Dictatorship and Democracy: Information Technology and Political Islam (Oxford Studies in... http://t.co/2HQmDxYI	0	0.202694831
Reopening of NATO supply line is not surprising, what else do you expect from present regime... Pure Jokes in the name of democracy	-1	0.193396444
Professors' Manifesto For the protection of democracy and the right of student protest http://t.co/b6SmHv7Y #ggi #maplespread #nycggi #OWS	0	0.423186798
Deals with the devil - Open Democracy http://t.co/FVU7IUJW	-1	0.31434708
RT @G_AnnWilkerson Quote of the Day: "Character, not circumstances, makes the man." - Booker T. Washington, Democracy and Education	0	0.389065284
RT @piratepartyau: Currently speaking, @AndrewDowning4 on our liquid democracy implementation, Polly. #piratecon	0	0.336595252
"I like the idea of democracy. You have to have someone everyone distrusts," said Brutha. "That way, everyone's happy." (Small Gods)	1	0.61188108
RT @JavedGhamidi: Some people believe that democracy is a concept alien to Islam. My response is here http://t.co/iR3aZZqU	1	0.277034926
Deepening Democracy: Global Governance and Political Reform in Latin America: Adams surveys the impact of transn... http://t.co/deuEKov6	1	0.362670303
RT @RobertLuchford: You got a nation, you got democracy. Now go and moan to syrian's or Zimbabwe's. Stop moaning bloody & make something of yourself!	1	0.117328898
A former member of MFDC rebels allowed to run in #Senegal's legislative elections. A good sign for #democracy? http://t.co/mRiFO0pe	1	0.384836002
@Reidy259y ah so freedom of speech only applies to those who believe in the union, you are familiar with the defintion of democracy	1	0.369563706
RT @simphiwedana: The revolution has fallen. This here democracy is a white institutionalized farce by sellout black leaders. #StateOfEmergency	-1	0.015521529
.@TheRealRoseanne @tomrn I 100% appreciate your advocacy, Madame President! Competition is what democracy is all about. Makes change happen!	1	0.536653838

Just played: Democracy Now Promo - Carla Myers, WNJR	0	0.188892761
RT @katjanouch: "I get suspicious when I hear word democracy, under such beautiful names lot of crimes are committed". ~~Nawal Al Sadaawi	-1	0.166975948
RT @joshtentoni: We can now say that the United States of America is no longer a democracy	-1	0.099073309
#Lesotho election gives southern #Africa a lesson in democracy. http://t.co/LDnBL9T7 @GlobalPost	1	0.437264446
RT @AnonymousIRC: Ohai Hong Kong. Tens of thousands demanding democracy on the street: http://t.co/WKdNiHvP Asia is waking up! (via @57UN) #Anonymous	1	0.19532799
RT @HealHealthcare: #ReasonsToVoteForObama & Vote GOP out: Repubs are selling out Democracy to the wealthiest people in the world #p24U	-1	0.32980223
RT @maulingmueller: When one motivated individual has the power to bring the economy to a grinding halt a country stops being a democracy.	-1	0.160447449
Official projections show New Democracy in lead in Greece: June 17 - Official projections show the conservative ... http://t.co/fI38rVhx	0	0.392283059
RT @MexicoVoices: Voices of #YoSoy132: The creator of the video of 131 students at the Ibero #studentprotests #democracy http://t.co/urIeZ0MU	0	0.387079405
RT @DalaiLama: Human values like democracy, freedom of speech and freedom of the individual, these are changes whose momentum cannot be stopped.	1	0.21312112
RT @BBCBreaking: New Democracy projected to win 127 seats, Syriza 72, Socialists 32 - second #Greece2012 exit poll http://t.co/sMjhmwIG	0	0.342351132
"A New Democracy - Pasok government also seems to be the preferred choice of Brussels..." http://t.co/8wIjfbmN 'New Democracy?' Orwellian.	1	0.334239819
Libya : Exclusive: Libya's Civil Crackdown Worries Democracy Advocates - Daily Beast: Daily BeastExclusive: Liby... http://t.co/bSWBsS9s	-1	0.31024294
@HollandTaylor if democracy requires votes at the ballot physically. Money doesn't vote people do. Voter apathy is to blame. IMHO #voteUSA	-1	0.151446481
RT @CoCreateUMBC: With @ChaiKay Tea & @cmeghan1 flying over America, dreaming the future of democracy, making astonishing plans for @umbcsga. #inspired #adp12	1	0.379031544
Hey kid, wanna buy a democracy? @upworthy @BillMoyers http://t.co/m0Dtzg9v	-1	0.086166676
@DEMSnREPS2BLAME Where would the Middle East be today if the USA didnt overthrow a democracy in Iran & instill a dictator? #tcot #tlot #p2	-1	0.168649088
UK & World News: Burmese democracy leader Aung San Suu Kyi to address Houses of Parliament on visit to UK http://t.co/XIFJNNot	0	0.229609204
Hong Kong holds big protest as new leader sworn in: A pro-democracy heckler interrupted a s... http://t.co/XGVEOfEA http://t.co/o48iIEH	0	0.140720965
RT @lauraped1: i pledge allegiance 2 the banks of the united states of corporations, & 2 the plutocracy 4 which it stands, one nation under goldman sachs.	-1	0.312385063
RT @MargaretAtwood: If you're Canadian, and involved with a charity: this is serious! Support democracy or kiss it goodbye. http://t.co/30qK89WC	1	0.108672648
[eng] Med: Tunisian entrepreneur, unemployment helps Salafis: Democracy depends on development, Germany good example http://t.co/WULaAPTT	1	0.419039309

Tell the American Medical Association to Label GMOs! http://t.co/sXjCcdhC @food_democracy Please RT #LabelGMOs	0	0.303687366
kei gala: rangle-- minority and democracy!!!! usa!!!!!!!!	1	0.100854662
RT @Obama2012: POTUS: "In this election, on every single challenge that we face, you've got the final say. That's the amazing thing about our democracy."	1	0.317294059
Welcome Ali, we r. @akchishti @aliarqam @AmjadAliZardari @mughalbha thanks a lot, we are a family fighting for democracy and secular ideals	1	0.347500061
Democracy is the process by which people choose the man who'll get the blame. Bertrand Russell	-1	0.360538069
RT @SegunAke: RT @segudemuren: You can't dictate what people should protest. This is the other part of democracy.	1	0.093882294
THANK YOU VERY MUCH! For The Last Time, America Is NOT A Democracy http://t.co/SA4sWPbc via @WestJournalism	-1	0.081654843
Dr Dre - Detox "JasonTheFace: @arielhelwani @danawhite which was later, Guns n Roses - Chinese Democracy or that Rosenthal stoppage?"	0	0.011541245
RT @sushi_goat: QUICK, USE THE UNACCOUNTABLE DEMOCRACY TO CHECK ON THE OUT-OF-CONTROL PRIVATE POWER	-1	0.254422542
Democracy Now! National and Global News Headlines for Friday, June 1: http://t.co/v1i126Vo via @youtube	1	0.306743475
RT @thobbuis: Democracy in Question in the Maldives: President Mohamed Nasheed Speaks This event will be webcast live http://t.co/vtBxhkzH	-1	0.388638491
GOP: Hungry Like the Wolf. GOP- The end of democracy? http://t.co/u9suzCsh via @AddInfoOrg #p2 #GOP #ff #tcot	-1	0.474921032
Democracy with a difference http://t.co/myDGbnqn #Fiji	0	0.340551892
RT @mashafeeg: @DrWaheedH Your kind of democracy is Government of the MNDF & MPS, by the MNDF & MPS, for the MNDF & MPS. #DamnGoodDemocray eh?	-1	0.422702785

APPENDIX C: INTERESTING FOR FUTURE CONSIDERATION

The Dunning-Kruger effect

In the backdrop to this thesis on creating instruments for the measurement of expressions on democracy, lie deeper questions such as whether democracy actually offers an optimal political system of governance. This author agrees with the famous statement of Churchill, stating that “Democracy is the worst form of government, except for all those other forms that have been tried from time to time.” With this sentiment in mind many researchers have examined the functioning of democracy. In the next section an interesting view on the limited cognitive capabilities of humans and its effect on democracy can be found.

An obvious constant element in any democracy is the *demos* - the people. What role the people should play and what role they can be expected to play is a question that has been asked throughout the history of democracy, as has been shown in this thesis. An enlightening study, regarding the mental aptitude of people, was conducted by David Dunning and Justin Kruger and published in 1999. According to their article, people hold overly favourable views of their abilities – which impact their interaction in many social and intellectual domains. Another interesting conclusion of this study shows that those participants with higher metacognitive competence recognised their limitations and therefore scored their abilities lower than those with a lower metacognitive competence (Kruger and Dunning, 1999: 1121). This cognitive bias which people hold towards their own capabilities is popularly called the Dunning-Kruger effect.

Seeing that it is the task of all people to elect the best leader to represent their interests in a modern representative democracy, it is problematic if these people are unaware of their own limitations and biases in their decision making capabilities. This problem is further aggravated by the fact that less competent people would feel more certain of their decisions than those who are more competent and also more aware of their limited decision-making abilities. This idea however is not new as philosophers such as Plato, in *Apology*, already noted similarly that:

“This man, on one hand, believes that he knows something, while not knowing (anything). On the other hand, I – equally ignorant – do not believe (that I know anything).”

With regards to the cognitive biases that all people suffer from, a study by Mato Nagal (2010) has attempted to show the limitations this places on democratic elections. In this study a mathematical model of democratic elections is created, building on insights such as the Dunning-Kruger effect. If one looks at the functioning of a democracy from the vantage point of the whole compound - therefore the combination of the whole *demos* - then Nagal (2010:255) argues that democratic elections are far from optimal.

The Nagal (2010:259-260) model finds that democratic elections lead to the election of mediocre political leaders. The democratic process does however prevent the election of below-average leaders, which makes it better than an authoritarian system. He believes that the only way to substantially improve these results would be to increase the capability quotient of the whole of society. Nagal briefly examines the idea of using communication technologies in order to gain information about the voters and representatives; and their capabilities. He (2010:259) states that introducing a natural selection system, in which the less capable are subsequently eliminated from the representative list, might lead to the appointment of the best possible leader.

This argument serves to show that there might be deeper structural problems to the system of democracy. This author cannot point towards a better system of governance at this time, but it should be noted that with movements such as the Occupy Wall Street Movement, it seems that people are calling for a system of more direct democratic participation. Still, even if democracy changed to be more participatory and direct, the critique above on the limited cognitive capabilities of the populace will still hold.

APPENDIX D: GROUPING NEGATIVE SENTIMENT

Money Corruption

RT @AstroSHayden: Mitt Romney and the whole GOP party is dangerous for America. When Koch Brothers can destroy democracy by buying GOP, DANGEROUS
RT @sunshineejc: Why do Americans Allow corporations to co-opt our democracy by granting them the rights of legal personhood and defining money as speech #p2
RT @rhrealitycheck: RT @MSignorile: Breaking: Conservatives plan 2 leave U.S., but can't find wealthy Western democracy w/ouniversal health care. #hrc #scotus
RT @TolpuddleTim: Today's market news: rich people worried that poor people might take some of their wealth off them as democracy intervenes in profit-making
RT @BenHowe: Liberals think democracy is failing if they can't take money out of worker paychecks against their will and use it to finance campaigns.
RT @jillstein2012: "We don't need to run America like a business or like the military. We need to run America like a democracy." PLS RT
RT @jillstein2012: "We don't need to run America like a business or like the military. We need to run America like a democracy." PLS RT
Democracy is a form of government that substitutes election by the incompetent many for appointment by the corrupt few.
The democracy in #MÃ©xico is null, our institutions are sold
.@oakeshottm Diamond made bad decisions, but how does political pressuring an employee of a private company constitute democracy?
Think Democracy only benefits the Rich and oppresses the Poor. Shameful amount of Poor in a Country like US for example.
FRAMERS intended each voter to have EQUAL voice in representative democracy. So why did ACTIVIST JUDGES give rich guys Super PAC MEGAPHONES?

Functional

RT @ForeignAffairs Bad judges! #Pakistan's supreme court is undermining democracy: http://t.co/TcDNYaog
@Kirkabout @jonsnowC4 People don't want less democracy, but do want less politicians. RT: My big brother, talking sense!!!
RT @DianeRavitch: When we lose neighborhood schools, we lose community. We lose a building block of democracy where people learn to take action. #soschat
The Moment of #Youth. Our Kids are Failing #Democracy. @DrPriceMitchell describes how to can engage youth in #service. http://t.co/xCcQ4RTo
RT @DianeRavitch: When we lose neighborhood schools, we lose community. We lose a building block of democracy where people learn to take action. #soschat
RT @ggreenwald: I wonder what percentage of Americans know their government is constantly bombing Yemen & killing civilians? #Democracy
Tradition means giving votes to the most obscure of all classes, our ancestors. It is the democracy of the dead. Tradition refuses to submit
#wirecall #p21 #ows RT @politicalmath: Man... I wish my vote counted as "democracy". Apparently only D votes count.
#Obama? #Romney? #Duopoly! Any #debates need to include #JillStein and #GaryJohnson as well! Any real #democracy needs more than 2 choices.

Pure Negative

RT @Drudge_Report: Sobbing man: 'Democracy died tonight'... http://t.co/jdFY5XKj
RT @Rana_J01: " #Egypt went..from having no democracy...to having a fatiguing and confusing amount of it." http://t.co/wLWosEKi from the brilliant Kevin
RT @ProdigalSon_Ro: I hate how they've abused n distorted what freedom n democracy is...I hate how they refuse to assist the vast majority of US!
RT @BHNPREDAATOR: When a country like #Iran which killed more than 120K Iranians & holds th rest hostage talks abt democracy N #Bahrain , clean up ur shit 1st
My friend Kasweka is right. This eviction and destroying of people's Kambashus is oppression. Theres no sign of democracy here... eeee
RT @mrdaveyd: As we talk abt NBA playoffs, Rihanna or George Zimmerman broken nose, right now in Chicago our way of life & democracy is being broken
RT @UrbanIntifada: Iraq Veteran: "I'm returning my medals because under the guise of democracy I stole Iraq's humanity and lost mine." #NoNato @OccupyChicago
This is not a democracy, it's a cheerocracy! ^.^
Another forgotten massacre: Kashmir. Within a so-called democracy http://t.co/yCpFhYKf
In #Egypt, democracy is not back to square one. It is actually worse off now. What a genius devilish ploy by former regime's supporters
@GarthGodsman I'm an Aussie living in #Ecuador for 4 years now. #Democracy working better here then it is in #UK, #USA or #Australia.
RT @YourAnonNews: Error 404 Democracy Not Found - #ACTA vs #OpTequila
Where is democracy?
RT @bisonranchr68: #IAmARepublicanBecause I love Democracy so much that I'm working hard to allow FASCISM to overthrow it
People seem tired of politics ... What does ths mean for democracy?
RT @TimKarr: This is *not* what democracy looks like: #NYPD raid activists' home in advance of #OWS protests: http://t.co/7acDINyX
Whats so great about democracy? We have democracy and America is still fucked up. But i guess capitalism is to blame for that.