The Diffusion of Information, Emotions and Opinions on Social Media

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Informatik und Angewandte Kognitionswissenschaft
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte kumulative Dissertation

von

Björn Roß
aus
Hannover

Gutachter: Prof. Dr. Stefan Stieglitz
Gutachter: Prof. Dr. Torsten Zesch

Tag der mündlichen Prüfung: 21. Oktober 2019

DOI: 10.17185/duepublico/71608
URN: urn:nbn:de:hbz:464-20200514-113711-9

Acknowledgements

I wish to express my sincerest gratitude to Prof. Dr. Stefan Stieglitz. Without your steady support, I would have never been able to complete this thesis. I would also like to thank Prof. Dr. Torsten Zesch for the many discussions we had.

I wish to thank the Research Training Group User-Centred Social Media at the University of Duisburg-Essen. UCSM gave me an environment in which I was constantly challenged and exposed to new ideas from different fields. Not only did our collaboration lead to many successful publications, but I made new friends who stood by me at all times.

I would like to extend my thanks to the Research Group Professional Communication in Electronic Media/Social Media, University of Duisburg-Essen. Research is often a team effort, and ProCo was the best team I could have hoped for. I will never forget the many ways in which we have helped each other along the way – nor the fun we have had.

I would like to thank my family and friends, who supported me throughout this period of my life, and especially Emily, not primarily for the many times you proofread my writing, although that was also invaluable, but for the constant encouragement.

Content

Figures

Tables

Abbreviations

API	Application programming interface
ESA	Emergency service agency
NLP	Natural language processing
SMA	Social media analytics
URL	Uniform Resource Locator

1 Introduction

1.1 Research context and motivation

The spread of information on social media shapes everyday decisions. As customers, we decide what to buy, as patients, we decide which doctors to consult, and as voters, we decide which choices to make at the ballot box partly based on the information we see on social media. Yet, the facts and opinions that we are presented with are a result of a complex web of factors. Driven by psychology, we make choices – which social media to sign up for and check regularly, who to follow or become friends with – that affect what we see. Driven by business interests, platforms employ filtering algorithms that gauge which content we are likely to respond positively to. This results in a steady mix of new content. Consciously or subconsciously, we then have to take the decision whether to 'like' it, forward it to friends, or interact with it in some other way. Whether we actually take the advice we see to heart, and believe the facts we are presented, is yet another question.

Against this backdrop, it is essential to understand which content spreads on social media, and thus affects social media users' decisions. The spread of content such as news stories on social media is also known as information diffusion (cf. Gruhl, Liben-Nowell, Guha, & Tomkins, 2004; Kim, Bae, & Hastak, 2018). In particular, understanding the patterns and mechanisms of information diffusion allows us to explain, predict and perhaps even counter the negative consequences of the spread of misinformation. One way in which researchers approach the topic of information diffusion is by studying real-life social media interactions. This is known as social media analytics, and it is emerging as a distinct research area that is grounded in a variety of disciplines such as computer science, information systems and psychology (Stieglitz et al. 2014). It is characterised by a distinct methodological approach, which is different, for example, from the laboratory experiments and survey research methods employed by psychologists, and it comes with its own advantages and disadvantages, opportunities and challenges. Vast amounts of unstructured data are difficult to analyse. Attempts to structure the data, to quantify text data, for example, may prove unreliable.

Yet, understanding which type of content spreads faster or reaches more people is a prerequisite for organisations such as emergency service agencies to be able to use social media effectively. It is at the same time increasingly being recognised that this environment could also be manipulated by organisations with less charitable aims. There have been reports of disinformation campaigns by political groups and even governments. How users could be helped in distinguishing between fact and fiction is an open question.

Moreover, while the accounts that post fictitious stories or attempt to divert users' attention might be controlled by humans, they might also be controlled by algorithms known as bots, or social bots (Ferrara et al. 2016). Previous research has especially described them in political contexts and sought to find method of identifying them. Our knowledge of how they could influence humans, perhaps also in other contexts such as commerce where people also take decisions based on what they see on social media, is limited.

A convincing explanation also needs to be firmly grounded in theory. Information systems research is characterised by a strong focus on the use of theory to explain or predict phenomena (Gregor 2006). I argue in this thesis that only by grounding social media research in theory, and by connecting the social media analytics methods from computer science with the theoretical knowledge about human interactions from research fields such as communication, we can make statements that generalise beyond the individual case that was studied.

This research is made more challenging by the fact that social media analytics is a relatively new research area. Methodological innovation is commonplace, approaches are often adopted from other research fields (Ghani et al. 2018; Stieglitz et al. 2014; Thelwall 2018). Yet, standards are only beginning to develop. There are few comprehensive discussions of this research field. This thesis advances the methods used in social media analytics, in order to help this young research area develop and mature. It systematically discusses challenges in social media analytics and possible solutions.

Social media analytics is not only necessary for theorising about people's interactions. It has two very practical sides to it. One is social media monitoring: organisations continuously monitor the discussions that take place on social media. Companies are interested in what customers are saying about their product, for example to manage their reputation and to receive useful feedback for future product development (Fan and Gordon 2014) Researchers have used the emotions expressed on social media to analyse mood swings across cultures (Golder and Macy 2011) There have even been controversial attempts to predict the outcomes of elections and other votes based on the opinions expressed on social media (Kalampokis et al. 2013). There are many established methods in social media analytics, such as sentiment analysis and social network analysis. However, social media analytics is a process consisting of multiple steps, and data analysis is only one stage. Before, there are challenges to be faced in data discovery, data collection, and data processing, and I argue in this thesis that understanding them is a prerequisite for any social media analytics practitioner.

Secondly, apart from monitoring what is being discussed on social media, it is also used as a tool to communicate information to the public. Companies do this as part of marketing efforts, and government agencies and non-governmental organisations such as the emergency management services apprise the public of important developments (Reuter et al. 2016). For example, during a terrorist attack, the police and fire fighters need to quickly inform people that they should avoid the area or encourage them to call a specific number if they have more information about the incident. A practical problem that arises in this context is how to formulate social media posts to reach as many people as possible. There are countless blog posts addressing this topic[1], but they are based on anecdotal evidence, not on systematically collected empirical data that was analysed using a published method. However, to develop reliable guidelines, a better understanding of how information, opinions and emotions diffuse on social media is necessary first.

1.2 Research questions

The central goal of this thesis is to understand and explain mechanisms and patterns of information diffusion. To be able to understand and explain, it is first necessary, in social media analytics, to collect, prepare, and analyse data (Stieglitz et al. 2014). Each of these steps poses its own challenges, which need to be identified and addressed. Since social media analytics is still an emerging research field, the solutions to these challenges are still far from clear. This was especially true in 2016, when I began working on this thesis. It has also been recognised that social media analytics is especially complex because it often combines multiple methods in a single study, including, for example automatic processing alongside human judgements, and there are nearly endless alternatives for how a given data set could be analysed (Thelwall 2018). The first research question in this thesis is therefore:

RQ1: What are the methodological challenges in research on information diffusion on social media?

Addressing this research question in a useful way does not only involve understanding the challenges but also pointing to concrete solutions. Some challenges arise in many contexts in social media analytics and have been described in the previous literature but never systematically collected and catalogued, while other more specific ones that arose during my work on my thesis also help address this research question.

Once these challenges have been adequately examined and solved, social media analytics can be used to address relevant questions. As explained above, an especially relevant goal is to understand the diffusion of information, emotions, and opinions. Such knowledge

[1] For example, https://blog.hootsuite.com/writing-for-social-media/ last accessed: 2019-09-02

could, for example, be used to optimise strategies for disseminating information to social media users, and to contain the spread of misinformation. Therefore, the following question is addressed in this thesis:

RQ2: What are the factors that shape the spread of information, emotion and opinions on social media?

Addressing this question involves quantifying the impact of various factors on the diffusion. The desire to understand the factors that shape the spread of information, emotions, and opinions on social media naturally gives rise to the question of who the individuals and groups are who are attempting to spread their points of view. Social bots are especially widely discussed as a potential tool for manipulating opinion formation on social media. This phenomenon is therefore examined in this thesis.

The spread of information is not limited to one social media site. Considering more than one site at a time leads to the question of how information diffuses between different types of online media. Finally, the question arises of how users could be helped to distinguish false stories from true ones, which would negatively affect the spread of misinformation.

The two main scientific contributions of this thesis are (1) an improved understanding of the challenges that are involved in social media analytics, and (2) a proposed mechanism for the influence of malicious individuals on public opinion formation, not only in politics, but also in other domains such as commerce, supported by theoretical and empirical studies on social bots.

1.3 Thesis structure and list of publications

This is the synopsis of a PhD thesis in cumulative form, following section 9 (3) of the regulations governing doctoral proceedings at the Faculty of Engineering of the University of Duisburg-Essen. The thesis consists of this synopsis and a series of research papers published in international academic journals and conference proceedings.

Chapter 2 of this synopsis explains the background regarding the diffusion of information, emotions and opinions. Important terms are defined and relevant previous research is summarised. The third chapter summarises the research methods used for the studies described in this thesis. It explains social media analytics, which is the dominant approach in this thesis, and compares the individual studies regarding the methods used. Chapter 4 presents the results. The results from the individual papers are not presented by themselves but are rather related to each other. It is shown how the studies built on one another. The fifth chapter discusses the meaning of the findings in a wider context and relates them

back to the research questions. It also reflects on their implications, both for research and practice, and the limitations of this thesis, and presents possible directions for future research. Chapter 6 draws a short conclusion.

Table 1 gives an overview of the papers that form part of this thesis in logical order. Because of differences between journals and conferences in turnaround time, this does not necessarily correspond to their order of publication. In my time as a PhD student, I was able to publish 16 research papers in international and national journals as well as conference and workshop proceedings. Nine of these papers are included in this thesis. The table shows the ranking of the publication venue, according to the German Academic Association for Business Research (VHB), their Journal Citation Reports (JCR) Impact Factors, and citation counts from Google Scholar, as well as the title, authors, year, and publication venue. The papers were co-authored with researchers from the Ludwig Maximilian University of Munich (LMU), Germany, the University of Agder (UiA), Norway, and Queensland University of Technology (QUT), Brisbane, Australia, as well as my colleagues at the University of Duisburg-Essen, in particular the Research Training Group User-Centred Social Media.

Of the nine papers, three were published in journals, namely the International Journal of Information Management (IJIM), Information Systems Frontiers (ISF), and the European Journal of Information Systems (EJIS). The latter is one of the eight journals selected by the Association for Information Systems (AIS) for the Senior Scholars' Basket of Journals, informally known as the "basket of eight". Six papers in this thesis were published in conference or workshop proceedings. Three of the conference papers are full papers (ICIS, HICSS, and ACIS), while three are short or research-in-progress papers (ICIS, ECIS, and NLP4CMC). The ICIS, notably, is "the most prestigious gathering of IS academics and research-oriented practitioners in the world", and it "attracts the top research papers in the field" (AIS 2019).

Table 1. List of publications

		Publication	Type	VHB Ranking [1]	JCR Impact Factor [2]	Citations (Google Scholar) [3]
1	Title:	Social Media Analytics - Challenges in Topic Discovery, Data Collection, and Data Preparation	Journal	C	5.063	104
	Authors:	Stieglitz, Stefan; Mirbabaie, Milad; Ross, Björn; Neuberger, Christoph				
	Year:	2018				
	Venue:	International Journal of Information Management (IJIM)				
2	Title:	Going Back in Time to Predict the Future - The Complex Role of the Data Collection Period in Social Media Analytics	Journal	B	2.539	5
	Authors:	Stieglitz, Stefan; Meske, Christian; Ross, Björn; Mirbabaie, Milad				
	Year:	2018				
	Venue:	Information Systems Frontiers (ISF)				
3	Title:	Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis	Workshop	N/A	N/A	83
	Authors:	Ross, Björn; Rist, Michael; Carbonell, Guillermo; Cabrera, Ben; Kurowsky, Nils; Wojatzki, Michael				
	Year:	2016				
	Venue:	Third Workshop on Natural Language Processing for Computer-Mediated Communication (NLP4CMC)				
4	Title:	The Diffusion of Crisis-Related Communication on Social Media: An Empirical Analysis of Facebook Reactions	Conference (full paper)	C	N/A	9
	Authors:	Ross, Björn; Potthoff, Tobias; Majchrzak, Tim A.; Chakraborty, Narayan Ranjan; Lazreg, Mehdi Ben; Stieglitz, Stefan				
	Year:	2018				
	Venue:	Hawaii International Conference on System Sciences (HICSS)				
5	Title:	Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts	Conference (full paper)	N/A	N/A	29
	Authors:	Stieglitz, Stefan; Brachten, Florian; Jung, Anna-Katharina; Ross, Björn				
	Year:	2017				
	Venue:	Australasian Conference on Information Systems (ACIS)				

6	Title:	Social Bots in a Commercial Context - A Case Study on SoundCloud	Conference (short paper)	B	N/A	1
	Authors:	Ross, Björn; Brachten, Florian; Stieglitz, Stefan; Wikström, Patrik; Moon, Brenda; Münch, Felix Victor; Bruns, Axel				
	Year:	2018				
	Venue:	European Conference on Information Systems (ECIS)				
7	Title:	Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks	Journal	A	2.603	5
	Authors:	Ross, Björn; Pilz, Laura; Cabrera, Benjamin; Brachten, Florian; Neubaum, German; Stieglitz, Stefan				
	Year:	2019				
	Venue:	European Journal of Information Systems (EJIS)				
8	Title:	Information Diffusion between Twitter and Online Media	Conference (short paper)	A	N/A	0
	Authors:	Jung, Anna-Katharina; Mirbabaie, Milad; Ross, Björn; Stieglitz, Stefan; Neuberger, Christoph; Kapidzic, Sanja				
	Year:	2018				
	Venue:	International Conference on Information Systems (ICIS)				
9	Title:	Fake News on Social Media: The (In)Effectiveness of Warning Messages	Conference (full paper)	A	N/A	1
	Authors:	Ross, Björn; Jung, Anna-Katharina; Heisel, Jennifer; Stieglitz, Stefan				
	Year:	2018				
	Venue:	International Conference on Information Systems (ICIS)				

[1] Ranking list released by the German Academic Association for Business Research (VHB), see https://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/

[2] 2018 impact factors released in June 2019 by Clarivate Analytics, see https://jcr.clarivate.com/

[3] Citation counts as of 30 August 2019, see https://scholar.google.be/citations?user=RQ2zK8QAAAAJ

2 Background

2.1 Social media

Social media is commonly defined as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content" (Kaplan and Haenlein 2010, p. 61). In this definition, Web 2.0 refers to the notion of using the World Wide Web "as a platform whereby content and applications are no longer created and published by individuals, but instead are continuously modified by all users in a participatory and collaborative fashion" (Kaplan and Haenlein 2010, p. 61), and user-generated content comprises "various forms of media content that are publicly available and created by end-users" (Kaplan and Haenlein 2010, p. 61). While a personal website is static and controlled by a single individual, a Facebook news feed can be thought of as having been written by a diverse group of authors each of which contributed blocks of text, images, and/or videos. Notably, this definition also includes collaborative encyclopaedias such as Wikipedia. It excludes instant messengers, at least when messages are only exchanged between two individuals (whose content is not publicly available, not even to a selected group of people). There are a number of related terms such as "social networks", "social networking sites", "social media networks", "social news sites", and "online communities". Discussing the relationships between them and the relative merits of competing definitions is outside the scope of this thesis, and the above definition is used because of its clarity and ongoing popularity.

Social media are no longer a technological novelty, but have now become established communication platforms. Facebook reported having 2.41 billion monthly active users as of 30 June 2019 (Facebook 2019), which is about one third of the world's population (CIA 2019). According to estimates, YouTube has an even higher number of website visits than Facebook (SimilarWeb 2019). Businesses attempt to engage with consumers by advertising on social media (Lee et al. 2018). Politicians use them to send messages to their supporters (Fromm et al. 2018). Emergency services use them to gain situational awareness and relay information to the public (Reuter et al. 2016). Users make important decisions partly based on the information, opinions, and emotions they encounter on social media: for example, which products to purchase, which physicians to consult, or which political candidate to vote for (Goh et al. 2013; Matook et al. 2015).

After an initial phase of enthusiasm during which social media were heralded as a return to "a world that was ruled by individuals instead of big corporations" (Kaplan and

Haenlein 2012, p. 103), much of the public and academic discussion about social media now focuses on its negative facets such as targeted disinformation campaigns (Berners-Lee 2017; Farkas et al. 2018; Mejias and Vokuev 2017). In a systematic overview of the negative consequences of social media use, Baccarella et al. (2018) stress the potential for conversations to be manipulated. An investigation into alleged Russian interference in the 2016 U.S. presidential election concluded that "the IRA [Internet Research Agency] operated a network of automated Twitter accounts (commonly referred to as a bot network) that enabled the IRA to amplify existing content on Twitter" (Mueller 2019). In other words, automated accounts are used to artificially inflate the reach and diffusion of content. These accounts are also known as social bots. Bot is short for (software) robot, and the term is used to denote many kinds of algorithmically controlled accounts (Ferrara et al. 2016). Information systems research studies social media networks (Kane et al. 2014) and the effects of social media on elections (Wattal et al. 2010). Given the widespread fears about the negative facets of social media, the question of which content diffuses on social media, and which actors may aim to influence this process, becomes especially relevant.

2.2 Perspectives on social media content diffusion

The diffusion of content on social media is a complex topic. In everyday speech, "diffusion" refers to "the action of spreading something throughout or over a wide area" (OED Online 2019a). Usage of the term "information diffusion" can be traced to the pre-Internet era, where it was used, among others, to refer to the spread of an unspecified message or piece of information through an abstract, mathematically modelled population (Rapoport 1953), and, in a sociological study, to the influence of social structure on the spread of farming knowledge (Lionberger and Coughenour 1957). As the following paragraphs show, it is used in the context of social media to refer to the spread of content such as pieces of information, URLs, opinions, or emotions. Authors use the term "information diffusion" with some variation in meaning, while other authors research the same topics but use different terms, yet their perspective is also critical for understanding the origins of the diffusion of information. Among the different strands of literature from different disciplines that tackle different facets of diffusion, three levels can be distinguished in particular.

The diffusion of a piece of information, an emotion, or an opinion begins with the decision by an individual to post a message online (Shen and Liang 2015; Ziegele et al. 2018). Others who see it then have to take, consciously or subconsciously, the decision whether or not to forward it (Bakshy et al. 2012). These decisions depend on various factors. For example, users are more likely to report that they will forward a video advert if it has a

pleasant emotional tone (Eckler and Bolls 2011) and users who perceive themselves to be in the minority with their opinion on a controversial topic may be less likely to post this opinion in an online discussion forum (Yun and Park 2011). The content a user has forwarded to others is usually visible on their social media profile, so motives such as self-presentation and impression management may play a role (Marwick and boyd 2011).

In comparison to this micro perspective, there is information diffusion research that relates the individual to their position in the overall network. This research might, for example, examine the paths that a given piece of information takes through a network, such as in the Independent Cascade and Linear Threshold models (Gruhl et al. 2004). The precise mathematical formulations of such models allow for tasks to be defined and solved, such as influence maximisation (Morone and Makse 2015).

The macro perspective on information diffusion abstracts from this network and focuses on the total extent of the diffusion and its temporal dynamics. These diffusion models include Rogers' diffusion of innovation theory (Rogers 2003), which can also be used to explain the diffusion of information (Gruhl et al. 2004). Other examples are the classification of topics into ongoing chatter and short spikes (Gruhl et al. 2004), or the Linear Influence Model (Yang and Leskovec 2010) which also falls into this category. Such models are also common in epidemiology, where the topology of the network is not necessarily known (Zafarani et al. 2014, p. 242). The research of Stieglitz and Dang-Xuan (2013) is also an example of this perspective. They found that emotionally charged Twitter messages were retweeted more quickly and more often than neutral ones. Similarly, Ferrara and Yang (2015) found that negative messages spread faster than positive messages, but positive ones reach larger audiences.

Many of these studies have the fact that they study the effect of various factors on diffusion in common. These are the factors that lead individuals to post or forward a message or not. At the network perspective, they may influence which path a piece of information takes. At the aggregate crowd level, they determine the total extent to which a piece of information spreads.

The first kind of such factors are individual characteristics. Overall activity on social networking sites is influenced by efficacy of self-presentation, or an individual's perception that they are in control of which image they present to others (Krämer and Winter 2008). Users with a self-interest incentive for using Facebook who, for example, seek to gain a sense of achievement by getting 'likes', are more likely to share commercial messages (Fu et al. 2017). In the case of the spread of opinion, an example would be the author's fear of being socially isolated due to being in the minority with their opinion (Yun and Park 2011).

Network characteristics may equally affect the diffusion of information. These characteristics include the relationship between the user posting the information and the user deciding to spread it (Bakshy et al. 2012). They also include the structure of the network relative to the spreader of a hashtag; for example, the network of early adopters of political hashtags was found to have a higher triangle count than early adopters of general hashtags (Romero et al. 2011).

Finally, there are content characteristics that affect the spread of information. Examples include, as mentioned above, the emotional valence of the message (Ferrara and Yang 2015; Stieglitz and Dang-Xuan 2013). These factors can be highly domain-specific: in health tweets, perceived efficacy (with regard to solving a health problem) has been found to be a significant predictor of diffusion size (Meng et al. 2018).

It has become clear that to develop a comprehensive understanding of information diffusion, it is necessary to understand the interplay of the different levels, and to take into account more than one type of factor that influences information diffusion. A purely technical approach based on social network analysis might describe which content spreads on social media but fail to explain the underlying reasons which are linked to the decisions of the users. Neither is a purely individualist view that ignores the shape of the network that connects us to others sufficient.

Since one of the topics of this thesis is manipulation on social media, and especially on bots, it should also be noted that the behaviour of bots differs from human behaviour in several ways, which could influence information diffusion. At the individual level, they have no interests or preferences of their own, but are programmed to search for specific content to share (e.g. every tweet that contains a hashtag) or to post specific messages. In terms of network characteristics, they are found in specific network positions, e.g. in densely interconnected groups (Ferrara et al. 2016). Finally, the content they share is also different from what people post. Previous research has, for example, found that messages by some bots are more likely than humans to include URLs in their tweets (Chu et al. 2012). These differences, when compared to humans, suggest that information by bots might also spread differently from information posted by humans.

3 Research design

This chapter summarises the research design: the overall research strategy and the methods used in the individual papers.

3.1 Research strategy

Figure 1 shows the relationship between the papers and the research questions. The first three papers are devoted to addressing the first research question. They summarise challenges that are encountered, and point the reader to possible consequences and solutions.

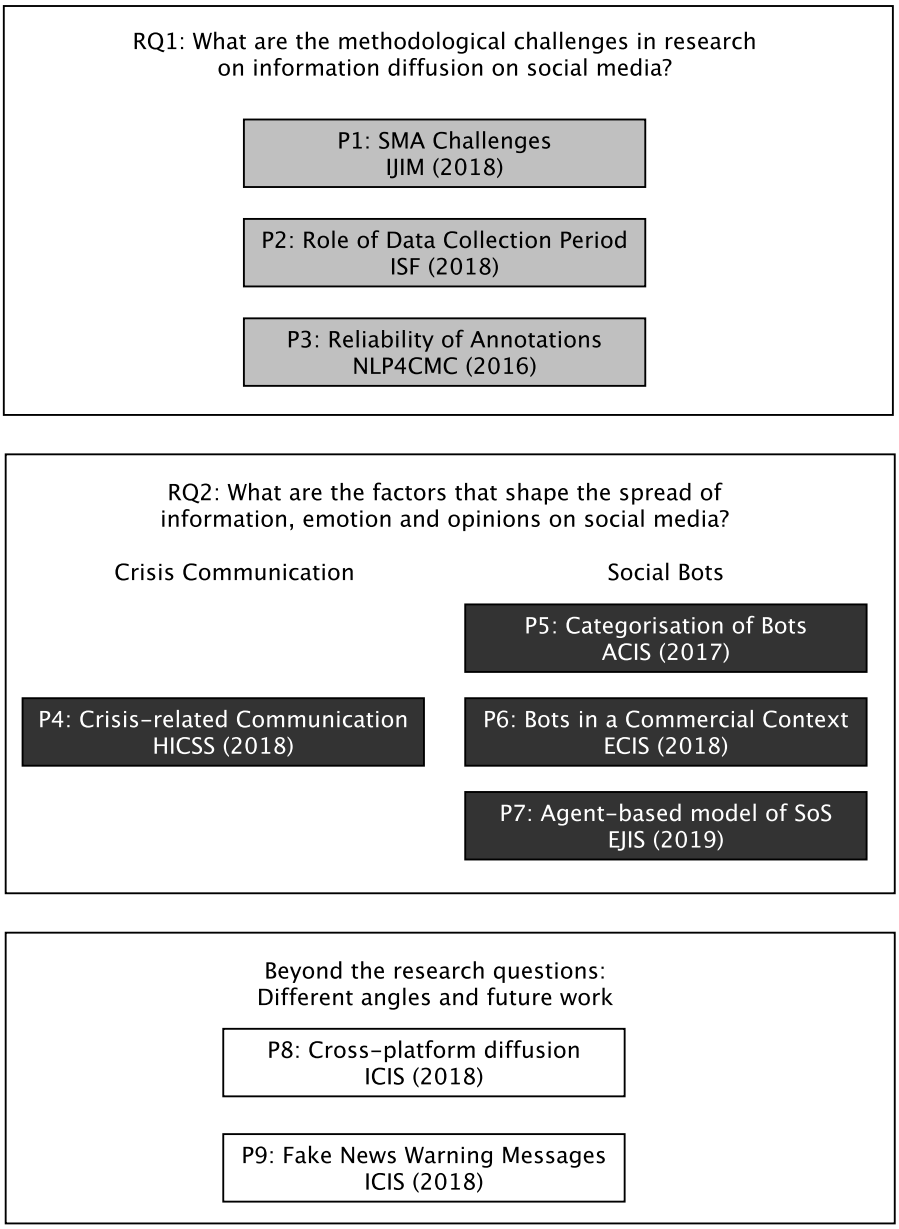

Figure 1. Relationship between individual papers and research questions

The next four papers address the second research question, regarding the factors that shape the diffusion of information, emotions, and opinions on social media. Two topics are discussed in these papers: crisis communication and social bots.

The final two papers are not directly related to either of the research questions. Instead, they go beyond the core of the thesis and approach the problem from two different, equally important angles. They offer preliminary glimpses beyond the horizon, and into topics that should be addressed in future research.

3.2 Applied research methods

Table 2 shows the research methods that were used in the papers. Data sources included Facebook, SoundCloud and news websites as well as Twitter, online surveys, and a simulation designed for a virtual experiment. The data analysis methods included statistical analysis, including regression models, and social network analysis.

Two of the papers in this dissertation are standalone literature reviews. Literature reviews can be useful tools to get an overview of a research field, not just for the researcher conducting them, but for a wider audience. By summarising past research and analysing it with regard to specific aspects, they can, for example, help highlight areas that have received less attention than others in the past. They can also critically reflect about the past and future development of a research area. The two literature review papers in this thesis followed the general principles described by vom Brocke et al. (2009, 2015) and Webster and Watson (2002). Academic data bases served as the data sources. Search results for specific search terms were scanned and the relevance of the papers judged based on their titles and abstracts. In case they were relevant, they were read in their entirety and their contents systematically summarised. The details of this procedure varied and can be found in the respective publications.

A central concept in my thesis is that of social media analytics. “Analytics” is the “collation and analysis of data or statistics, esp. by computer, typically for financial or commercial purposes; the data that results from this” (OED Online 2019b). According to Stieglitz et al. (2014, 2018), social media analytics in particular is an “emerging research field that aims on combining, extending and adapting methods” for the “analysis of social media data”. In this thesis, social media analytics is viewed as a collection of methods and tools for analysing data that is gathered by the researcher or otherwise interested party from the social networking site. The data reflects actions and interactions that have taken place on the platform, without the intervention of the researcher. This data has also been described as “online personal and social data” (Ruths and Pfeffer 2014) or “digital trace data” (Howison et al. 2011). In the dichotomy between experimental and observational

research, this data is observational. By this definition, four of the papers in this dissertation can be classified as using social media analytics to address their research questions, and they more or less explicitly follow the process model described by Stieglitz et al. The data sources in these studies were Twitter, Facebook, and Soundcloud. A fifth paper, a literature review, contributes to the development of the field of social media analytics by examining its challenges.

Social media analytics, in this definition, contrasts with other methods and approaches for understanding the usage of social media and the diffusion of information in online social networks. A notable example is the family of experimental approaches, in which the researcher constructs a controlled environment and systematically measures the effect of an intervention. Three of the papers in this thesis follow an experimental approach. In one case, this experiment was entirely virtual. Such a simulation gives the researcher the greatest possible degree of control over the environment, although it is necessary to carefully ground the assumptions in reality and validate the model to be able to make inferences that are relevant outside the simulation model (Wilensky and Rand 2015). In the other two cases, these experiments took the form of an online survey. Participants were confronted with realistic representations of social media messages and asked to record their response. In both cases, the participants were divided into groups, and each group was presented with slightly different material – stimuli in the terminology of psychology – to examine the effect of varying material on the result.

Table 2. Overview of applied research methods

Paper no.	Research approach	Data collection method(s)	Data analysis method(s)
1	Standalone Literature Review	Academic databases	Qualitative and quantitative summary, social network analysis
2	Social Media Analytics	Twitter API	Statistical analysis
3	Survey experiment	Twitter API and online survey	Statistical analysis
4	Social Media Analytics	Facebook API	Statistical analysis
5	Standalone Literature Review	Academic databases	Qualitative summary
6	Social Media Analytics	SoundCloud API	Quantitative summary and social network analysis
7	Virtual experiment	Simulation	Quantitative summary and social network analysis
8	Social Media Analytics	Twitter API and news websites	Statistical analysis
9	Survey experiment	Online survey	Statistical analysis

4 Research Results

This research addresses the methodological challenges that arise when examining social media data, as well as the factors that affect the spread of information, emotion and opinions. In the following section, the findings on these questions are presented.

4.1 Challenges related to the methods

This section summarises and discusses the challenges examined in papers 1-3. These challenges are related to social media analytics. This research field, described by Stieglitz et al. (2014), is not just of interest to academic researchers. Despite the wide variety of approaches, it offers a common core of concepts and best practices.

Stieglitz et al. (2014) suggested some future research directions for this field. According to them, future research needs to take the dynamic nature of the data into account. Another one of the directions suggested was that machine learning-based approaches for classifying social media content need to be improved. Papers 1-3 answer this call for more research into the methods of social media analytics.

The definition of social media analytics as a research field speaks of the conviction that different methods and approaches for analysing social media data have much in common, and benefit from further development. Stieglitz et al. (2014) explicitly called for a "significant increase in the level of inter-disciplinary research co-operation". It can be argued that many papers are isolated case studies that collect a very specific data set for a very specific purpose and analyse it with a very specific method. Still, the people carrying out this research often perform similar steps.

As a relatively new field, research in social media analytics faces its own challenges (see e.g. Ruths and Pfeffer (2014). Paper 1 argues that there are already many academic papers that discuss the specific challenges that researchers face when they employ specific methods from social media analytics to analyse their data. However, analysis is the last step of the multi-step social media analytics framework. Paper 1 further argues that there is a relative scarcity of research challenges into the equally important steps that come before it, and that researchers would benefit from an overview.

Paper 1 argues that because of the continuing growth in social media usage, the collected data can be considered "big data". Challenges of big data are already well-researched, and the paper thus draws on concepts from this body of literature, namely, the four V's of big data (McAfee and Brynjolfsson 2012; Saha and Srivastava 2014). *Volume* refers to the quantity of data, i.e. the storage space required. *Velocity* refers to the speed with which

new data is created and captured. *Variety* refers to the many different forms and shapes that the data can take. Finally, *veracity* refers to uncertainty especially regarding data quality. Yet, these V's are broad categories rather than precise definitions of problems that point researchers to concrete solutions. Paper 1 was created out of the conviction that the latter would be useful for the field to succeed.

As a result, a literature review was conducted, and the challenges mentioned in the social media analytics literature were analysed with the concepts from the big data literature in mind. Apart from a short quantitative overview of the papers found, the results were presented in the form of a table that maps the papers to the big data V's, according to their challenges. This table was then used to identify similar papers. A qualitative approach was used for a more detailed analysis of the findings. This resulted in the identification of the following five major challenges:

1. Bridge the gulf between the social and the computational sciences
2. Discover relevant topics and events
3. Choose an appropriate software architecture and storage technology
4. Obtain high-quality data
5. Visualise the data meaningfully

Each of these challenges was found in many papers, sometimes from widely different application domains. Paper 1 contains a detailed discussion of each challenge, summarising the context in which they were mentioned, and discussing possible solutions. Paper 1 also discusses these challenges in light of the original social media framework (see Figure 2).

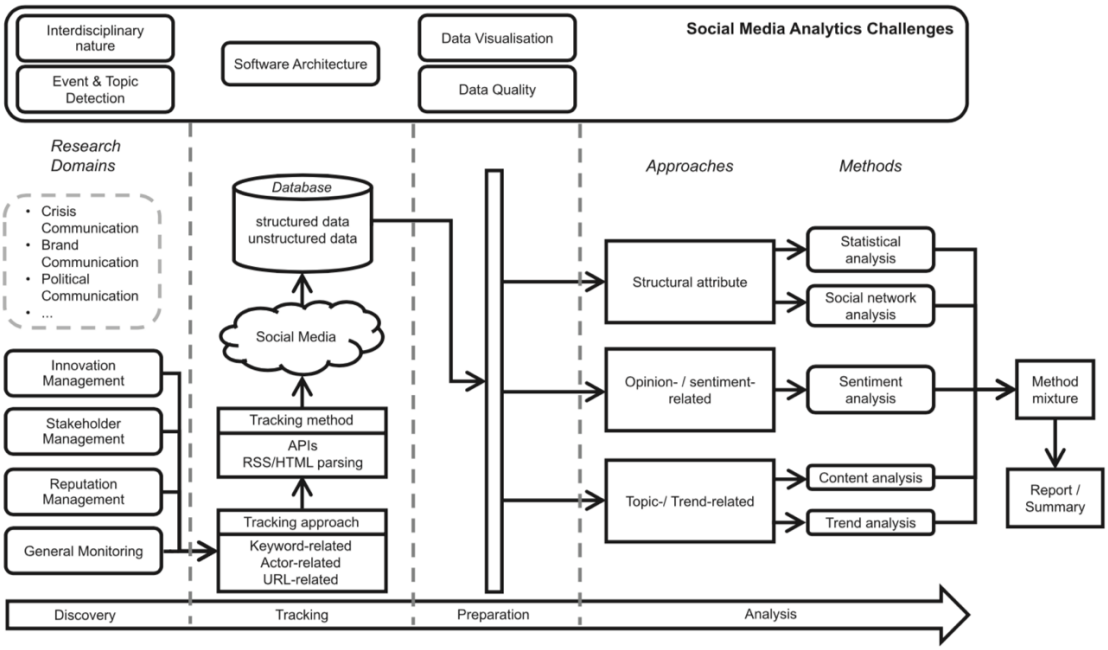

Figure 2. Social Media Analytics Framework (Source: Paper 1)

One of the very first decisions to take when analysing social media data is which data to collect. The results from Paper 1 provide some insights into this: The challenge of discovering relevant topics and events to study is in part addressed by the event detection literature, which offers methods to detect events based on texts and time series. However, sometimes the event of interest is known, and the question that the researcher faces is which data to include and which data to exclude, based on the time of publication of the data (or social media message). Similarly, Paper 1 describes the challenge of obtaining high-quality data, but the solutions in the literature mostly refer to the problems of ignoring missing data – data that is known to be missing, to be precise – and method of inferring this data. Paper 2 therefore goes beyond these results and addresses the specific question of which time period is best for data collection. Since the results of Paper 1 were based on the challenges that are frequently mentioned in the social media analytics literature, and the choice of time period was not among the identified challenges, it appears that this challenge is not frequently mentioned. Paper 2 similarly notes that there are few studies that have addressed this problem. However, this does not mean that this choice is irrelevant or does not pose a problem.

To the contrary, Paper 2 highlights the ongoing debate on the extent to which social media can offer an unbiased representation of public opinion. The paper is in part motivated by the observation that existing studies differ widely in the time frame that is chosen to collect the data which is subsequently used to make the predictions. To rephrase one of the arguments made in Paper 2 in more general terms, research that makes use of opinions

expressed on social media should take psychological theories into account that explain why people are likely to disclose their opinions on social media, and also when. Paper 2 studied the relationship between an artist's performance in the Eurovision Song Contest, a televised competition between participants from different European countries, and the tweets surrounding this event. The winner is decided by an audience vote, and the event is discussed enthusiastically on Twitter by part of the audience, which suggests that the contents of the tweets might have some bearing on the voting results. The analysis revealed that there is indeed a positive relationship between the number of tweets about an artist and their performance. However, tweets collected during the event were shown to be much less useful for the prediction than the tweets from before the event (Table 3, H3).

Table 3. Overview of hypotheses and results from Paper 2

Hypothesis	Results 2015	Results 2016	Conclusion
H1: There is a consistent, replicable positive relationship between the number of artist-related tweets and a better artist ranking in the audience voting.	Supported	Supported	Supported
H2: There is a consistent, replicable positive relationship between the sentiment of artist-related tweets and a better artist ranking in the audience voting.	Supported	Not supported	Not supported
H3: The explanatory power of artist-related tweets from prior to the event is higher than for those from during the event. This relationship is also valid across more than one year.	Supported	Supported	Supported

This second finding is perhaps the key result from this paper, with potential ramifications far beyond the scope of predicting events. The data collection period can have a profound impact on the results of a social media analytics study. Special care thus needs to be taken in justifying the time period. This result is all the more relevant when considering that this choice is arguably sometimes overlooked, sometimes made arbitrarily.

In the method that was employed in the paper, Paper 2 offers another innovation compared with much of the social media analytics literature. Because data from 2015 and 2016 were compared, the 2016 analysis of the data can be considered a replication of the 2015 analysis. There have been growing calls for researchers to repeat their studies (Dennis and Valacich 2014). The result that data from before the event resulted in better

predictions than data from during the event held up when tested both on the 2015 and the 2016 data. However, another result of the paper is that, in fact, one of the other hypotheses supported by 2015 data (that there is a relationship between sentiment and artist performance) was rejected when the same calculations were carried out using the 2016 data. This observation further reinforces the result that the choice of time period matters – to paraphrase Paper 2, not just on the small scale of choosing when to start and when to finish data collection, but also on a much larger scale.

When laying out future research directions for social media analytics, Stieglitz et al. (2014) called for the "machine learning-based classification of social media textual content and recognition of social network patterns" to be improved beyond the state of the art then. Paper 1 found that social media data can be noisy and unreliable.

In this spirit, Paper 3 examined the reliability of annotations of social media data for machine learning-based classifications of hate speech. The year 2015 had seen an unprecedented influx of refugees and migrants into Germany, and the results were evident on social media. Users discussed the effects of this development on society, and some expressed hate toward refugees. Regulators became active, with the European Commission asking internet companies to sign a *Code of Conduct on countering illegal hate speech online* (European Union Internet Forum 2016) and Germany passing a new law, partly for this purpose (BBC 2018). Their approach consisted of requiring platforms to quickly check posts once they are reported by users, necessitating large-scale content moderation. Against this backdrop, functioning methods for automatically detecting hate speech online would appear to be a useful tool.

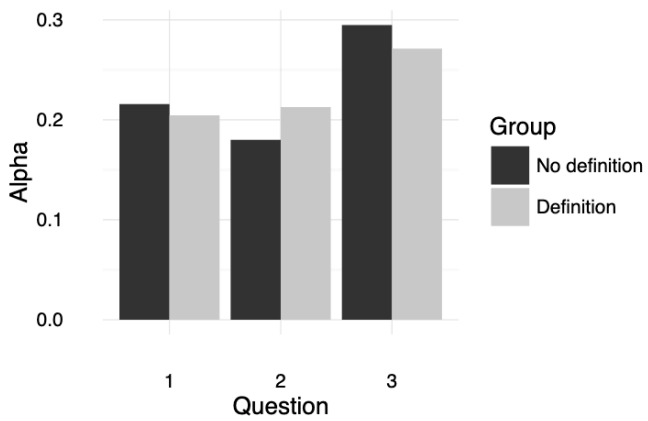

Figure 3. Reliability (Krippendorff's α) for the different groups and questions in Paper 3

Paper 3 found that when participants were asked to decide whether or not a specific Twitter message was hate speech, the reliability of these annotations, as measured by commonly used metrics, was low (Figure 3). Crucially, the reliability was still far below conventional minimum standards when the users were presented with the definition of hate speech that was actually used by Twitter.

In summary, papers 1-3 show various challenges that are commonly encountered in social media data and in part they propose possible solutions.

4.2 Diffusion of information, emotions, and opinions

This section addresses the diffusion of information, emotions, and opinions. It summarises the findings from papers 4-9.

4.2.1 Crisis communication

Communicating information to the public in a timely manner and disseminating this information to as many people as possible becomes an especially critical goal in crisis situations. Disasters and ongoing threats force emergency service agencies (ESAs) and other actors such as municipalities to react rapidly to provide the public with accurate information and to contain the spread of rumours and misinformation. Social media are among the methods used for achieving these aims (Xiao et al. 2018).

Paper 4 argues that social media analytics can play a crucial role in helping ESAs understand the reactions to their content on social media. It presents an analysis of Facebook posts made by emergency service agencies and municipalities during and after three terrorist attacks. The process model by Stieglitz et al. (2014) that was extended with the challenges in Paper 1 is closely followed. Determining how the diffusion of posts depends on their content required a manual annotation of posts, much like the hate speech tweets in Paper 3. However, in this case, a satisfactory reliability was achieved, which allowed the analysis to proceed.

A negative binomial regression analysis was carried out for Paper 4 to explain the number of shares and the number of reactions to the posts, based on the Facebook page that published the post, the logarithm of the length of the post in characters, the presence or absence of images and videos, and the content categories from the annotation. It revealed significant effects of the Facebook page and text length on the number of shares, and significant effects of the Facebook page, the presence of images and videos and the post content category of *condolences* on the number of reactions. These results are restated in the paper in simpler terms, with practitioners in mind:

- Keep your posts concise. Doubling their length will decrease their number of shares by a third, all else equal.
- Use image and/or video along with text in your post. An image will approximately double user reactions, while a video will quadruple them.
- The most important factor, however, that determines both the number of shares and other reactions is the number of followers and other variables outside your immediate control. Prepare accordingly and encourage as many people as possible to subscribe to your updates by liking or following your page.

Paper 4 further examined the relationships between various emotional reactions to posts, reflected by different emojis that indicate, among others, happiness, anger, and sadness. It found that this emoji usage data contains useful information for understanding and quantifying the different types of content in the posts, since some pairs of emojis were uncorrelated in their usage and the posts that received an especially high proportion of reactions of each type clearly differed from one another in their textual content. The results further suggested that participants expressed different emotions at different stages of the discussion, corroborating the notion from Paper 2 that the time period of data collection is important.

4.2.2 "Social bots": Automated communication

Despite the clear advantages for communicating information to the public and for discussions between users about topics such as politics, there are growing concerns that social media could be used to spread political messages in an automated way, and even to interfere with the process of public opinion formation.

Paper 5 argues that the topic is a relatively new one, and thus there are still various, sometimes incompatible definitions, especially regarding the distinction between bots and *social* bots. To understand the differences and the subtleties in the definitions, and to understand which types of bots are theoretically possible and to what extent they are reportedly used, a literature review was carried out. Paper 5 noted that there was a fairly consistent increase in research on social bots over the observed time period (2010 – mid-2017).

The literature review resulted in the classification of bots shown in Table 4. It consists of two dimensions, intent and imitation of human behaviour. Intent (Ferrara et al. 2016) captures whether the intentions of those creating the bots are benign, neutral or malicious. Examples of benign bots would be bots that perform useful services, such as aggregating content. They could equally be used to disseminate disaster-related information to the

public in a scenario akin to that in Paper 4. Malicious bots are those that are designed "with a purpose to harm" (Paper 5). The second dimension, imitation of human behaviour, reflects the degree to which the bots are "designed to pass off as a human being" (Boshmaf et al. 2013), at least at a first glance.

Table 4. Categorisation scheme of social media bot accounts (Paper 5)

		Intent (Ferrara et al. 2016)		
		Malicious	Neutral	Benign
Imitation of human behaviour (Boshmaf et al. 2013)	High: *Social* bots	• Astroturfing bots (Ratkiewicz et al. 2011) • Social botnets in political conflicts (Abokhodair et al. 2015) • Infiltration of an organisation (Elyashar et al. 2015) • Influence bots (Subrahmanian et al. 2016) • Sybils (Alarifi et al. 2016; Goga et al. 2015) • Doppelgänger bots (Goga et al. 2015)	• Humoristic bots (Veale et al. 2015)	• Chat bots (Salto Martínez and Jacques García 2012)
	Low to none	• Spam bots (Wang 2010) • Fake accounts used for botnet command & control (Sebastian et al. 2014) • Pay bots (Subrahmanian et al. 2016)	• Nonsense bots (Wilkie et al. 2015)	• News bots (Lokot and Diakopoulos 2016) • Recruitment bots (Flores-Saviaga et al. 2016) • Public Dissemination Account (Yin et al. 2014) • Earthquake warning bots (Haustein et al. 2016) • Editing Bots, Anti-Vandalism Bots on Wikipedia (Tsvetkova et al. 2017)

According to Paper 5, bots that mimic human behaviour should be considered "social bots" (since they simulate at least some social skills or are presented as social actors). Following this definition, not all bots on social media are social bots, and other bots on social media are best referred to using an umbrella term such as "social media bot accounts". Table 4 also shows some of the papers found in the literature review, and the category that the bots they deal with best fits in, according to the authors' descriptions of the bots. It becomes clear that there have already been several research papers that fit this definition of social bots, including especially the malicious social bots whose potential influence has been the subject of widespread worries.

After these definitions in Paper 5, Papers 6 and 7 deal specifically with *social* bots. In particular, paper 6 presents an empirical analysis of bots currently active on social media, while paper 7 analyses their potential impact from a theoretical point of view.

Paper 6 focuses on bots active on SoundCloud, a music sharing and commenting platform with several social networking features. This analysis is notably different from many other analyses of bots, because it concerns a potential commercial application. The desire to influence human behaviour is not limited to the political domain, and there has been evidence that social media users partly rely on user-generated content that they encounter online when making non-political decisions such as which physician to choose (Carbonell and Brand 2018) or which product to buy (Erkan and Evans 2016). Paper 6 argues that it is therefore of interest to examine potential commercial uses of bots.

In response to this research question, Paper 6 finds a group of SoundCloud profiles with a particular commenting behaviour. They post a very high number of comments, but these comments are often textual duplicates of one another. This observation is formalised as a comment uniqueness score. Classifying accounts by their comment uniqueness score resulted in the observation that accounts with low comment uniqueness also tend to have far fewer followers, upload far fewer tracks and create fewer playlists, but repost far more existing tracks than other accounts. This behaviour is consistent with the expected behaviour of bots and contrary to a typical usage scenario by a real human, and the accounts are therefore considered potential bots. Moreover, Paper 6 presents a network of the users that uploaded the most commented tracks and the users that commented on these tracks. This network (Figure 4) shows a clear pattern of a centre of eight track authors whose tracks are frequently commented on by these potential bots.

While the results from this research-in-progress paper are not final, they suggest that bots play a role on SoundCloud, where a potential commercial use is not out of the question. These bots are "social" bots in the definition from Paper 5, in the sense that they engage in behaviour usually reserved for humans (i.e., commenting). An examination of some of these bot profiles revealed that they are not immediately recognisable as bots based on their profile picture and description, although their repetitive commenting behaviour clearly indicates some degree of automation.

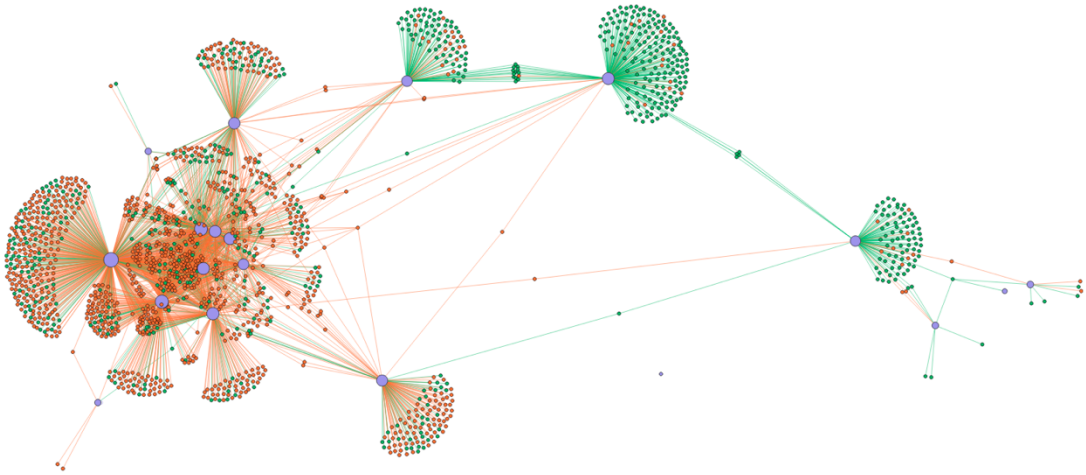

Figure 4. Network visualisation of the 25 most commented tracks on SoundCloud. Nodes represent users, edges comments. Purple nodes are track authors. Red nodes are users who posted many highly repetitive comments (high comment uniqueness score, potential bots). Green nodes are users who posted many diverse comments (low comment uniqueness score). Details can be found in Paper 6.

However, Paper 6 also stresses that it is unclear whether these suspected bots can successfully convince human accounts that a track is more popular than it actually is among real listeners, although this would seem to be a prerequisite for convincing them to listen to it and for the artist to benefit financially. Paper 5 pointed out that one of the bot tactics described in the literature is astroturfing, or creating the appearance of a "grassroots" campaign that does not truly exist.

Following these considerations, Paper 7 describes a potential mechanism by which bots could influence human behaviour, or more specifically, public opinion formation on controversial issues. The spiral of silence theory is translated into a networked agent-based model. The simulated agents observe the opinions that are expressed on the topic of interest in their environment, much like people do according to the spiral of silence theory. The agents then constantly update their own willingness to express their own opinion according to whether this opinion agrees with what is expressed by their peers.

Paper 7 describes how this process leads to the emergence of a spiral of silence, that is, the simulation ultimately reaches a stable state in which one of the two opinions is expressed by far more individuals than the other opinion, notwithstanding the fact that each opinion is held by approximately half the human population. The paper further reveals that this final state looks markedly different, depending on the density of the network: the

denser the network, the stronger the dominance of the majority opinion over the minority opinion, as evidenced by the numbers of individuals willing to express them.

Crucially, the paper then shows how the opinion that comes to be expressed by the majority, that is, accepted as a norm or consensus, can be influenced by inserting bots into the network. In the simulation, these bots are modelled as agents that do not pay attention to their environment at all, since they are programmed to advocate their "opinion" unrelentingly.

Paper 7 examines how the likelihood of success of such a strategy depends on factors such as the number of bots, the position of the bots in the network, the human likeness of bots. If there are no bots in the network, each of the two opinions will become the majority opinion 50% of the time. A key finding of Paper 7 is that if an individual bot is as effective in influencing a human contact as a real human, then a mere 2–4 % of bots are sufficient to increase this probability to 66%. Figure 5 shows how this probability depends on their position and overall network density.

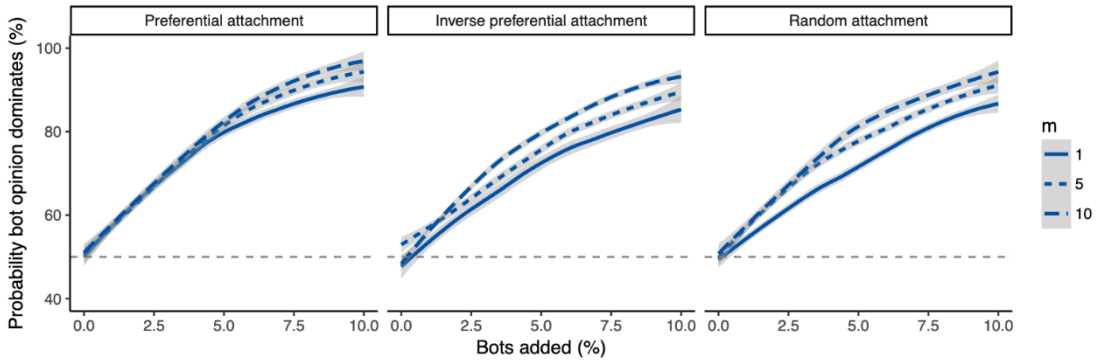

Figure 5. Influence of bots added at different positions in the network. Curves are Loess smoothers (span = 0.75) fit to the empirical probabilities to highlight the trend (Cleveland 1979). Grey areas around the curves are 95% confidence intervals. Straight dashed lines indicate the 50% baseline probability

4.2.3 Beyond understanding diffusion: Means of control

This section describes two studies that go beyond the core of this thesis to examine the same topic – the diffusion of information, emotions and opinions – from different angles. The studies that employed social media analytics in a broad sense (Papers 2, 4, and 6) each considered a single social networking site. However, many users use several online sources to acquire information on topics such as politics (Statista 2018). To go one step

further than the above studies, it is therefore necessary to consider more than one social media site at a time. Paper 8 therefore addresses the question of inter-media information diffusion.

In addition, the previously described papers are primarily concerned with understanding the spread of information, emotions, and opinions on social media. They barely touch the question of how this diffusion could be influenced by interested parties from an empirical angle. Although recommendations for improving the reach of disaster-related social media posts are made in Paper 4, these propositions are only empirically validated on observational data. Paper 7 also addresses the topic of influencing opinion formation, but offers a theoretical perspective, not empirical evidence. Paper 9 therefore addresses measures to influence the diffusion of content and offers evidence from an experiment.

Inter-media information diffusion can take many forms, but the one addressed in Paper 8 refers specifically to the observation that Twitter messages are often taken up by journalists, for example, to supplement a news story with reactions or opinions or because the tweets themselves are considered newsworthy. Equally, Twitter users discuss news stories online. Paper 8 argues that to suitably examine these phenomena, both data sources need to be considered at the same time. Therefore, both tweets and news stories were captured and cross-references detected automatically, based on URLs. The results suggest that those cross-references are very common. Around 10% of the tweets contained a URL belonging to a news site. However, the news websites differ greatly both in how often they mention Twitter content in their stories and how often their stories are discussed on Twitter. Paper 8 shows in which cases these differences are statistically significant, and proposes more research into the causes of these inter-media diffusion events. For example, examining which type of news story is most likely to be shared and discussed on social media is likely to yield new insights into the interaction of social media and news content. It could also help with identifying the causes of the sometimes viral diffusion of misinformation or “fake news”.

Paper 9 focuses on a possible measure which could be used to stop the spread of such fake news, which has also been described as one of the activities of “social bots” (Paper 5). Social media companies could display warning messages when users share news stories that have been disputed by reputable fact-checking sources. One such measure was announced by Facebook in December 2016 and discontinued in December 2017. The goal of this study was to examine if such warning messages are indeed effective at helping users distinguish between fake news and real news. The study further examined whether a variant of this warning message that is more informed by recent research into the effective design of such messages could be more successful than Facebook’s own design. The

study used signal detection theory to study this phenomenon in an experiment conducted by online survey. Two groups of users were shown warning messages along with some of the stories, and a third group was shown no warning message at all. In line with previous research, it was hypothesised that the presence of a warning message should result in an improved ability to detect fake news, but also cause an increase in the number of false alarms, that is, true stories mistakenly identified as fake, because it should make users generally more distrustful of the news. Only the supposedly superior design informed by recent research should successfully lower both the number of falsely accepted fake news ("misses") and the number of false alarms. However, the results indicate that neither warning message design is effective. Flagging the news as disputed could not be shown to help them distinguish fake news from real news.

5 Discussion and implications

This thesis has two overarching research questions, regarding methodological challenges in information diffusion research and the factors that shape the spread of information, emotions, and opinions. Each of the following two sections is devoted to one of these questions.

5.1 Overcoming challenges in social media analytics

The first research question concerned the methodological challenges in research on information diffusion on social media. Three papers have been put forward to address this question. Although Paper 1 identified several challenges and the solutions that have been proposed, the need for more research into how they can be addressed is also emphasised. The results demonstrate that the early stages of the research process, before the analysis even begins, are important. When authors describe their methodology in a research paper, they often devote much more space to data analysis than the steps before it. If the level of detail in the description is indicative of the amount of effort that was put into the respective step, then the quality of the results might suffer.

When the three papers on the first research question are considered together, a clear picture emerges. Even in 2019, social media analytics is still a relatively young field, and standards are only just beginning to be established. Norms will need to be developed, and are perhaps emerging naturally, that describe how to proceed in a way that ensures high-quality research. The onus is on the authors, reviewers and editors alike to ensure that all steps of the research process are appropriately described and that the right choices are made. Examples of developments in this area include systematic comparisons of methods for gathering Twitter data, including the nonobvious ways this choice may affect research results (Felt 2016; Morstatter et al. 2013; Pfeffer et al. 2018), and an ethnographic study of social media researchers, in which they were interviewed about their data collection and sharing practices (Weller and Kinder-Kurlanda 2015). The process for social media research, including earlier stages, is increasingly being formalised and critically discussed (Thelwall 2018), and studies that explicitly make use of such a framework and describe all steps of the research process in detail have appeared in high-profile publication outlets (e.g. Suseno et al. 2018).

In this vein, papers 2 and 3 point out two very specific issues that need attention when choosing and describing an early stage of the research process. When the goal is to make predictions from social media data, the data collection period matters (Paper 2). It needs to be justified well. Hypothetically, it is also possible to demonstrate that the collection period does not matter in a specific case, by repeating the study with different collection

periods, and comparing the results. This suggestion ties in with growing calls made in recent years for the replication of studies to show whether results can be reproduced. Paper 2 reproduced its own result from the 2015 data (that pre-event tweets are more useful for the prediction than tweets from during the event) on the 2016 data but failed to reproduce another result (that the sentiment of tweets about an artist is related to their ranking). Ideally, researchers should also reproduce the results of other researchers.

However, reproducing results can be hard with social media data (Weller and Kinder-Kurlanda 2016). Some social media platforms only allow a select group of researchers access to their APIs, while others restrict data sharing and archiving. Researchers with large Twitter data sets, for example, are only allowed to share the IDs of tweets, not the text and other metadata, requiring others to re-download this data directly from Twitter. The attempt to "rehydrate" (Twitter 2019) will fail for tweets and accounts that have since been deleted, and metadata such as the number of followers may have changed considerably (Zubiaga 2018). Bruns (2019) discusses the impact of API restrictions on academic research in detail, spurred by their tightening after the *Cambridge Analytica* scandal. There is no easy solution, but one way forward might be to use more standardised data sets. A little known example is the collection of Twitter data in the Internet Archive (Sequiera and Lin 2017).

Paper 1 also pointed out that the social media research community is traditionally divided between computer scientists and social scientists, or, bluntly put, between those with the technical background to deal with large amounts of data generated by human interactions, and those with the theoretical understanding to put the data into its proper context and interpret it. However, there are increasing efforts to link the two. A decade ago, the field of "computational social science" was a vision for the future (Lazer et al. 2009). Today, it has become reality (van Atteveldt et al. 2019; Mann 2016).

In classification problems in machine learning, the quality of the training data is paramount. In an NLP task such as predicting whether or not a tweet is hate speech, ensuring this quality is not always easy (Paper 3). Yet, the quality of the data can have profound implications for the results. There are various ways of obtaining labelled examples, and they can each affect the quality of the final classifier. The availability and quality of training data is closely related to a point made in Paper 1. Apart from the skills divide, there is another divide in the research community, the data divide. Among those who have the skills to deal with the data, there are those who have access to it and those who perhaps cannot afford that. The decision by some social media platforms to share data only with selected researchers exacerbates this problem (Bruns 2019). In the case of training data, large corporations have the financial resources to pay human annotators for labelling large

amounts of data (Vengattil and Dave 2019), or they might even already have the data, as a by-product of another service they offer (Giannoulakis and Tsapatsoulis 2016). At the same time, Paper 3 also highlights that simply throwing more money at the problem might not make it go away: If the concept is ill-defined, and even we humans cannot even agree on it, no amount of training data is likely to provide a classifier with the necessary information.

5.2 Information, misinformation and influence

The second research question asked about the factors that shape the spread of information, emotion and opinions on social media.

A naïve approach to answering this question would be to compile a list of possibly relevant factors and test them systematically to identify those that are valid on every existing social media platform, in every situation that social media is used in. However, such an approach would likely be futile for several reasons. Social media data is of an enormous variety (see Paper 1), there can be immense differences in results even within one application domain and platform as soon as the time period of data collection is varied (see Paper 2), and complex textual statements are difficult to compress into numbers (see Paper 3).

Instead, the papers that I put forward to address the first research question have made it clear that research needs to take a more nuanced approach. Since the answers to the second research question are likely to differ widely between application domains, efforts to find them should focus on especially relevant applications, such as those where understanding information can help keep the population up to date in times of crisis, or where political decision-making processes are endangered. Paper 4 therefore focuses on social media crisis communication, and Papers 5–7 on the topic of social bots.

In Paper 4, the factors contributing to information diffusion are measured directly. By using the number of shares or the number of reactions as the response variable, and factors such as text length and the presence or absence of an image as predictors, it is possible to measure the influence each of them has in isolation, all else equal. The paper demonstrates the viability of social media analytics data in the context of researching aggregate information diffusion. It is not the first to do so, and should rather be seen as a new contribution to an already existing line of research, e.g. Stieglitz and Dang-Xuan (2013). In agreement with the findings from this line of research, the most emotional of the different types of content researched – condolences, as opposed to warning or encouragement – also received by far the highest number of (emotional) reactions.

The results also suggest that users expressed different emotions a few days after the crisis event than immediately afterwards. This is reminiscent of the key result from Paper 2: the data collection period matters. If some predictive model was based on the posts about the three crisis events studied in Paper 4, the appropriate data collection period would need to be considered carefully.

There are worries that the tendency for some pieces of information to spread rapidly may be exploited. Paper 5 confirmed that such worries also exist in research: Various papers describe attempts to hijack online discussions with the help of automated accounts, although there are also many kinds of other bots. This paper is the first to combine the two dimensions of intent – benign, neutral, or malicious (cf. Ferrara et al. 2016) – and the degree to which the account imitates human behaviour – high, low, or not at all (cf. Boshmaf et al. 2013) – into a comprehensive six-category system. It provides an overview of this topic, and allows individual bots to be located in a large framework and systematically compared to others based on their characteristics.

When looking at a social media data set, a suspicious set of accounts could be identified (Paper 6). Although there is no definite proof that they are bots, their behaviour is consistently like that attributed to and expected from bots. The degree to which these likely bots imitate human behaviour is low (due to the excessive repetitiveness of their comments) but an attempt was clearly made. While further analysis will be necessary to determine whether this is enough to influence the behaviour of human users, this finding demonstrates the viability of bots in a commercial context that could have financial implications for the parties involved. Given how central advertisement is to the business models of social media platforms, it appears likely that future developments in the realm of social bots will be exploited for commercial purposes. Research into the commercial use of bots is still scarce. In a rare exception, Cresci et al. (2019) report a practice they call “cashtag piggybacking” to promote low-value stocks with suspected bot accounts on Twitter.

One possible mechanism by which human users could be influenced was described in Paper 7. By triggering a spiral of silence, they might be able affect the dynamics of public opinion to their liking. Much of the public discussion about social bots has revolved around suspected attempts to directly influence the results of elections, perhaps by spreading misinformation (Stocking and Sumida 2018). In clear contrast to such speculation, our paper shows that a much more subtle form of manipulation could be possible. By influencing users in their willingness to speak about a specific topic, the bots might affect what is talked about. The discussion on social media will also influence the wider public debate, for example when journalists and bloggers turn to social media for topics to write

about (Paper 8). The setting in Paper 7 is different from the music industry example in Paper 6, since there is an assumption that a contentious issue is being discussed and each actor is assumed to have one of two diametrically opposite opinions with a probability of 50% each. This is more likely to be true of, for example, a proposal for a law, to which half the population might be strictly opposed while the other half favours it clearly, than of something of which people might have more nuanced opinions. Nevertheless, Paper 7 shows how bots could influence opinion formation. The model itself is based on an earlier model of the spiral of silence by Sohn and Geidner (2016), but the application to bots is new. This study answers the call for more research into the dynamic nature of the spiral of silence theory (Matthes and Hayes 2014; Scheufele 2007). The method distinguishes it clearly from studies that developed bot classification systems or used such systems to measure the number of bots in specific conversations. The discipline of Information Systems studies how the internet affects political campaigns and elections (Wattal et al. 2010) as well as their commercial applications (Paper 6), and research into social bots plays an important role in this regard.

The fact that bots may be influencing the discussion on social media, not only by posting messages themselves, but also by influencing what the human users are willing to say, also ties in with the results from Papers 1 and 2. Not only the period of data collection matters, but also the set or subset of users whose expressed thoughts and feelings are taken into account to generate statements or predictions. This adds to the uncertainty in social media data and the challenges in using it for research. An "awareness of what is actually being analysed" (Ruths and Pfeffer 2014) is as important than ever.

In the past years, social media companies such as Twitter have recognised the threat posed by malicious social media use, and begun to take action against it (Twitter 2018). It remains to be seen whether social bots will be established permanently as a successful tool of manipulation, with programmers perhaps playing cat and mouse with platform operators.

The majority of research papers about social bots address the malicious kind. There is potential in researching benign bots that perform useful tasks. If we accept them as part of the social media landscape, we will need to discuss how they should be dealt with. For example, a bot that automatically replies to misinformation with links to fact-checking websites might be accepted by most users as helpful, but when it comes to politically sensitive topics such as climate change, this bot could quickly be labelled a political agent by its opponents. Who, then, should be responsible for deciding which bot is helpful, and which one harmful and should be deleted? A similar debate is currently taking place on the topic of hate speech (see Paper 3), since the European Union has decided to leave the

responsibility of deciding which post is deleted to platform operators (European Union Internet Forum 2016). To what extent should political advertisements on social media be regulated? These questions will need to be addressed. It also needs to be stressed that a large role is played by the users' media literacy. The results in Paper 7, after all, also showed that whether a social media user is influenced by a bot to the same extent as they are by a human plays a crucial role in determining the bots' success in shaping public opinion.

Paper 9 suggests a similar conclusion, as the attempt to influence users' likelihood of accepting a news story on social media as true by showing a warning message were unsuccessful. If social media are to be truly user-centred, then it is up to the users to hone the ability to distinguish the real from the fake, and the automated from the authentic.

5.3 Contribution to research

Of the ways in which the studies in this dissertation contribute to research, I would like to highlight two especially. The first is an improved understanding of the challenges involved in social media analytics approaches: the many difficulties faced by researchers from the very beginning of their research process, the importance of choosing the right data collection period, and the difficulty in obtaining high-quality training data for text classification. Only if we understand the problems in our data – the subtle side effects of choices, or the many ways in which the right method with the wrong data can lead to the wrong conclusions – can we avoid mistaking noise for signal and obtain accurate results.

Another key contribution is that this dissertation proposes a mechanism for the influence of malicious individuals and groups on public opinion. There has been a multitude of studies about identifying bot accounts, or describing their behaviour, Paper 6 included. Paper 7 goes a step further by giving their supposed influence on human behaviour a theoretical foundation. The spiral of silence is a plausible mechanism by which these individuals could influence opinion formation. The results indicate the conditions under which this influence is more likely to be large, or small. It should be noted that although the paper largely discusses this phenomenon in the context of social bots, automated accounts are not the only ones who could abuse the mechanism. There have been reports of paid trolls who were involved in disinformation campaigns on political and other topics. It is hard to say how much of their activity is automated. If the goal is to research political influence, research to determine whether an account is human or a bot might be asking the wrong question. Perhaps more relevant is the question whether they can really succeed in influencing users in this manner in real-life discussions. In the coming months and years, this will need to be carefully observed.

5.4 Practical implications

In terms of practical implications, a few are especially worth highlighting. Social media analytics is not confined to academic research, and neither are its challenges. The visualisation (Figure 2) can be used as a tool to ensure that while following the process, none of them are overlooked. Paper 1 devotes a section to each challenge and points the analyst, whether they are an academic researcher or a private-sector social media manager, to possible solutions. Approaches to the problems discussed in papers 2 and 3, which equally may arise in many applications outside academic research, are also discussed there.

In terms of reaching a social media audience, simple guidelines for how to phrase social media posts effectively were presented as part of the results of Paper 4. The observation that the emotional posts containing condolences attracted more reactions than factual posts about warnings or encouragement leads to a problematic paradox: ESAs are interested in maximising audience engagement, yet social media users apparently prefer emotional messages. Should ESAs consequently insert emotional content such as condolences into their feeds in an effort to keep users engaged? Such strategies are already used in social media marketing, where emotional attachment is thought to strengthen an individual's relationship with a brand (Hudson et al. 2015). Deliberately posting tear jerkers could arguably be seen as unethical. It might boil down to the age-old question of whether the end justifies the means.

At the same time, it needs to be considered that some of the content on social media may be automatically generated, or automatically spread beyond its natural reach. Legislators, journalists, and social media users need to be aware of activities of malicious accounts such as social bots in order to come to well-informed decisions. They should be aware, for example, that a high number of 'likes' is not necessarily an indicator of agreement; it could simply be an indicator of the effort that someone is willing to expend to convince the public that there is such widespread agreement. Social media marketers will need to observe future developments closely. Since using bots to promote products or buying likes is considered 'foul play' by users, businesses could have to deal with competitors buying likes for them to damage their reputation.

5.5 Limitations and an outlook

Despite my best efforts, this thesis is of course not without limitations. Each of the papers contains a discussion of its limitations, and the most important ones are highlighted here. First, many of the papers describe social media analytics studies. Social media companies' policies rarely allow researchers to share their data. In a world dominated by a growing open science movement, this state of affairs is a pity. It severely limits researchers'

opportunities to replicate others' studies, as demanded in Paper 2. However, it is still possible to replicate the results on other data, and the results can sometimes be different. For example, do the results about the diffusion of Facebook posts by emergency service agencies (Paper 4) also hold when a different type of crisis is examined, such as a natural disaster? Can bot-like accounts similar to the ones found on SoundCloud (Paper 6) also be found on other social media? These limitations are at the same time open questions that invite more research on the topic. Secondly, paper 7 is a simulation. It shows how bots could influence humans, not how they did so in a real-life setting. This has advantages, such as being able to observe the internal states of an individual, but also obvious disadvantages. There is much research left to do to examine under which conditions this proposed mechanism indeed takes place in this form on social media.

The discussions of the challenges mentioned some solutions but other challenges remain unresolved. The field of social media analytics will need to continue to develop. While topic discovery and event detection can look back to a long history, there will need to be many more discussions about data collection practices and data processing standards, especially since the subject of our research, social media, continues to evolve. Images are becoming increasingly important, as can be seen, for example, in the rise of Instagram, and there are signs that the future will see more and more video content. As deep learning techniques such as generative adversarial networks allow the production of realistic fabricated videos (Suwajanakorn et al. 2017), misinformation may become increasingly difficult to distinguish from authentic news, and researchers working on its detection will need to keep pace.

Social media is also offering content creators more and more possibilities to gauge their audience's reactions. An example of this is Facebook's Reactions feature, where users can select an emoji instead of simply 'liking' a post. Incidentally, developments such as this also afford researchers a better look into how social media users react to things they see online. If more data like this became available, it could be immensely valuable for social media analytics, since the reaction of a user to a post can otherwise only be measured in a laboratory setting, which inevitably takes place on a much smaller scale. However, this research can only be carried out effectively if Facebook allows academics adequate access to its data (Bruns 2019).

Regarding the influence of bots in particular, the simulation of social media interactions in particular proved a valuable tool, based on the notion of complex adaptive systems. Only the micro-level behaviour of the individual was programmed into it. The macro-level behaviour of the crowd emerges and can be observed. The motivation for this approach also partly stems from the fact that an observation on such a large scale is both

practically impossible and ethically questionable, especially when the topic is opinion manipulation (cf. Flick 2016). A central problem of social media research is that it is not possible to peer into the minds of those who remain silent (Chung 2015). This simulation approach can be used for addressing more research questions about social bots and other questions about social media behaviour. For example, what happens when the number of people advocating each of two diametrically opposed opinions are not equal? In such a scenario, it could be much harder for the bots to influence the users' opinions. How do dynamic graphs, which mirror the real-life changes to social networks over time (Kossinets and Watts 2006), affect opinion formation? How do sudden external events affect the network, for example, under what circumstances can they overturn an already established consensus? The simulation approach makes it possible to explore these questions, although proper validation is crucial. Further, reproducibility is not an issue, as code can be freely shared between researchers.

In summary, there are many avenues for future research, and the quick pace of developments in this field is likely to create even more. If challenges to research can be solved, and the factors that influence information diffusion in various contexts be identified, then research can play a crucial role in combating the "dark side" of social media, such as the spread of misinformation.

References

Abokhodair, N., Yoo, D., and McDonald, D. W. 2015. "Dissecting a Social Botnet," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, pp. 839–851 (doi: 10.1145/2675133.2675208).

AIS. 2019. "International Conference on Information Systems," (available at https://aisel.aisnet.org/icis/; retrieved September 2, 2019).

Alarifi, A., Alsaleh, M., and Al-Salman, A. 2016. "Twitter turing test: Identifying social machines," *Information Sciences*, (372), pp. 332–346 (doi: 10.1016/j.ins.2016.08.036).

van Atteveldt, W., Margolin, D., Shen, C., Trilling, D., and Weber, R. 2019. "A Roadmap for Computational Communication Research [draft version]," *Computational Communication Research*, (1:1).

Baccarella, C. V., Wagner, T. F., Kietzmann, J. H., and McCarthy, I. P. 2018. "Social media? It's serious! Understanding the dark side of social media," *European Management Journal*, (36:4), pp. 431–438 (doi: 10.1016/j.emj.2018.07.002).

Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L. 2012. "The role of social networks in information diffusion," in *Proceedings of the 21st Annual Conference on World Wide Web*, Lyon, France: ACM, pp. 519–528 (doi: 10.1145/2187836.2187907).

BBC. 2018. "Germany starts enforcing hate speech law," (available at https://www.bbc.com/news/technology-42510868; retrieved September 2, 2019).

Berners-Lee, T. 2017. "Three challenges for the web, according to its inventor," (available at http://webfoundation.org/2017/03/web-turns-28-letter/; retrieved February 2, 2019).

Boshmaf, Y., Muslukhov, I., Beznosov, K., and Ripeanu, M. 2013. "Design and analysis of a social botnet," *Computer Networks*, (57:2), pp. 556–578 (doi: 10.1016/j.comnet.2012.06.006).

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., and Cleven, A. 2009. "Reconstructing the giant: On the importance of rigour in documenting the literature search process," in *Proceedings of the European Conference on Information Systems (ECIS)*.

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Cleven, A. 2015. "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research," *Communications of the Association for Information Systems*, (37) (doi: 10.17705/1CAIS.03709).

Bruns, A. 2019. "After the 'APIcalypse': social media platforms and their fight against critical scholarly research," *Information, Communication & Society*, (22:11), pp. 1544–1566 (doi: 10.1080/1369118X.2019.1637447).

Carbonell, G., and Brand, M. 2018. "Choosing a Physician on Social Media: Comments and Ratings of Users are More Important than the Qualification of a Physician," *International Journal of Human-Computer Interaction*, (34:2), pp. 117–128 (doi: 10.1080/10447318.2017.1330803).

Chu, Z., Gianvecchio, S., Wang, H., and Jajodia, S. 2012. "Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?," *IEEE Transactions on Dependable and Secure Computing*, (9:6), pp. 811–824 (doi: 10.1109/TDSC.2012.75).

Chung, J. E. 2015. "Antismoking campaign videos on YouTube and audience response: Application of social media assessment metrics," *Computers in Human Behavior*, (51), pp. 114–121 (doi: 10.1016/j.chb.2015.04.061).

CIA. 2019. "The World Factbook," (available at https://www.cia.gov/library/publications/resources/the-world-factbook/geos/xx.html; retrieved September 2, 2019).

Cleveland, W. S. 1979. "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, (74:368), pp. 829–836 (doi: 10.1080/01621459.1979.10481038).

Cresci, S., Lillo, F., Regoli, D., Tardelli, S., and Tesconi, M. 2019. "Cashtag piggybacking: Uncovering spam and bot activity in stock microblogs on twitter," *ACM Transactions on the Web*, (13:2) (doi: 10.1145/3313184).

Dennis, A., and Valacich, J. 2014. "A Replication Manifesto," *AIS Transactions on Replication Research*, (1) (doi: 10.17705/1atrr.00001).

Eckler, P., and Bolls, P. 2011. "Spreading the Virus: Emotional Tone of Viral Advertising and Its Effect on Forwarding Intentions and Attitudes," *Journal of Interactive Advertising*, (11:2), pp. 1–11 (doi: 10.1080/15252019.2011.10722180).

Elyashar, A., Fire, M., Kagan, D., and Elovici, Y. 2015. "Guided socialbots: Infiltrating the social networks of specific organizations' employees," *AI Communications*, (29:1), pp. 87–106 (doi: 10.3233/AIC-140650).

Erkan, I., and Evans, C. 2016. "The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption," *Computers in Human Behavior*, (61), pp. 47–55 (doi: 10.1016/j.chb.2016.03.003).

European Union Internet Forum. 2016. "Code of Conduct on Countering Illegal Online Hate Speech," (available at https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/countering-illegal-hate-speech-online_en; retrieved February 2, 2019).

Facebook. 2019. "facebook Newsroom: Stats," (available at https://newsroom.fb.com/company-info/; retrieved September 2, 2019).

Fan, W., and Gordon, M. D. 2014. "The power of social media analytics," *Communications of the ACM*, (57:6), pp. 74–81 (doi: 10.1145/2602574).

Farkas, J., Schou, J., and Neumayer, C. 2018. "Cloaked Facebook pages: Exploring fake Islamist propaganda in social media," *New Media & Society*, (20:5), pp. 1850–1867 (doi: 10.1177/1461444817707759).

Felt, M. 2016. "Social media and the social sciences: How researchers employ Big Data analytics," *Big Data & Society*, (3:1), pp. 1–15 (doi: 10.1177/2053951716645828).

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. 2016. "The rise of social bots," *Communications of the ACM*, (59:7), pp. 96–104 (doi: 10.1145/2818717).

Ferrara, E., and Yang, Z. 2015. "Quantifying the effect of sentiment on information diffusion in social media," *PeerJ*, (1:e26) (doi: 10.7717/peerj-cs.26).

Flick, C. 2016. "Informed consent and the Facebook emotional manipulation study," *Research Ethics*, (12:1), pp. 14–28 (doi: 10.1177/1747016115599568).

Flores-Saviaga, C., Savage, S., and Taraborelli, D. 2016. "LeadWise: Using online bots to recruit and guide expert volunteers," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, (Vol. 26-Februar), pp. 257–260 (doi: 10.1145/2818052.2869106).

Fromm, J., Melzer, S., Ross, B., and Stieglitz, S. 2018. "Trump versus Clinton – Twitter Communication During the US Primaries," in *Network Intelligence Meets User Centered Social Media Networks*, R. Alhajj, H. U. Hoppe, T. Hecking, P. Bródka, and P. Kazienko (eds.), Cham: Springer, pp. 201–217 (doi: 10.1007/978-3-319-90312-5_14).

Fu, P. W., Wu, C. C., and Cho, Y. J. 2017. "What makes users share content on facebook? Compatibility among psychological incentive, social capital focus, and content type," *Computers in Human Behavior*, (67), pp. 23–32 (doi: 10.1016/j.chb.2016.10.010).

Ghani, N. A., Hamid, S., Targio Hashem, I. A., and Ahmed, E. 2018. "Social media big data analytics: A survey," *Computers in Human Behavior* (doi: 10.1016/j.chb.2018.08.039).

Giannoulakis, S., and Tsapatsoulis, N. 2016. "Evaluating the descriptive power of Instagram hashtags," *Journal of Innovation in Digital Ecosystems*, (3:2), pp. 114–129 (doi: 10.1016/j.jides.2016.10.001).

Goga, O., Venkatadri, G., and Gummadi, K. P. 2015. "The doppelgänger bot attack: Exploring identity impersonation in online social networks," in *Proceedings of the 2015 Internet Measurement Conference*, pp. 141–153 (doi: 10.1145/2815675.2815699).

Goh, K. Y., Heng, C. S., and Lin, Z. 2013. "Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content," *Information Systems Research*, (24:1), pp. 88–107 (doi: 10.1287/isre.1120.0469).

Golder, S. A., and Macy, M. W. 2011. "Diurnal and seasonal mood vary with work,

sleep, and daylength across diverse cultures," *Science*, (333:6051), pp. 1878–1881 (doi: 10.1126/science.1202775).

Gregor, S. 2006. "The nature of theory in Information Systems," *MIS Quarterly*, (30:3), pp. 611–642 (doi: 10.2307/25148742).

Gruhl, D., Liben-Nowell, D., Guha, R., and Tomkins, A. 2004. "Information diffusion through blogspace," in *Thirteenth International World Wide Web Conference Proceedings, WWW2004*, New York, NY, USA: ACM, pp. 491–501 (doi: 10.1145/988672.988739).

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., and Larivière, V. 2016. "Tweets as impact indicators: Examining the implications of automated 'bot' accounts on Twitter," *Journal of the Association for Information Science and Technology*, (67:1), pp. 232–238 (doi: 10.1002/asi.23456).

Howison, J., Wiggins, A., and Crowston, K. 2011. "Validity issues in the use of social network analysis with digital trace data," *Journal of the Association of Information Systems*, (12:12), pp. 767–797 (doi: doi.org/10.17705/1jais.00282).

Hudson, S., Roth, M. S., Madden, T. J., and Hudson, R. 2015. "The effects of social media on emotions, brand relationship quality, and word of mouth: An empirical study of music festival attendees," *Tourism Management*, (47), pp. 68–76 (doi: 10.1016/j.tourman.2014.09.001).

Jung, A.-K., Mirbabaie, M., Ross, B., Stieglitz, S., Neuberger, C., and Kapidzic, S. 2018. "Information Diffusion between Twitter and Online Media," in *Proceedings of the International Conference on Information Systems (ICIS)*.

Kalampokis, E., Tambouris, E., and Tarabanis, K. 2013. "Understanding the predictive power of social media," *Internet Research*, (23:5), pp. 544–559 (doi: 10.1108/IntR-06-2012-0114).

Kane, G. C., Alavi, M., Labianca, G., and Borgatti, S. P. 2014. "What's different about social media networks? A framework and research agenda," *MIS Quarterly*, (38:1), pp. 274–304.

Kaplan, A. M., and Haenlein, M. 2010. "Users of the world, unite! The challenges and opportunities of Social Media," *Business Horizons*, (53:1), pp. 59–68 (doi: 10.1016/j.bushor.2009.09.003).

Kaplan, A. M., and Haenlein, M. 2012. "Social media: Back to the roots and back to the future," *Journal of Systems and Information Technology*, (14:2), pp. 101–104 (doi: 10.1108/13287261211232126).

Kim, J., Bae, J., and Hastak, M. 2018. "Emergency information diffusion on online social media during storm Cindy in U.S.," *International Journal of Information Management*, (40), pp. 153–165 (doi: 10.1016/j.ijinfomgt.2018.02.003).

Kossinets, G., and Watts, D. J. 2006. "Empirical analysis of an evolving social network," *Science*, (311:5757), pp. 88–90 (doi: 10.1126/science.1116869).

Krämer, N. C., and Winter, S. 2008. “Impression Management 2.0,” *Journal of Media Psychology*, (20), pp. 106–116 (doi: 10.1027/1864-1105.20.3.106).

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Van Alstyne, M. 2009. “Life in the Network: The Coming Age of Computational Social Science,” *Science*, (323:5915), pp. 721–723 (doi: 10.1126/science.1167742).

Lee, D., Hosanagar, K., and Nair, H. S. 2018. “Advertising Content and Consumer Engagement on Social Media: Evidence from Facebook,” *Management Science*, (64:11), pp. 5105–5131 (doi: 10.1287/mnsc.2017.2902).

Lionberger, H. F., and Coughenour, C. M. 1957. “Social Structure and Diffusion of Farm Innovation,” Columbia, MO.

Lokot, T., and Diakopoulos, N. 2016. “News Bots,” *Digital Journalism*, (4:6), pp. 682–699 (doi: 10.1080/21670811.2015.1081822).

Mann, A. 2016. “Core Concept: Computational social science,” *Proceedings of the National Academy of Sciences*, (113:3), pp. 468–470 (doi: 10.1073/pnas.1524881113).

Marwick, A. E., and boyd, D. 2011. “I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience,” *New Media & Society*, (13:1), pp. 114–133 (doi: 10.1177/1461444810365313).

Matook, S., Brown, S. A., and Rolf, J. 2015. “Forming an intention to act on recommendations given via online social networks,” *European Journal of Information Systems*, (24:1), pp. 76–92 (doi: 10.1057/ejis.2013.28).

Matthes, J., and Hayes, A. F. 2014. “Methodological conundrums in spiral of silence research,” in *The Spiral of Silence: New Perspectives on Communication and Public Opinion* (doi: 10.4324/9780203125007).

McAfee, A., and Brynjolfsson, E. 2012. “Big data: The management revolution,” *Harvard Business Review*, (90:10), pp. 60–68.

Mejias, U. A., and Vokuev, N. E. 2017. “Disinformation and the media: the case of Russia and Ukraine,” *Media, Culture & Society*, (39:7), pp. 1027–1042 (doi: 10.1177/0163443716686672).

Meng, J., Peng, W., Tan, P.-N., Liu, W., Cheng, Y., and Bae, A. 2018. “Diffusion size and structural virality: The effects of message and network features on spreading health information on twitter,” *Computers in Human Behavior*, (89), pp. 111–120 (doi: 10.1016/j.chb.2018.07.039).

Morone, F., and Makse, H. A. 2015. “Influence maximization in complex networks through optimal percolation,” *Nature*, (524:7563), pp. 65–68 (doi: 10.1038/nature14604).

Morstatter, F., Pfeffer, J., Liu, H., and Carley, K. M. 2013. “Is the sample good enough?

Comparing data from twitter's streaming API with Twitter's firehose," in *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*, pp. 400–408.

Mueller, R. S. 2019. "Report on the Investigation Into Russian Interference In The 2016 Presidential Election," Washington, DC.

OED Online. 2019a. "diffusion, n.," Oxford University Press (available at https://www.oed.com/view/Entry/52543#eid6760342; retrieved September 2, 2019).

OED Online. 2019b. "analytics, n.," Oxford University Press (available at https://www.oed.com/view/Entry/273413#eid189501945; retrieved September 2, 2019).

Pfeffer, J., Mayer, K., and Morstatter, F. 2018. "Tampering with Twitter's Sample API," *EPJ Data Science*, (7:1), p. 50 (doi: 10.1140/epjds/s13688-018-0178-0).

Rapoport, A. 1953. "Spread of information through a population with socio-structural bias: I. Assumption of transitivity," *The Bulletin of Mathematical Biophysics*, (15:4), pp. 523–533 (doi: 10.1007/BF02476440).

Ratkiewicz, J., Conover, M. D., Meiss, M., Gonçalves, B., Flammini, A., and Menczer, F. 2011. "Detecting and Tracking Political Abuse in Social Media," in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, Spain, pp. 297–304.

Reuter, C., Ludwig, T., Kaufhold, M.-A., and Spielhofer, T. 2016. "Emergency services' attitudes towards social media: A quantitative and qualitative survey across Europe," *International Journal of Human-Computer Studies*, (95), pp. 96–111 (doi: 10.1016/j.ijhcs.2016.03.005).

Rogers, E. M. 2003. *Diffusion of Innovations*, (5th ed.), New York: Free Press.

Romero, D. M., Meeder, B., and Kleinberg, J. 2011. "Differences in the mechanics of information diffusion across topics," in *Proceedings of the 20th international conference on World wide web - WWW '11*, New York, New York, USA: ACM Press, p. 695 (doi: 10.1145/1963405.1963503).

Ross, B., Brachten, F., Stieglitz, S., Wikström, P., Moon, B., Münch, F., and Bruns, A. 2018. "Social Bots in a Commercial Context – A Case Study on Soundcloud," in *Proceedings of the European Conference on Information Systems (ECIS)*.

Ross, B., Jung, A.-K., Heisel, J., and Stieglitz, S. 2018. "Fake News on Social Media: The (In)Effectiveness of Warning Messages," in *Proceedings of the International Conference on Information Systems (ICIS)*.

Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., and Stieglitz, S. 2019. "Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks," *European Journal of Information Systems* (doi: 10.1080/0960085X.2018.1560920).

Ross, B., Potthoff, T., Majchrzak, T. A., Chakraborty, N. R., Lazreg, B., and Stieglitz, S. 2018. "The Diffusion of Crisis-Related Communication on Social Media: An Empirical Analysis of Facebook Reactions," in *Proceedings of the 51st Hawaii International Conference on Information Systems*, pp. 2525–2534 (doi: 10125/50207).

Ross, B., Rist, M., Carbonell, G., Cabrera, B., Kurowsky, N., and Wojatzki, M. 2016. "Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis," in *Bochumer Linguistische Arbeitsberichte 17 (BLA 17), Proceedings of NLP4CMC III: 3rd Workshop on Natural Language Processing for Computer-Mediated Communication*, M. Beißwenger, M. Wojatzki, and T. Zesch (eds.), Bochum, Germany, pp. 6–9 (doi: 10.17185/DUEPUBLICO/42132).

Ruths, D., and Pfeffer, J. 2014. "Social media for large studies of behavior," *Science*, (346:6213), pp. 1063–1064 (doi: 10.1126/science.346.6213.1063).

Saha, B., and Srivastava, D. 2014. "Data quality: The other face of Big Data," in *Proceedings of the 2014 IEEE 30th International Conference on Data Engineering*, IEEE, March, pp. 1294–1297 (doi: 10.1109/ICDE.2014.6816764).

Salto Martínez, R., and Jacques García, F. A. 2012. "Development and implementation of a chat bot in a social network," in *Proceedings of the 9th International Conference on Information Technology, ITNG 2012*, pp. 751–755 (doi: 10.1109/ITNG.2012.147).

Scheufele, D. 2007. "Spiral of Silence Theory," in *The SAGE Handbook of Public Opinion Research*, p. 640 (doi: 10.1057/bp.2010.5).

Sebastian, S., Ayyappan, S., and Vinod, P. 2014. "Framework for design of Graybot in social network," in *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*, pp. 2331–2336 (doi: 10.1109/ICACCI.2014.6968575).

Sequiera, R., and Lin, J. 2017. "Finally, a Downloadable Test Collection of Tweets," in *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '17*, New York, New York, USA: ACM Press, pp. 1225–1228 (doi: 10.1145/3077136.3080667).

Shen, F., and Liang, H. 2015. "Cultural Difference, Social Values, or Political Systems? Predicting Willingness to Engage in Online Political Discussion in 75 Societies," *International Journal of Public Opinion Research*, (27:1), pp. 111–124 (doi: 10.1093/ijpor/edu012).

SimilarWeb. 2019. "youtube.com July 2019 Overview," (available at https://www.similarweb.com/website/youtube.com; retrieved September 2, 2019).

Sohn, D., and Geidner, N. 2016. "Collective Dynamics of the Spiral of Silence: The Role of Ego-Network Size," *International Journal of Public Opinion Research*, (28:1), pp. 25–45 (doi: 10.1093/ijpor/edv005).

Statista. 2018. "Nutzung von Social Media als Informationsquelle nach Plattform in Deutschland 2018," (available at

https://de.statista.com/statistik/daten/studie/878465/umfrage/nutzung-von-social-media-als-informationsquelle-nach-plattform-in-deutschland/; retrieved September 2, 2019).

Stieglitz, S., Brachten, F., Ross, B., and Jung, A.-K. 2017. "Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts," in *Proceedings of the Australasian Conference on Information Systems (ACIS)*.

Stieglitz, S., and Dang-Xuan, L. 2013. "Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior," *Journal of Management Information Systems*, (29:4), pp. 217–248 (doi: 10.2753/MIS0742-1222290408).

Stieglitz, S., Dang-Xuan, L., Bruns, A., and Neuberger, C. 2014. "Social Media Analytics," *Business & Information Systems Engineering*, (6:2), pp. 89–96 (doi: 10.1007/s12599-014-0315-7).

Stieglitz, S., Meske, C., Ross, B., and Mirbabaie, M. 2018. "Going Back in Time to Predict the Future - The Complex Role of the Data Collection Period in Social Media Analytics," *Information Systems Frontiers* (doi: 10.1007/s10796-018-9867-2).

Stieglitz, S., Mirbabaie, M., Ross, B., and Neuberger, C. 2018. "Social media analytics – Challenges in topic discovery, data collection, and data preparation," *International Journal of Information Management*, (39), pp. 156–168 (doi: 10.1016/j.ijinfomgt.2017.12.002).

Stocking, G., and Sumida, N. 2018. "Social Media Bots Draw Public's Attention and Concern," Pew Research Center.

Subrahmanian, V. S., Azaria, A., Durst, S., Kagan, V., Galstyan, A., Lerman, K., Zhu, L., Ferrara, E., Flammini, A., and Menczer, F. 2016. "The DARPA Twitter Bot Challenge," *Computer*, (49:6), pp. 38–46 (doi: 10.1109/MC.2016.183).

Suseno, Y., Laurell, C., and Sick, N. 2018. "Assessing value creation in digital innovation ecosystems: A Social Media Analytics approach," *The Journal of Strategic Information Systems*, (27:4), pp. 335–349 (doi: 10.1016/j.jsis.2018.09.004).

Suwajanakorn, S., Seitz, S. M., and Kemelmacher-Shlizerman, I. 2017. "Synthesizing Obama," *ACM Transactions on Graphics*, (36:4), pp. 1–13 (doi: 10.1145/3072959.3073640).

Thelwall, M. 2018. "Social media analytics for YouTube comments: potential and limitations," *International Journal of Social Research Methodology*, (21:3), pp. 303–316 (doi: 10.1080/13645579.2017.1381821).

Tsvetkova, M., García-Gavilanes, R., Floridi, L., and Yasseri, T. 2017. "Even good bots fight: The case of Wikipedia," *PLOS ONE*, (S. Gómez, ed.), (12:2), p. e0171774 (doi: 10.1371/journal.pone.0171774).

Twitter. 2018. "Platform manipulation," (available at https://transparency.twitter.com/en/platform-manipulation.html; retrieved

September 2, 2019).

Twitter. 2019. "Developer Terms: More about restricted uses of the Twitter APIs," (available at https://developer.twitter.com/en/developer-terms/more-on-restricted-use-cases.html; retrieved September 9, 2019).

Veale, T., Valitutti, A., and Li, G. 2015. "Twitter: The best of bot worlds for automated wit," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, N. Streitz and P. Markopoulos (eds.), (Vol. 9189), Cham: Springer, pp. 689–699 (doi: 10.1007/978-3-319-20804-6_63).

Vengattil, M., and Dave, P. 2019. "Facebook 'labels' posts by hand, posing privacy questions," *Reuters* (available at https://www.reuters.com/article/us-facebook-ai/facebook-labels-posts-by-hand-posing-privacy-questions-idUSKCN1SC01T; retrieved September 2, 2019).

Wang, A. H. 2010. "Detecting spam bots in online social networking sites: A machine learning approach," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, S. Foresti and S. Jajodia (eds.), (Vol. 6166), Berlin, Heidelberg: Springer, pp. 335–342 (doi: 10.1007/978-3-642-13739-6_25).

Wattal, S., Schuff, D., Mandviwalla, M., and Williams, C. B. 2010. "Web 2.0 and politics: The 2008 U.S. presidential election and an e-politics research agenda," *MIS Quarterly*, (34:4), pp. 669–688.

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly*, (26:2), pp. xiii–xxiii.

Weller, K., and Kinder-Kurlanda, K. E. 2015. "Uncovering the challenges in collection, sharing and documentation: The hidden data of social media research?," in *Standards and Practices in Large-Scale Social Media Research: Papers from the 2015 ICWSM Workshop*, pp. 28–37.

Weller, K., and Kinder-Kurlanda, K. E. 2016. "A manifesto for data sharing in social media research," in *WebSci 2016 - Proceedings of the 2016 ACM Web Science Conference* (doi: 10.1145/2908131.2908172).

Wilensky, U., and Rand, W. 2015. *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*, Cambridge, MA: The MIT Press.

Wilkie, A., Michael, M., and Plummer-Fernandez, M. 2015. "Speculative Method and Twitter: Bots, Energy and Three Conceptual Characters," *The Sociological Review*, (63:1), pp. 79–101 (doi: 10.1111/1467-954X.12168).

Xiao, Y., Cauberghe, V., and Hudders, L. 2018. "Humour as a double-edged sword in response to crises versus rumours: The effectiveness of humorously framed crisis response messages on social media," *Journal of Contingencies and Crisis Management*, (26:2), pp. 247–260 (doi: 10.1111/1468-5973.12188).

Yang, J., and Leskovec, J. 2010. "Modeling Information Diffusion in Implicit Networks," in *Proceedings of the 2010 IEEE International Conference on Data Mining*, IEEE, December, pp. 599–608 (doi: 10.1109/ICDM.2010.22).

Yin, P., Ram, N., Lee, W.-C., Tucker, C., Khandelwal, S., and Salathé, M. 2014. "Two Sides of a Coin: Separating Personal Communication and Public Dissemination Accounts in Twitter," in *Advances in Knowledge Discovery and Data Mining: 18th Pacific-Asia Conference, PAKDD 2014*, V. S. Tseng, T. B. Ho, Z.-H. Zhou, A. L. P. Chen, and H.-Y. Kao (eds.), Cham: Springer International Publishing, pp. 163–175 (doi: 10.1007/978-3-319-06608-0_14).

Yun, G. W., and Park, S.-Y. 2011. "Selective Posting: Willingness to post a message online," *Journal of Computer-Mediated Communication*, (16:2), pp. 201–227 (doi: 10.1111/j.1083-6101.2010.01533.x).

Zafarani, R., Abbasi, M. A., and Liu, H. 2014. *Social Media Mining*, (Draft Vers.), Cambridge: Cambridge University Press (doi: 10.1017/CBO9781139088510).

Ziegele, M., Weber, M., Quiring, O., and Breiner, T. 2018. "The dynamics of online news discussions: effects of news articles and reader comments on users' involvement, willingness to participate, and the civility of their contributions," *Information, Communication & Society*, (21:10), pp. 1419–1435 (doi: 10.1080/1369118X.2017.1324505).

Zubiaga, A. 2018. "A longitudinal assessment of the persistence of twitter datasets," *Journal of the Association for Information Science and Technology*, (69:8), pp. 974–984 (doi: 10.1002/asi.24026).