Introduction

Consider the following tweets posted on 9 November 2016, one day after Donald Trump won the US presidential election:

“President Trump wants to know if you have any last words Mr Soros?”
#RevengeWillBeSweet #WhiteGenocide #RapeJihad #RWDS #Trump #Trump16
[+ image of a Nazi shooting a Jewish person]
#Trump卐The end of #WhiteGenocide in America. #Nazi #SiegHeil

We won! This is a BIG win for the white race as a whole. And we won’t stop. We will take back what is ours! #MAGA #WhitePride #14words
Anti-Whites are shitting themselves right now. They do not like whites taking back their country!! #WhitePride #Trump2016

Gonna go kill some niggers, mexicans, and muslims tommorow trump will just pardon me lol cant wait wooo #MAGA

The examples indicate the prevalence of fascist, racist, nationalist ideology in public discussions of Trump’s victory. Given that the world economic crisis of 2008 has turned into a political crisis that has brought about the intensification of nationalism, xenophobia, racism and fascism, it is an important task for critical research to study how and why these phenomena exist. Social media is a kind of mirror of what is happening in society. Studying social media content is therefore a good way of studying society. But whenever we conduct social research, ethical issues regarding anonymity, informed consent, and privacy may arise.
Research ethics is a key aspect of social science. Not only is there a general etiquette of publishing, but also ethical questions that arise in the collection of data. The emergence of what some call ‘social media’ and ‘big data’ has complicated research ethics. In this contribution, I reflect on research ethics in respect to the study of online ideologies, especially in the context of ‘negative’ social movements and forms of online expression that are fascist, racist, nationalist, anti-socialist, and anti-Semitic in character.

Doing online research complicates research ethics. So when for example conducting a Critical Discourse Analysis (CDA) of White supremacist content, the question arises whether you have to obtain informed consent for including and analyzing a fascist tweet. Writing an email asking ‘Dear Mr. Neo-Nazi, can you please give me your informed consent so that I can quote your fascist tweet?’ may not just result in rejection, it could also draw the attention of fascists towards you as a critical researcher and put you in danger.

This chapter deals with the question of how to deal with research ethics in qualitative online research. First, the chapter discusses the limits of established research ethics guidelines (Section 2). Second, it outlines foundations of critical-realist internet research ethics (Section 3). Third, it provides some examples of how to use such a framework (Section 4). Finally, some conclusions are drawn (Section 5).

2 Established research ethics guidelines

An obvious approach of how to deal with questions of research ethics in qualitative online research is to look at established research ethics guidelines provided by academic associations.

The Association of Internet Researchers’ ethical recommendations (2012: 6–7) contains a list of questions that one can ask when conducting online research and points out ethical problems that may arise:

People may operate in public spaces but maintain strong perceptions or expectations of privacy. Or, they may acknowledge that the substance of their communication is public, but that the specific context in which it appears implies restrictions on how that information is – or ought to be – used by other parties. Data aggregators or search tools make information accessible to a wider public than what might have been originally intended. … Social, academic, or regulatory delineations of public and private as a clearly recognizable binary no longer holds in everyday practice. … Yet there is considerable evidence that even ‘anonymised’ datasets that contain enough personal information can result in individuals being identifiable. Scholars and technologists continue to wrestle with how to adequately protect individuals when analysing such datasets. … These are important considerations because they link to the fundamental ethical principle of minimizing harm.

We can find two important points here:

1 In the online world, the boundary between the private and the public realm is messy. The question therefore arises if all Twitter content can be considered public content, as in a newspaper, or if there may also be content that is more private and intended for a limited audience.

2 Anonymization becomes difficult online because data is stored on servers and is searchable. In the case of Twitter, search engines such as backtweets (http://backtweets.com) allows us to search for archived tweets. Anonymity of cited content therefore becomes difficult to ascertain.
But does this mean that any qualitative analysis and quoting from Twitter violates research ethics? Or does one have to attain informed consent for each tweet one uses from others? The AoIR-document points out the complexity of online research ethics, but it does not provide any guidelines on how to actually deal with such questions.

The British Sociological Association (2002: §41) recommends in its *Statement of Ethical Practice* that researchers studying the internet should keep themselves updated on relevant issues:

Members should take special care when carrying out research via the Internet. Ethical standards for internet research are not well developed as yet. Eliciting informed consent, negotiating access agreements, assessing the boundaries between the public and the private, and ensuring the security of data transmissions are all problematic in Internet research. Members who carry out research online should ensure that they are familiar with ongoing debates on the ethics of Internet research, and might wish to consider erring on the side of caution in making judgements affecting the well-being of online research participants.

This short paragraph certainly does not help an internet researcher in any particular situation in which s/he deals with ethical issues. The International Sociological Association’s 2001 *Code of Ethics* argues in respect to informed consent:

The security, anonymity and privacy of research subjects and informants should be respected rigourously, in both quantitative and qualitative research. The sources of personal information obtained by researchers should be kept confidential, unless the informants have asked or agreed to be cited. Should informants be easily identifiable, researchers should remind them explicitly of the consequences that may follow from the publication of the research data and outcomes. … The consent of research subjects and informants should be obtained in advance.

The ISA code does not mention the specificities of online research. Anonymity often does not exist online. Obtaining informed content when working with a large online dataset is for the most part practically impossible due to time restrictions. In the online world, the private and the public spheres do not uphold clear boundaries.

The American Sociological Association’s (1999) *Code of Ethics* says the following about anonymity and informed consent:

11.06 Anonymity of Sources (a) Sociologists do not disclose in their writings, lectures, or other public media confidential, personally identifiable information concerning their research participants, students, individual or organizational clients, or other recipients of their service which is obtained during the course of their work, unless consent from individuals or their legal representatives has been obtained. (b) When confidential information is used in scientific and professional presentations, sociologists disguise the identity of research participants, students, individual or organizational clients, or other recipients of their service. … 12.01 Scope of Informed Consent (a) Sociologists conducting research obtain consent from research participants or their legally authorized representatives (1) when data are collected from research participants through any form of communication, interaction, or intervention; or (2) when behavior of research participants occurs in a private context where an individual can reasonably expect that no observation or reporting is taking place. …
public places or use publicly-available information about individuals (e.g., naturalistic observations in public places, analysis of public records, or archival research) without obtaining consent.

The ASA code does not specifically mention online research. It does not recognize that in online research it is not straightforward to keep cited content anonymous. However, it does make a good point in remarking that there is a difference in obtaining informed consent in respect to the question of whether communication, interaction and behaviour take place in a private context or in a public place. In relation to social media, this means that one needs to ask which communications are private and which ones are public.

Overall, the discussion shows that established ethics guidelines do not direct much attention to the particularities of online research ethics.

3 Towards critical-realist internet research ethics

There are two extremes in internet research ethics. The one extreme argues that one must obtain informed consent for every piece of data one gathers online. The other argues that what is online is out there and can and should be analyzed without regard to ethical considerations.

Zimmer (2010) discusses the question of whether or not it is ethical to harvest Twitter data without informed consent:

Yes, setting one’s Twitter stream to public does mean that anyone can search for you, follow you, and view your activity. However, there is a reasonable expectation that one’s tweet stream will be “practically obscure” within the thousands (if not millions) of tweets similarly publicly viewable. Yes, the subject has consented to making her tweets visible to those who take the time and energy to seek her out, those who have a genuine interest to connect and view her activity through this social network. But she did not automatically consent, I argue, to having her tweet stream systematically followed, harvested, archived, and mined by researchers (no matter the positive intent of such research). That is not what is expected when making a Twitter account public, and it is my opinion that researchers should seek consent prior to capturing and using this data.

Some of the people commenting on this blog post heavily disagreed with Zimmer’s (2010) perspective:

It’s like a blog. (Originally, Twitter was called “the microblogging service”.) You can quote and attribute from blogs, but you can’t pretend it’s your work … As for someone deciding to analyse me from my tweets and publish the results – well, not much i can do about the analysis

The web is not an environment that supports a reasonable expectation of privacy in public. Unmistakeably not. Nor does twitter as a subculture gesture toward such an expectation. Once tweeted, a birdsong is gone forever. No deleting or taking back what’s been broadcast to the world. If someone seeks privacy, they should seek another method of communication. TWITTER IS PUBLIC – NO QUESTION ABOUT IT. Tweets (from the public stream) are like to be treated like blogs (microblogs) and webpages – PUBLIC. No consent required for analyzing them, unless of course they are DMs (which are like emails – confidential) or sent to your “followers only” …. You tweet because you want to get your message out, and not only to our friends (ever heard of retweets?). This is
‘Dear Mr. Neo-Nazi’

VERY different from discussion boards, chat rooms, or even Facebook. … I simply dispute that ANYBODY who tweets (regardless of whether he has read the privacy policy or not) does so under the expectation of privacy or having a “limited” audience (if they want to do that, there is a privacy setting for that). Anybody who tweets sees on a daily basis that others are retweeting their tweets or quoting from their tweets also appear in search engines and on the twitter homepage itself.

The discussion shows that there is a conflict between research ethics fundamentalists and big data positivist. Research ethics fundamentalists tend to say:

You have to attain informed consent for every piece of social media data you gather because we cannot assume automatic consent. Users tend not to read a platform’s privacy policies – they may assume that some of their data is private, and they may not agree to their data being used in research. Even if you anonymize the users you quote, many might still be identified in the networked online environment.

There are limits of informed consent. It can censor critical research and cause harm for a researcher conducting critical online research if s/he contacts a user, asking: ‘Dear Mr. Misogynist/Nazi/Right-Wing Extremist etc.! I am a social researcher gathering data from Twitter. Can you please give me your informed consent for quoting your violent threat against X?’ The researcher may be next in line for being harassed or threatened.

A solution would be to only use aggregated data. But such an approach is biased towards quantitative methods and computational social science. Critical discourse analysis and critical interpretative research thereby become impossible.

Big data positivists tend to say: ‘Most social media data is public data. It is like data in a newspaper. I can therefore gather big data without limits. Those talking about privacy want to limit the progress of social science.’ This position disregards any engagement with ethics and is biased towards quantification (meaning big data positivism, digital positivism). Zimmer and Proferes (2014) conducted a meta-study of 382 works focusing on Twitter research. Only 4 per cent of the works discussed any ethical aspects. While privacy fetishism is one extreme, another extreme is the complete ignorance of research ethics, a kind of ‘anything-goes’ attitude towards the question what researchers are allowed to do.

Privacy fetishism holds the danger of censoring and disabling critical research. It can endanger the critical researcher and result in violence directed against him/her by fascists, racists, nationalists, etc. Downright ignoring research ethics is often associated with a positivist approach to online research that focuses on the digital Lasswell formula: who says what online, who do they say it to, how many likes, followers, re-tweets, comments, and friends do they have? The problem of this formula is that it leaves out questions such as the following: how are meanings expressed? What power structures condition the communication? What are the communicator’s motivations, interests and experiences? What contradictions does the communication involve?

We need critical-realist digital media research guidelines that go beyond research ethics fundamentalism and big data positivism. The approach needs to be realist in the sense that it avoids the two extremes of fundamentalism and positivism. The approach has to both engage with research ethics and enable the conduction of actual online research. The approach is critical in that it takes care to formulate guidelines in such a way as to enable and foster critical online research. By critical online research, we can understand any study that investigates digital media in the context of power structures (Fuchs 2017b).
In February 2016, I was part of a group of 16 scholars that met for a workshop funded by the Economic and Social Research Council (ESRC) at the University of Aberdeen. The task was that we create social media research ethics guidelines. The group consisted of a diverse range of scholars taking different perspectives on research ethics. Overall, the group managed to formulate some guidelines for a critical-realism research ethics framework (Townsend et al. 2016).

As one of the starting points for a realist perspective, we found a recommendation in the British Psychological Society’s 2009 Code of Ethics and Conduct helpful: ‘Unless informed consent has been obtained, restrict research based upon observations of public behaviour to those situations in which persons being studied would reasonably expect to be observed by strangers’ (BPS 2009: 13). The British Psychological Society’s 2013 Ethics Guidelines for Internet-Mediated Research applies this principle to online research:

Where it is reasonable to argue that there is likely no perception and/or expectation of privacy (or where scientific/social value and/or research validity considerations are deemed to justify undisclosed observation), use of research data without gaining valid consent may be justifiable.

BPS 2013: 7

Based on this insight, we formulated the following general guideline in the framework Social Media Research: A Guide to Ethics:

The question as to whether to consider social media data as private or public comes down, to some extent, to whether or not the social media user can reasonably expect to be observed by strangers (British Psychological Society 2013; Fuchs forthcoming). Things to consider here are: is the data you wish to access on an open forum or platform (such as on Twitter), or is it located within a closed or private group (e.g. within Facebook) or a closed discussion forum? Is the group or forum password protected? Would platform users expect other visitors to have similar interests or issues to themselves? Does the group have a gatekeeper (or admin) that you could turn to for approval and advice? How have users set up their security settings? Data accessed from open and public online locations such as Twitter present less ethical issues than data which are found in closed or private online spaces. Similarly, data posted by public figures such as politicians, musicians and sportspeople on their public social media pages is less likely to be problematic because this data is intended to reach as wide an audience as possible. If the data you wish to access is held within a group for which you would need to gain membership approval, or if the group is password protected, there are more ethical issues to take into consideration.

Townsend et al. 2016: 10

Practically speaking, this means that analyzing private messages and conversations in a closed group of recipients on Twitter requires informed consent. Most tweets, especially those using hashtags, aim at public visibility and therefore do not require informed consent in online research. How should one deal with Twitter users’ identifiability? As good practice, one should not mention usernames, except for well-known public persons and institutions. One can instead use a pseudonym. It may still be possible to identify who posted a particular text that the researcher uses, but as this requires additional effort on the part of the person who wants to find out, the researcher does not directly identify the user.

Here is a specific example of how to apply these guidelines:
‘Dear Mr. Neo-Nazi’

Context: A researcher conducts a critical discourse analysis of a dataset of tweets using the hashtags #DonaldTrump; #TrumpTrain; #VoteTrump2016; #AlwaysTrump; #MakeAmericaGreatAgain or #Trump2016. These are analysed in order to find out how Trump supporters argue for their candidate on Twitter. Concerns: Can we consider this data public? Are there any issues of sensitivity or risk of harm? Do we need to seek informed consent before quoting these tweets directly?

Solution: Trump supporters use these hashtags in order to reach a broad public and convince other people to vote for Trump. It is therefore reasonable to assume that such tweets have public character: the authors expect and want to be observed by strangers in order to make a political point that they want others to read. The researcher can therefore directly quote such tweets without having to obtain informed consent. It is, however, good practice to delete the user IDs of everyday users, who are not themselves public figures.

Townsend et al. 2016: 15

4 Example cases of critical-realist internet research ethics

I want to outline an example of how I have dealt with research ethics in qualitative online studies that used critical discourse analysis. I will here deliberately abstract from the actual research results and merely focus on the ethical questions.

The study ‘Fascism 2.0: Twitter Users’ Social Media Memories of Hitler on his 127th Birthday’ (Fuchs 2017a) analyzed how Twitter users communicated about Hitler on his 127th birthday. It utilized empirical ideology critique as its method. I used the tool Texifter to obtain all tweets from 20 April 2016 that mentioned any of the following hashtags: #hitler OR #adolfhitler OR #hitlerday OR #1488 OR #AdolfHitlerDay OR #HeilHitler OR #SiegeHeil OR #HappyBirthdayAdolf OR #HitlerNation OR #HappyBirthdayHitler OR #HitlersBirthday OR #MakeGermanyGreatAgain OR #WeMissYouHitler. The search resulted in 4,193 tweets that were automatically imported into Discovertext, from where I exported them along with meta-data into a csv file. Using such hashtags on Hitler’s birthday clearly aims at creating public attention. We can therefore say that the use of these hashtags in the context of Hitler’s birthday constitutes a public space. Informed consent for analyzing such postings is therefore not needed.

The study ‘Red Scare 2.0: User-Generated Ideology in the Age of Jeremy Corbyn and Social Media’ (Fuchs 2016b) asked: how has Jeremy Corbyn during the Labour Leadership Election been framed in an ideological manner in discourses on Twitter and how have such ideological discourses been challenged? The study stands in the context of the negative framing of Corbyn during and following his run for the Labour Party leadership. With the help of Discovertext, I gathered 32,298 tweets based on the following search query: Corbyn AND anti-Semite OR anti-Semitic OR chaos OR clown OR commy OR communism OR communist OR loony OR Marx OR Marxist OR pinko OR red OR reds OR socialism OR socialist OR Stalin OR Stalinist OR terrorist OR violent OR violence. The data gathering was active for 23 days, from 22 August 2015 (23:25 BST) until 13 September 2015 (12:35 BST). Corbyn was announced as the winner on 12 September 2015 (11:45 BST). It is reasonable to assume that users who tweet about Jeremy Corbyn during times when he is subject to increased public attention are directing their communication at the public. Also in this case, informed consent is therefore not required.

The study entailed a focus on the ten most active and most mentioned pro- and anti-Corbyn users (see Table 40.1). In the analysis, I anonymized individual users who are not well-known public figures and did not anonymize public figures (such as Glenn Greenwald, Rupert Murdoch, David Schneider) and institutions (such as the Daily Telegraph, Russia Today, The Independent).
The most active users were Twitter bots (redscarebot, mywoodthorpe). A bot based on an algorithm conducts certain online behaviour. Given that technologies do not maintain ethics, they likewise do not have expectations about privacy. They therefore do not need to be anonymized.

The study ‘Racism, Nationalism and Right-Wing Extremism Online: The Austrian Presidential Election 2016 on Facebook’ (Fuchs 2016a) stands in the context of the Austrian presidential election 2016 that saw a run-off between the Green party candidate Alexander Van der Bellen and the Freedom Party of Austria’s (FPÖ) far-right candidate Norbert Hofer. The paper asks: how did voters of Hofer express their support on Facebook? The FPÖ is the prototype of a European far-right party that bases its ideology on nationalism and xenophobia. Under the leadership of Jörg Haider (1986–2000), it was expanding and growing in popularity. Its current leader is Heinz Christian Strache.

I used Netvizz in order to collect comments on postings related to Hofer’s presidential candidacy. I accessed Norbert Hofer and Heinz Christian Strache’s Facebook pages on 30 May 2016, and used Netvizz for extracting comments to postings made between 25 and 30 May. Given that the collected comments were posted in the days after the presidential election’s second round, it is likely that the dataset contains data that refers to the political differences between Hofer and Van der Bellen. I selected postings by Hofer and Strache that were particularly polarizing. This selection resulted in a total of 15 postings: ten by Strache, five by Hofer. There were a total of 6,755 comments posted as responses to these 15 Facebook postings. So the analyzed dataset consisted of 6,755 items.

The Facebook pages of Norbert Hofer and Heinz Christian Strache are public pages. All postings and comments on these pages are visible to everyone visiting them, not just to those who ‘like’ them. One does not have to have a Facebook profile to access the two pages, as they can also be viewed without logging into Facebook. All postings and comments are thus visible in public. Furthermore, politicians are public figures. Citizens expect them to be present in the public. This includes that they post in public on social media and offer possibilities for public communication on their profiles. Given the public character of Strache and Hofer’s Facebook pages, it is reasonable to assume that someone posting a comment on such a page can expect to be observed by strangers. In such a case, a researcher does not have to obtain informed consent for analyzing and quoting comments. Given that the users are not public figures themselves, but only make public comments when posting on a politician’s public Facebook page, I do not mention the usernames in the analysis. Netvizz does not save the usernames and so the collected dataset does not contain any identifiers.
5 Conclusion

Objectively speaking, the far right is fairly effective when it comes to utilizing social media for political communication. Yet if one looks at the body of works published in social movement media studies, one gets the impression that political communication in the internet age is by far dominated by politically progressive, left-wing, social movements. There are comparatively few studies that focus on the internet and far-right politics (Caiani & Kröll 2015). The far right’s use of the internet has hardly been studied and is a blind spot in social movement media studies. W. Lance Bennett and Alexandra Segerberg’s book *The Logic of Connective Action* (2013) mentions Occupy 70 times, but the Golden Dawn, Jobbik, the National Front, UKIP, Svoboda, Farage, or Le Pen not a single time. The *Encyclopedia of Social Movement Media* (Downing 2011) presents 600 pages of analyses of ‘alternative media, citizens’ media, community media, counterinformation media, grassroots media, independent media, nano-media, participatory media, social movement media, and underground media’ (Downing 2011: xxv). The focus here is on all sorts of progressive and left-wing media, from the likes of the Adbusters Media Foundation to Zapatista media. The editor John Downing (2011: xxvi) admits that ‘much less examination of media of extreme right movements occurs in this volume than there might be’, but he does not explain why this is the case, why it is problematic, and how it could be changed.

Most social movement researchers like to do feel-good research. They study progressive left-wing movements that they like and are sympathetic towards, consider such studies as a form of solidarity, and tend to simply celebrate how these groups organize and communicate. Such studies make the researchers feel good and politically engaged. But celebratory studies of these movements hardly help us to understand the difficult contradictions that left-wing activism faces in the capitalist world. They neglect analyzing right-wing movements and groups that pose a threat to democracy. And thus this is the blind spot of social movement media studies.

One might now be tempted to argue that far-right groups are not part of social movement studies because they tend to be hierarchic, have a populist leader, and aim at a society that is governed from the top in an authoritarian or even fascist manner. However, such a definitional exclusion overlooks that also left-wing progressive movements often develop certain hierarchies and forms of leadership. Left-wing movements too attempt to define the social as a progressive political concept by arguing that the far right has anti-social political goals. The ‘social’ in social movements means nothing more than the circumstance that social movements are groups that act collectively in order to change society and move it in a certain direction. It tells us nothing about these groups’ political content. The point is that in a contradictory world, social movements are contradictory. They contest how society is developing. Two options that are today possible are the democratic socialist option of participatory democracy and the authoritarian option of fascist barbarism. Social movement studies should focus on studying the diverse range of political movements.

Studying online politics poses ethical challenges in respect to privacy/the public, anonymity and informed consent. Conventional research ethics guidelines often ignore qualitative online research or have little to say on the topic. Conducting studies of online nationalism, racism, xenophobia and fascism poses additional challenges because these phenomena are inherently violent. Debates on internet research ethics face two extremes. On the one side, research ethics fundamentalism obstructs qualitative online research. On the other, big data positivism lacks a critical focus on qualitative dimensions of analysis. The alternative is a critical-realist online research ethics that informs critical studies of digital media.
References


