The Investigator's Dilemma - A Review of Social Media Analytics Research Ethics in Information Systems

Julian Marx
University of Duisburg-Essen, Germany

Milad Mirbabaie
Paderborn University, Germany
milad.mirbabaie@uni-paderborn.de

Abstract

Social media have become not only integral parts of our private and professional lives, but also an indispensable source of data for empirical research across a variety of academic disciplines. Applying a Social Media Analytics (SMA) methodology, however, imposes heavy ethical challenges on researchers. Scholars in the Information Systems (IS) discipline must deal with a patchwork of ethical frameworks, regulations, and (missing) institutional support. To initiate a debate on how to develop a common understanding of SMA research ethics, this paper compiles a scoping review of extant literature and suggests a research agenda for IS scholarship on this matter. The review yields a total of eight fundamental principles of ethical SMA research, which provide a starting point to guiding individual researchers towards more ethical conduct. At the same time, this work unearths a multitude of intricate dilemmas that are currently unresolved. The findings of this review will encourage IS scholarship to find its own voice in the debate about social media research ethics.

Keywords: Social Media Analytics, Ethics, Information Systems.

1 Introduction

The impact of social media use on how communication is shaped within societies is continuously reinforced. Nowadays, social media are used in various domains, for example in business (Beier & Wagner, 2016; Richter et al., 2011), politics (Bruns & Burgess, 2015; S. Stieglitz & Dang-Xuan, 2013), or crisis management (Mirbabaie et al., 2020; Reuter et al., 2018). Hence, Information Systems (IS) research continues to closely investigate social media as both a phenomenon and data source.

Communication on social media is often complex and voluminous, creating high quantities of data. At this juncture, the Social Media Analytics (SMA) approach combines a set of data-driven analysis techniques, including the collection, analysis, and visualisation of social media data (Stieglitz et al., 2014). Particularly in IS research, the SMA approach has helped scholars to navigate the process of social data analysis, aiming at the improvement of making sense of the data (Zeng et al., 2010). This data handling, however, results in unprecedented challenges for researchers. Stieglitz et al. (2018) elaborated that SMA researchers have to deal with

1) a high heterogeneity of information,

2) data spread across different platforms,

3) many actors who are involved, and

4) a highly dynamic information diffusion.

These challenges exemplify that handling social media data constitutes a unique way of conducting research. Social media data may include personal and identifying information such as a research subject's name, date of birth, or occupation. Oftentimes, a researcher can hardly estimate to what extent this personal information is of a sensitive nature (Lomborg & Bechmann, 2014). Rather mundane social media postings can turn out to contain highly sensitive information. Health, gender, or financial information, to name a few examples, can be used to enable privacy or security harm (Ohm, 2015). This makes SMA subject to high ethical standards for the way research is conducted. So far, however, clear definitions and guidance about how to deal with social media data in research projects is missing. In the IS literature, an infinitesimal number of scholars addresses ethics with explicit focus on the field of social media analytics (Anderson et al., 2019; Zafeiropoulou et al., 2015). There is hardly any consensus about ethical principles within online and computer research or, more specifically, SMA (Gruzd et al., 2020). Ethical guidance for SMA resembles a farrago with roots in disparate disciplines, individual judgement, and the need to comply with (supra-)local legislations such as the GDPR in countries of the European Union. Consequently, we argue, researchers conducting SMA face an additional challenge:

5) ambiguity of ethical principles.

In the past decade, ethical conduct in SMA research appeared to be of secondary relevance due to the early age of the field, little institutional control, and lack of research towards consensus. The former factors, however, diminish with SMA being an accepted methodology in high ranked IS journals and, at the same time, higher ethical standards expected by publishers and project funders. From a researcher's standpoint, however, following ethical guidelines in SMA research is inherently problematic. Throughout the SMA process, the individual investigator finds herself in various ethical dilemmas. The latter, also referred to as moral dilemmas, are situations that involve conflicts between different moral requirements (Foot, 1983). For instance, it is practically impossible to receive informed consent by many social media users *and* explore the communication patterns of a particular event. Neither it is possible to comply with the terms and conditions of (most) social media platforms *and* collect supposedly public social media data. As a result, we argue, SMA researchers are subject to another challenge:

6) the urgency to resolve ethical dilemmas.

In order to reduce the ambiguity of existent ethical principles that apply to SMA research in IS, and to lay bare current ethical dilemmas of SMA research, this paper aims to answer the following research questions:

RQ1: What is known from existing scholarship in IS and its reference disciplines about ethical principles in Social Media Analytics research?

RQ2: Which ethical dilemmas (if any) need to be resolved by IS scholars in Social Media Analytics research?

To answer these questions, we conducted a scoping literature review, which aims to map extant scholarship in a field primarily based on volume and characteristics of the published research. The review covers 58 relevant publications from IS and reference disciplines. By using keyword, forward and backward searches, we collected a sample of literature that portrays how the topic of SMA research ethics is represented in IS outlets and what cumulative tradition this debate builds on. This paper aims to contribute to existing literature by bringing

together existing efforts of ethical guidance in SMA research that are mainly unconnected to the interdisciplinary nature of the field. Moreover, this work attempts to raise awareness for the ethical dilemmas that result from combining ethical principles and practical SMA research. In order to mitigate this conflict, we propose a research agenda for SMA research ethics in IS and provide propositions about how to approach possible dilemmas. Thus, we contribute to both the debate on SMA ethics within the IS discipline and the individual practice of SMA researchers.

The paper is structured as follows. First, we provide a methodological background about the conduct of scoping reviews, which underlies our study. Additionally, we report our systematic search and exclusion criteria. Second, we present our findings, i.e. ethical principles and dilemmas that have been discussed in SMA research or emerge from the juxtaposition of the reviewed literature. We conclude with a research agenda, including recommendations for further research and a discussion of limitations.

2 Scoping Review Methodology

To provide a general conspectus on the present state of research in terms of SMA ethics, we choose the approach of a scoping review. It constitutes a literature-based approach to synthesise research evidence and map the extant scholarship in a field, including volume and characteristics of the research (Arksey & O'Malley, 2005; Paré et al., 2014). According to Mays et al. (2001), a scoping review turns out to be useful if a research topic has not been extensively reviewed or is particularly complex or of heterogenous nature. The topic of this study, that is, ethical conduct in SMA research, meets all those criteria, and therefore, is an appropriate subject of a scoping review. The expected contribution of this approach comprises the summary and dissemination of extant research findings, the identification of research gaps, and the scope and value of systematic reviews narrowed down to specific research questions within the domain (Arksey & O'Malley, 2005; Levac et al., 2010).

This work is based on the framework proposed by Arksey and O'Malley (2005) and additional suggestions made by Levac et al. (2010). The framework provides five phases to conduct a scoping review: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, and (5) collating, summarizing, and reporting the results.

2.1 Research questions

This scoping review is guided by the questions 'what is known from existing IS scholarship about ethical principles in Social Media Analytics Research?' and 'which ethical dilemmas (if any) need to be resolved by IS scholars in Social Media Analytics Research?'. We are aware that SMA research is a field of interdisciplinary efforts and is of concern for both social sciences (e.g., sociology, media and communication studies, business studies, economics, political science, and social psychology) and natural and applied sciences (e.g., computer science, information systems, linguistics, statistics, or physics) (Stieglitz et al. 2014). However, each of these disciplines follows its own methodological idiosyncrasies and approaches to SMA research with differing perspectives. To contribute to the body of knowledge of IS research at the best possible rate, we use the IS discipline as a starting point, to scope the debate within this discipline and its research identity that focuses on "*how IT systems are developed and how individuals, groups, organizations, and markets interact with IT*" (Sidorova, Evangelopoulos, Valacich, & Thiagarajan, 2008, p. 467). However, through iterative forward and backward searches (vom Brocke et al., 2015), publications from reference disciplines will be identified

that inform the debate in IS. This will also yield insights about what those reference disciplines are[1]. This is vital as neither SMA nor ethics are fields that are discussed only in IS.

We decided to divide this work into two areas of interest, that is, the identification of existing *ethical principles* and the juxtaposition of potential *ethical dilemmas* present in the debate about SMA research. By doing so, we argue, this scoping review will not only be able to uncover what the debate is composed of but also to expose possible contradictions, and to give recommendations for further directions.

2.2 Data sources and search criteria

We initiated our literature search on August 26th, 2020 by using the tool *litbaskets.io*[2] . It is a browser-based application connected to *scopus*[3] and tailored for IS researchers by providing a curated set of IS-related journals (Boell & Wang, 2019). We chose the largest preconfigured basket *3XL*, which includes, at the time of the search, 847 journals that refer to the IS discipline. Having a broad basket of outlets serves the purpose of a scoping review to focus on breath rather than depth (Arksey & O'Malley, 2005). Subsequently, we performed the same initial search on the AIS Electronic Library (AISeL) to expand the search to outlets that might not be considered by litbaskets.io but might be part of the leading database for IS research. The combination of using *litbaskets.io* and the AISeL allowed us to include all relevant IS outlets in our search. For the search query, terms were used that were considered by the authors to describe the topic of social media analytics and ethics: "Social Media Analytics" OR "Social Media Analysis" AND "moral" OR "ethic" OR "ethical" OR "ethics". The search query was generated by *litbaskets.io* and then inserted on *scopus*. For AISeL, the same query was tailored manually to fit the characteristics of the database. We made the a priori decision that a screening of relevant articles should only cover the first 100 hits (as sorted by relevance) per database. This is because the search algorithms of databases vary in broadness to a great extent. The above search string yielded 4 results on scopus and 8,044 on AISeL. Of those hits, the first 100 results per database were coded as relevant or not, following suggestions from the literature that further screening is unlikely to reveal a greater number of relevant publications (Pham et al., 2014; Stevinson & Lawlor, 2004). Duplicates were excluded prior to the initial screening.

2.3 Eligibility assessment

The screening was performed independently by the two authors and included a review of the title and abstract, a brief paraphrase of the topic and an assessment, whether the article is considered relevant or not. For the keyword search, an article had to clearly deal with social media analysis techniques and, at least in the abstract, address ethical considerations of the research practice. This is a complete list of all inclusion and exclusion criteria for the assessment after the keyword search:

---

[1]The term "reference discipline refers to *"those disciplines X that provide foundational, methodological, or other inputs to another discipline/s Y such that the state of knowledge in Y is advanced through inputs provided by X. In other words if Y cites X in order to develop and advance the state of its knowledge, X becomes a reference discipline for Y, and vice versa."* (Grover et al. 2006, p.337)

[2]https://www.litbaskets.io/, last accessed 09/12/2020

[3]http://scopus.com/, last accessed 09/12/2020

- studies had to be peer-reviewed and published in journals or conference proceedings
- studies must deal with social media and include an ethical reflection of data analysis techniques
- due to limited resources for translation, articles published in languages other than English were excluded.
- to be able to map the current debate on this topic, we limited the search to articles published after 2010

After screening the first 20 articles, a reviewer agreement kappa of 0.95 was calculated, which is considered as high reviewer agreement (> 0.8) (Randolph, 2005). Subsequently, a forward and backward snowball search was conducted with each relevant article to collect further appropriate publications by screening the reference lists (backward search) and seeking for relevant publications that have cited these papers (forward search) (vom Brocke et al., 2015). The inclusion and exclusion criteria for the forward and backward searches are the following:

- the time limitation (articles published after 2010) was repealed.
- we did not exclude articles that did not deal with social media

The reason for this is that we found that the debate on SMA ethics is essentially not new. It builds on established concepts and principles, which have been developed pre-social media. Therefore, to understand the debate on SMA research ethics, the cumulative tradition of ethical conduct must be considered. Moreover, we altered the following criterion:

- we did not exclude non-scientific sources in the backward search.

The reason for this is that a handful of non-scientific publications were cited by various research papers, which indicates that those publications are seminal for the debate on SMA research ethics.

The two researchers met on a regular basis and discussed the approach to resolve possible inconsistencies. As no changes in the coding process were necessary after the reliability pretest, the two coders arrived at an overall kappa of 0.90 after reviewing all data. To manage and perform the coding, Microsoft Excel was used, whereas the authors used Mendeley as a literature management system.

3 Results

3.1 Search results and selection of relevant articles

In total, the search yielded 58 relevant articles. From those articles, 15 originated from the database search, 15 from the first iteration of forward and backward search, 21 from the second iteration, 4 from a third iteration, and 3 from a final iteration. All articles could be gathered in full-text, and thus were included in the review. Many studies were excluded, especially in the initial screening, as the search yielded a lot of publications with a methodical focus on social media analytics techniques but without any consideration of ethics. Figure 1 provides an overview of the article selection process.

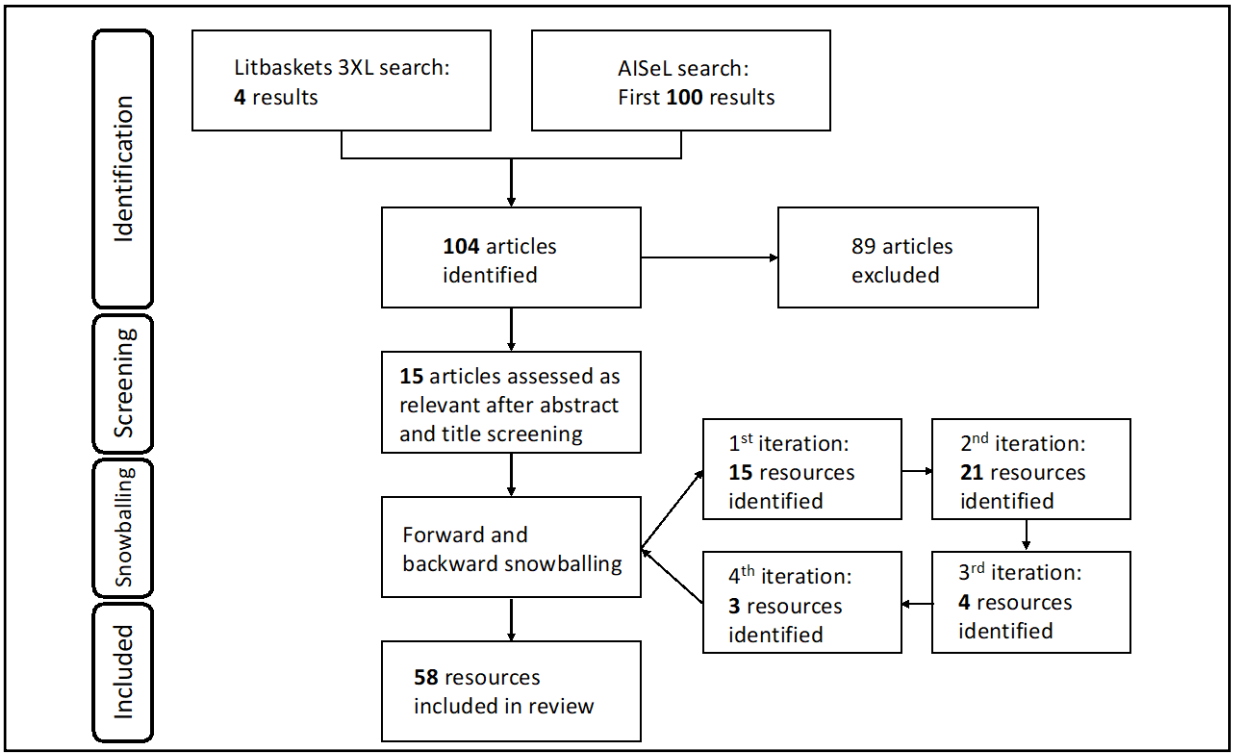

Figure 1. Literature search process

The literature sample shows a high variety of studies in terms of outlets and originating disciplines addressing ethical issues that form a basis for potential debate in IS research. A particularly small portion of the resources stem from core IS outlets, whereas most of the research was published in journals from reference disciplines. Table 1 provides an overview and characteristics of selected articles.

Discipline	Total number of publications	Types of publications	Authors
Information Systems (IS)	6	Conference Proceedings (6)	Anderson et al., 2019; Gruzd et al., 2020; Kaschesky et al., 2013; Malsbender, Voigt, Beverungen, & Rosemann, 2013; Shropshire, 2014; Zafeiropoulou et al., 2015
Computer Science	5	Journal Article (3), Conference Proceedings (2)	Carpenter & Dittrich, 2011; Fiesler, Lampe, & Bruckman, 2016; Libert, 2015; Mavroudis & Milne, 2016; Vitak, Shilton, & Ashktorab, 2016
Information Science	4	Journal Article (4)	Crawford & Finn, 2015; Fiesler & Proferes, 2018; Zimmer, 2010; Zimmer & Proferes, 2014
Media Studies	10	Book (5), Book Section (2), Journal Article (3)	Bakardjieva & Feenberg, 2000; Couldry, 2010; Couldry & Powell, 2014; Curran, Fenton, & Freedman, 2012; danah boyd, 2010; Dijck & Poell, 2013; Dubois, Gruzd, & Jacobson, 2020; Fuchs, 2014; Gillespie, 2014; Trottier, 2012
Sociology	3	Journal Article (3)	Allen, 1996; Kennedy, Elgesem, & Miguel, 2017; Kennedy & Moss, 2015
Psychology	4	Journal Article (4)	Beninger et al., 2014; Conway & O'Connor, 2016; Roberts, 2015; Suler, 2004
Philosophy	2	Journal Article (2)	Elgesem, 2002; Nissenbaum, 2011
Medicine	18	Journal Article (18)	Brotsky & Giles, 2007; Conway, 2014; Eysenbach & Till, 2001; Franz, Marsh, Chen, & Teo, 2019; Golder, Ahmed, Norman, & Booth, 2017; Golder,

			Scantlebury, & Christmas, 2019; Guillemin & Gillam, 2004; Heilferty, 2011; Ioannidis, 2013; Mayer & Till, 1996; Mikal, Hurst, & Conway, 2016; Moreno, Fost, & Christakis, 2008; Moreno, Grant, Kacvinsky, Moreno, & Fleming, 2012; Rothstein & Shoben, 2013; Swirsky, Hoop, & Labott, 2014; Vayena, Salathé, Madoff, & Brownstein, 2015; Vitak, Proferes, Shilton, & Ashktorab, 2017; Yeung, 2018
Law	4	Book (1), Journal Article (3)	Cohen, 2012; Heimer & Petty, 2010; Kerr, 2001; Solove, 2013
Economics	1	Journal Article (1)	Hair & Clark, 2007
Other	1	Newspaper Article (1)	Johnson, 2010

Table 1. Overview and characteristics of selected articles

Scoping the literature on SMA ethics provides several useful insights on how the debate on this matter unfolds. First, IS can solely contribute conference papers (6) to the debate. Consequently, IS researchers heavily rely on the work provided by reference disciplines. Even though computer science contributed a couple of journal articles (3), the predominant discipline on research ethics and SMA is medicine with 18 journal articles. Figure 2 provides an additional point of view on the development of the debate.

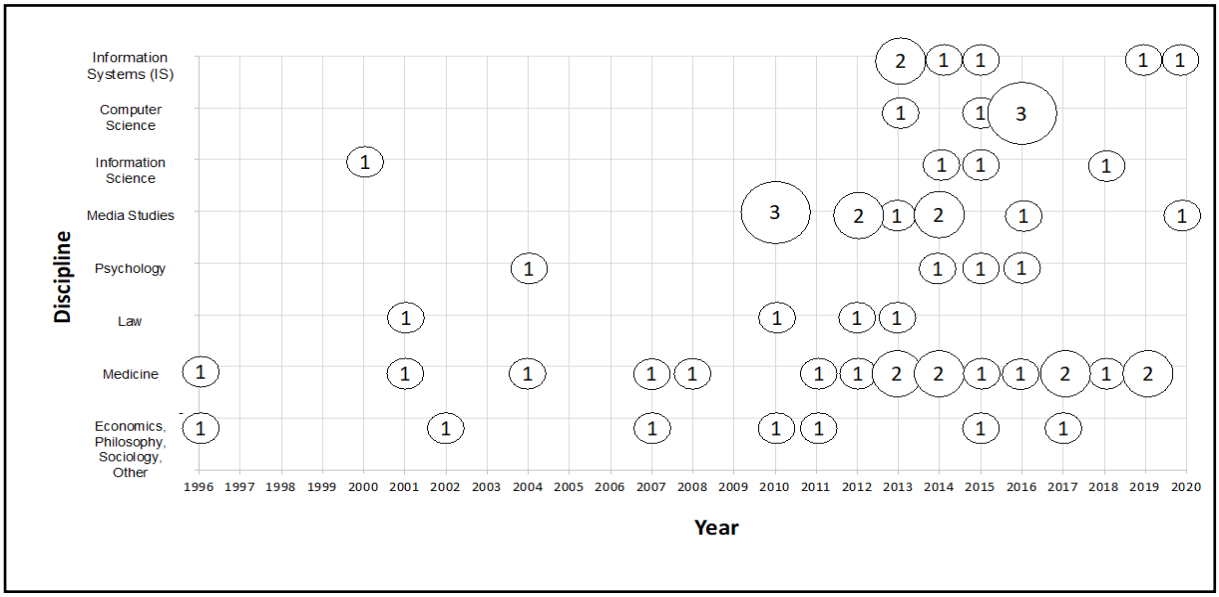

Figure 2. Bubble plot of identified studies and their respective disciplines.

The distribution shown in figure 2 exemplifies how the debate of SMA ethics evolved across disciplines. Studies discussing ethics of online research pre social media (1996-2004) were predominantly published either in medicine or across various disciplines with lesser touchpoints to IS. Later, in the post social media phase (2007-2020), debates began in other disciplines, including media studies and computer science. Remarkably, IS and computer science are the disciplines with the latest start in any research activity on SMA research ethics. Furthermore, all articles published in IS on this matter are part of conference proceedings, leaving no journal articles contributing to this debate.

3.2 Ethical principles of SMA research considered in IS

To map the contents of the selected articles and identify the ethical principles covered by extant literature, we inductively derived emerging themes from the full texts. We screened the articles for content that was particularly labelled as ethical 'principles', 'challenges', or 'issues'. Subsequently, we categorised all principles that apply to SMA research. This content analysis approach was performed according to the suggestions regarding inductive category formation made by Mayring (2014). Here, we did not paraphrase the contents but jumped directly to the level of abstraction that we aimed for (ethical principles and challenges). We revised the categories after coding 50% of the papers and calculated an interrater agreement kappa of 0.79, which counts as substantial agreement (Randolph, 2005). Overall, this analysis resulted in eight main categories. Each category represents one ethical principle that was addressed in the literature. Table 2 provides an overview of the identified principles and in which studies it they were mentioned.

Ethical Principle	**Supporting Literature**	**Total no. of times mentioned**
Preserving privacy	Allen 1996; Mayer & Till 1996; Kerr 2001; Eysenbach & Till 2001; Elgesem 2002; Hair & Clark 2007; Moreno et al. 2008; Zimmer 2010; Johnson 2010; Heilferty 2011; Nissenbaum 2011; Moreno et al. 2012; Trottier 2012; Malsbender et al. 2013; Solove 2013; Swirsky et al. 2014; Beninger et al. 2014; Zimmer & Proferes 2014; Conway 2014; Shropshire 2014; Libert 2015; Roberts 2015; Vayena et al. 2015; Kennedy & Moss 2015; Vitak et al. 2016; Conway & O'Connor 2016; Fiesler et al. 2016; Mikal et al. 2016; Golder et al. 2017; Franz et al. 2019; Dubois et al. 2020	31
Obtaining informed consent	Eysenbach & Till 2001; Elgesem 2002; Hair & Clark 2007; Moreno et al. 2008; Zimmer 2010; Rothstein et al. 2013; Malsbender et al. 2013; Carpenter & Dittrich 2011; Ioannidis 2013; Beninger et al. 2014; Zimmer & Proferes 2014; Conway 2014; Roberts 2015; Crawford & Finn 2015; Kennedy & Moss 2015; Vitak et al. 2016; Conway & O'Connor 2016; Golder et al. 2017; Fiesler & Proferes 2018; Anderson et al. 2019; Franz et al. 2019	21
Minimising harm and maximising benefit	Allen 1996; Mayer & Till 1996; Elgesem 2002; Hair & Clark 2007; Swirsky et al. 2014; Beninger et al. 2014; Vayena et al. 2015; Kennedy & Moss 2015; Vitak et al. 2016; Golder et al. 2017; Fiesler & Proferes 2018; Golder et al. 2019	12
Producing good research	Moreno et al. 2008; Shropshire 2014; Roberts 2015; Vayena et al. 2015; Vitak et al. 2016; Conway & O'Connor 2016; Golder et al. 2017	7
Transparency	Allen 1996; Bakardjieva & Feenberg 2000; boyd 2010; Eysenbach & Till 2001; Hair & Clark 2007; Brotsky & Giles 2007; Carpenter & Dittrich 2011; Couldry & Powell 2014; Shropshire 2014; Libert 2015; Roberts 2015; Vayena et al. 2015; Zafeiropoulou et al. 2015; Vitak et al. 2016; Golder et al. 2017	15
Fairness	Moreno et al. 2008; Crawford & Finn 2015; Vayena et al. 2015; Kennedy et al. 2017; Fiesler & Proferes 2018; Anderson et al. 2019	6
Data Minimisation	Zimmer 2010; Zimmer & Proferes 2014	2
Diversity	boyd 2010; Kaschesky et al. 2013; Golder et al. 2017	3

Table 2. Identified ethical principles and number of mentions.

Preserving privacy

Safeguarding the privacy of research subjects in the context of SMA turns out to be the most frequently discussed ethical principle in IS literature and its reference disciplines. This principle particularly addresses the integrity and autonomy of each individual social media user and aims at the "*protection of the individual's ability to form reasonable expectations about how information about him or her will be treated in different channels*" (Elgesem, 2002, p. 201). A related notion mentioned in the literature is confidentiality, that is, to protect personal identifiers and sensitive information, even though they are publicly available on social media (Moreno et al., 2008). Consequently, research should be designed in a way that inhibits the ex post re-identification of subjects (Zimmer, 2010). Safeguarding the privacy rights of data subjects is crucial as potential risks such as data aggregation, profiling and selective targeting may arise (Nissenbaum, 2011), alongside the risk of general surveillance (Trottier, 2012). From a researcher's perspective, anonymisation and pseudonymisation techniques are the obvious measure to respect this principle (Roberts, 2015). Based on the papers included in this review, a consensus emerges that only public data should be part of SMA research (Conway & O'Connor, 2016; Mikal et al., 2016; Fiesler et al., 2016). However, this maxim entails an often discusses and partly unresolved dilemma: what data stemming from social media can be considered public and what data is private? Many factors such as differing contexts or applied philosophies determine the outcome of this decision (Solove, 2013). As Fiesler (2016) point out, studies have shown *"striking misunderstandings of copyrights in one's own content—for example, that any content posted online automatically becomes public domain"* (p.1452). Here, the individual perception of privacy, including researchers, often differs from how it is defined in social media platforms' terms and conditions. Especially under consideration of the different nature of public, private, and semi-public spaces created by social media, this problem remains important for prospective research on SMA ethics.

Dilemma: *A researcher faces the often-difficult task to draw a line between public and private social media space, which determines the eligibility of data.*

Obtaining informed consent

Building on the argument of privacy protection, the literature on hand defines informed consent as another key principle of ethical SMA research. The procedure of obtaining informed consent is rooted in research involving human participants (Eysenbach & Till, 2001). The literature that contemporary IS papers rely on largely stem from the discipline of medicine, in which standards were established for clinical studies involving social media. In its most basic manifestation, informed consent is achieved through explaining the characteristics of a particular study to research subjects and asking them for permission (Hair & Clark, 2007). To obtain informed consent from social media users, researchers need to reach out to them individually and ask for consent to obtain their data. One could argue that using social media and confirming a platform's terms and conditions would equal this type of consent. However, research suggests that users are not aware of how their data is used, and therefore, might not consent (Anderson et al., 2019). The same might apply to consent forms handed out by researchers. Such forms are often difficult to understand for social media users and cannot fully convey the implications of the study (Fiesler & Proferes, 2018). Another stream of SMA research, that is, aiming at identifying patterns in large scale social datasets rather than (n-)ethnographic research, faces another pivotal dilemma: It is practically impossible to obtain consent from millions of social media users (Franz et al., 2019; Golder et al., 2019).

Dilemma: *A researcher is asked to obtain informed consent of social media users while often experiencing its impracticability.*

Given this dilemma, some scholars have reflected upon the question whether it is necessary to obtain consent from a large social media user base or if other measures should be taken. Kennedy and Moss (2015), for example advocate for a democratisation of SMA, which is subject to (1) greater public supervision and regulation, (2) available and accessible to a greater public, and (3) used to educate the public to become reflexive, active and knowing. SMA projects that use retrospective data to learn about events such as crises, elections, or social movements, can obtain consent from users who would be highlighted in form of quoted postings or descriptions of their roles and profiles.

Nevertheless, dissent remains whether to treat SMA as human subject research (e.g. and ethnographic field study about an online community) or humanities research (e.g. a social network analysis of an election campaign) (Heimer & Petty, 2010; Rothstein & Shoben, 2013). This distinction clearly determines the ethical standards that should be applied to a given study. Most ethical principles that have emerged in the reviewed literature have been developed against the backdrop of human subject research designs. Here, SMA research ethics need to differentiate and give room for reflexivity (Guillemin & Gillam, 2004) and 'in-flux' ethical considerations (Vitak et al., 2016) on the part of the researcher – dependent on context and data.

Dilemma: *When performing SMA, a researcher must distinguish between human subject research and humanities research to make ethical decisions.*

In case a SMA endeavour is classified as involving human subjects and informed consent is obtained, a researcher might face an additional problem: Consent bias. Participants may show altered behaviour in the way they use social media once they have given consent to be researched (Rothstein & Shoben, 2013). For example, in closed communities or groups an unintended behaviour change might occur as a result of the presence of a researcher. As we learn from the discipline of medicine, debates revolve around loosening the requirements for informed consent in case of research that is informational rather than interventional (Ioannidis, 2013).

Dilemma: *As a result of obtaining consent, research findings of SMA studies may be compromised by consent bias.*

Minimising harm and maximising benefit

The third pillar of ethical SMA research is the principle of avoiding harm and, at the same time, creating benefit for the research subject, the community or public under scrutiny, and society. Possible (perceived) harm for social media users may comprise unsolicited attention on the Web and, "abuse" or bullying, but also exploitation from organisations or use by authorities (Golder et al., 2017). In contrast to medical studies involving social media, in which physical harm constitutes a serious risk of data misuse, IS and computer science might be more concerned with the psychological and social nature of harm. Vitak et al. (2016), for instance refer to the "Golden Rule", that is "*do to others what you would have them do to you.*" (p. 945). This deontological approach may fundamentally differ from other ethical philosophies. Separate philosophies can lead to different outcomes and weighing of harm and benefit. In contrast to a normative approach, a researcher could also apply a utilitarian perspective and pursue the greatest good for the greatest number of people (Hair and Clark, 2007). However,

the overall tone of research on this matter shows a tendency towards always pursuing the common good (Vayena et al., 2015; Vitak et al., 2016).

Dilemma: *A researcher must determine what it means to be 'ethical' and consider sometimes conflicting ethical philosophies in order to make decisions.*

Producing good research

A deeply entrenched ethical principle that applies to SMA research is the researcher's responsibility to conduct 'good' research. It is mentioned primarily in the studies that were published pre-social media but appeared as important reference literature ever since (Allen, 1996; Mayer & Till, 1996). Not to be confused with benefitting the common good, this principle aims at a high validity of deployed research methods as well as rigor (Golder et al., 2017). Moreover, data retrieved from social media must be interpreted with the context of the data in mind (Anderson et al., 2019). This "contextual integrity" (Vayena et al., 2015, p.3) is imperative to collect, analyse, and theorise data with the needed sensitivity for the context of its origin, that is, platform characteristics, user history, socio-demographic specificities etc. Regarding this principle, too, problems may arise such as "collapsed contexts" (boyd, 2010, p. 11), if different viewpoints about what is appropriate or ethical collide. Contributing factors to this problem are a missing consensus on guidelines and misalignment of bureaucratic and professional ethics among ethics boards (Heimer & Petty, 2010; Vitak et al., 2017) and additional regulations such as GDPR. The fact that the individual researcher must adhere to top-down (supposed) ethical and legal standards and do justice to her own ethical standards may confront her with an ethical dilemma.

Dilemma: *A researcher must distinguish between the regulation of science (boards), legal frameworks, and the regulation of ethics, which may be misaligned.*

Transparency

Early ethnographic studies in online and social media involved the intrusion of researchers in virtual communities to study them. A prominent example for this kind of research was the deception inside the "pro ana" virtual community described by Brotsky and Giles (2007). Here, a researcher became part of a community related to the eating disorder anorexia nervosa without disclosing herself as a researcher. This seminal study lead to a debate about whether deception can be justified if a greater good is being served. Earlier studies that are part of this review clearly emphasise that the phenomenon shall not be perturbed (Allen, 1996). However, after Brotsky and Giles' seminal paper, most authors do not categorically oppose the means of deception if it is very well justified (Vitak et al., 2016). In recent years, however, the field of SMA involves much more analyses that include a much larger number of research subjects. Carpenter and Dittrich (2011) note that this type of research "*diminishes […] transparency and creates a distance between the researchers and potentially impacted parties*" (p.2). As this distance leads to researchers not considering their work as human subject research, they argue, new and consistent ethical standards need to be developed.

Dilemma: *A SMA methodology might create an unprecedented researcher-subject relationship which is not covered by existing ethical standards.*

Fairness

According to the findings by Kennedy et al. (2017), social media users expect a fair use of their data when SMA is applied. In other words, SMA research is fair when it "*meets users'*

expectations about the collection and use of social media data" (Kenney et al., 2017, p.17). Once social media users publish information in a particular context online, their data rights will be violated in case the analysis removes the data from this context. Avoiding this practice has also been coined as "contextual integrity" (Nissenbaum, 2011). Moreover, the fairness principle foresees that data shall not be exploited and researchers aim to handle the data in a just manner (Fiesler & Proferes, 2018). At the same time, studies that questioned social media users found that the general public is largely unaware about the possibility that their social media activity might be subject to research work (Fiesler & Proferes, 2018; Golder et al., 2017). Research projects increasingly implement awareness campaigns to inform the general public about their research. This can happen via social media, blogs, or other public relations measures.

Dilemma: *Too few users are aware about the fact that their social media postings might be analysed by researchers, leaving the researcher with the question: who is responsible for this more general awareness campaign?*

Data Minimisation

With the increasing number of research subjects in datasets obtained from social media, the protection of personal profile data becomes a challenge for researchers (Zimmer, 2010). This entails the need for automated and reliable anonymisation and pseudonymisation techniques. Moreover, this development also poses the question of how much data is needed to answer a particular research question. Consequently, Zimmer & Proferes (2014) conducted a systematic analysis of Twitter-based academic research, finding a trend towards increasingly larger social media datasets. Additionally, with the implementation of GDPR in 2018, the data minimisation has become a legal matter. Both research and data protection regulations point towards the importance of the data minimisation principle, that is, to treat personal data as "*adequate, relevant, and limited to what is necessary in relation to the purposes for which they are processed*" (Art. 5, GDPR). As Franz et al. (2019) point out, a dilemma emerges from the responsibility that is placed upon the researcher.

Dilemma: *Despite the power and resources of platform vendors, the responsibility to ensure that social media users' data are obtained minimised and protected is placed on the researcher.*

Diversity

Hair and Clark (2007) examined ethnographic research in online communities in terms of the range of ethical dilemmas and challenges. Their argument stresses the point that each Facebook group, Twitter feed and YouTube comment section is different and will involve dissimilar ethical challenges. The fact that different technological structures create diverse patterns of behaviour was also made by boyd (2010). Consequently, it becomes an intricate task for the researcher to apply standards, guidelines and adhere to regulations to the analysis of these highly diverse environments. Therefore, Hair and Clarke (2007) argue, it might be helpful to follow a relativistic approach to ethics that involves independent decisions and the open negotiation of the research practice within a discipline.

Moreover, according to the reviewed literature, SMA approaches should consider the diverse nature of society and consider the fact that some groups and minorities might be underrepresented when following certain sampling methods. The terminology regarding the principle of diversity is quite inhomogeneous. Mentioned are concepts that refer to power and agency such as "invisible audiences" (boyd, 2010), people's "voice" (Couldry, 2011; Couldry & Powell, 2014), and "segments of society" (Kaschesky et al., 2013). In the reviewed literature,

Kennedy and Moss (2015) underline a dilemma that is grounded in a power difference between researcher and data subject.

Dilemma: *SMA can yield an unequal relationship between societal groups and the researcher, creating a lack of agency and equal data power.*

4 Discussion

4.1 Towards a shared code of ethical practice in SMA research

Scoping the literature about SMA research ethics has yielded several striking patterns, which characterise the current state of the debate on this matter inside the IS discipline. In the following, three patterns that could be identified from the literature will be discussed.

First, the IS discipline has missed out on making its mark in the debate on SMA research ethics. Whereas research published in IS journals contributed important methodological groundwork to the interdisciplinary field of SMA (e.g. Stieglitz et al. 2014; Stieglitz et al. 2018), no further efforts have been made to scrutinise SMA methods in terms of their ethical requirements. This became apparent in the results section, with IS research contributing to the debate about ethical principles in SMA research with only six conference papers. Other disciplines such as psychology or medicine have developed a rich discussion on social media research ethics in leading journals (e.g. Conway & O'Connor, 2016; Golder et al. 2017). Admittedly, those are disciplines of much greater proportions than IS but, at the same time, are much further away from SMA being a core concern of their disciplines' identities. To be able to shape the debate on SMA ethics to a much greater extent, IS journals should encourage for more research on this matter by calls for papers and collaboration with other disciplines. The urgency to find better consensus about how SMA research can be made more ethical provides an opportunity for IS to set an example in a core area of its expertise. At this point in time, however, IS remains dependent on reference disciplines to discuss the SMA methodology. Here, IS researchers can use their experience in designing artefacts that resolve dilemmas about privacy, diversity, or data minimisation "by design", e.g. with machine learning algorithms.

Second, extant ethical principles do not reflect the specificities of large-scale SMA approaches. The literature on hand largely develops ethical principles against the backdrop of ethnographic social media research. For example, questions about informed consent have been largely answered in medical research (e.g. Conway, 2014; Franz et al., 2019), but remain unresolved when it comes to "digital humanities", in which large datasets are researched to unearth patterns or mechanisms of a larger user base.

A cut set of ethical principles that are referenced in the literature have been developed pre social media (1996-2004). Therefore, one could pose the legitimate question: What is different with social media? Why do ethical theories and principles from other domains not apply to social media? As this review showed, some ethical principles in fact apply to SMA research as to other methodologies. However, SMA approaches that involve the mass processing of personal data, for example concerning public events, social movements and the like, a considerable number of dilemmas emerge when applying traditional research ethics or even standards that have been derived from ethnographic social media research. One example being the inability to obtain informed consent for large data samples (Anderson et al. 2019).

Consequently, much more scrutiny is needed to exactly identify the distinguishing features of SMA research and finding a shared code of practice within the field.

Third, the protection of the individual researcher has been erased from the equation. The ethical principle that we identified called 'minimising harm and maximising benefit' is almost entirely focused on the integrity of the research subject. In fact, only Golder et al. 2017 reflected upon the protection of the researcher. Surely most researchers handling social media data are experienced in their field of study and can foresee psychological consequences of their research efforts. However, coding large-scale data sets and social media contents oftentimes involves the labour of novice researchers and students. Considering the fact that a multitude of SMA studies involve data revolving around societal crises, e.g. terroristic attacks, natural disasters, or political uprisings, the low efforts concerning researcher protection are astounding. Being confronted with bulk quantities of disturbing imagery and language can take a toll on a researcher's psychological well-being. At this point, automated systems such as image recognition algorithms could filter disturbing material prior to the researchers being exposed to it. However, more consideration needs to be placed upon the protection of the research team.

4.2 A research agenda for the Information Systems discipline

Ultimately, the above-mentioned findings from extant literature give some direction to address the challenges of SMA research ethics and move towards a shared code of ethical conduct in this field. Hereby, we encourage especially IS researchers to reflect their own SMA practices to contribute to a more open and fruitful debate about SMA research ethics. In the following, we propose research questions that might be of interest to the IS discipline. Based on the findings of this literature review, we grouped those exemplary questions along three levels of analysis (see table 3).

Level of analysis	Research question
Institutional	How can the interest in research on SMA ethics be incentivised by IS journals and research associations?
	What are ethical dilemmas that can be addressed by interdisciplinary research projects?
	How can SMA methodologies be made more explainable to improve the collaboration between researchers and ethics boards?
Technological	How can SMA artefacts be designed to incorporate ethical principles 'by design'?
	How can the individual researcher be protected throughout the SMA research process, e.g. through automated filtering?
Individual/Project	How can individual projects be empowered to initiate awareness campaigns to educate the public about the practices, risks, and benefits of SMA research?

Table 3. Examples for possible research questions about SMA research ethics.

This collection of possible research questions only represents a limited number of issues that need to be addressed. Additionally, the identified dilemmas serve as starting points for further research as well. It is up to us, the researchers, to raise questions and add them to the list, to lead fruitful discussions about our own methods, and to reduce the distance between us and our research subjects by transparent communication and exchange. Institutional entities, too, are responsible to do their part and intensify their efforts to assume responsibility to find a shared code of SMA research practice. This will take some weight of the shoulders of the individual researcher and benefit all parties involved in the complex but so highly promising field of SMA.

4.3 Limitations

Our study comes with limitations. In this case, we considered literature that explicitly addresses SMA research ethics. Our search and eligibility criteria, therefore, may have led to the exclusion of studies that primarily applied SMA and reflected about research ethics in a secondary manner. We might have missed relevant hits due to the limitation of screening only the first 100 hits per database, which we justify with practicability and suggestions from the literature. To focus on a purely scientific debate, we did not include institutional guidelines and frameworks or legal texts such as GDPR. Those writings might have given us more depth regarding certain principles, however, a high number of reviewed articles referred to relevant guidelines and regulations and incorporated them into their findings.

5 Conclusion

In this study, we conducted a scoping literature review to determine the state of the art in Social Media Analytics research ethics. A total of 58 identified pieces of literature was charted, summarised, and discussed. The findings uncovered that eight ethical principles are relevant for SMA research, that is, preserving privacy, obtaining informed consent, minimising harm and maximising benefit, producing good research, transparency, fairness, data minimisation, and diversity. At the same time, our analysis revealed that several ethical dilemmas emerge for the individual researcher once each principle is applied. This is true especially for characteristic SMA research that involves large-scale data set involving big samples of data subjects. Based on our review results, we found that (1) the IS discipline has missed out on making its mark in the debate on SMA research ethics, (2) extant ethical principles do not reflect the specificities of large-scale SMA approaches, and (3) the protection of the individual researcher has been erased from the equation. Finally, we derived a research agenda for Information Systems to address those shortcomings and advance the important debate on SMA research ethics.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 823866.

References

Allen, C. (1996). What's wrong with the "golden rule"? Conundrums of conducting ethical research in cyberspace. *Information Society*, *12*(2), 175–188. https://doi.org/10.1080/713856146

Anderson, J., Casas Saez, G., Anderson, K., Palen, L., & Morss, R. (2019). Incorporating Context and Location Into Social Media Analysis: A Scalable, Cloud-Based Approach for More Powerful Data Science. *Hawaii International Conference on System Sciences (HICCS)*, 2274–2283. https://doi.org/10.24251/hicss.2019.275

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology: Theory and Practice*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Bakardjieva, M., & Feenberg, A. (2000). Involving the virtual subject. *Ethics and Information Technology*, *2*(1), 233–240. https://doi.org/10.1023/a:1011454606534

Beier, M., & Wagner, K. (2016). Social Media Adoption : Barriers to the Strategic Use of Social Media. *European Conference on Information Systems (ECIS)*, 1–18.

Beninger, A. K., Fry, A., Jago, N., Lepps, H., Nass, L., Silvester, H., Beninger, K., Fry, A., Jago, N., Lepps, H., Nass, L., & Silvester, H. (2014). Research using social media; users' views. *NatCen Social Research*.

Boell, S. K., & Wang, B. (2019). www.litbaskets.io, an IT Artifact Supporting Exploratory Literature Searches for Information Systems Research. *Australasian Conference on Information Systems (ACIS)*, 663–673. https://www.acis2019.org/Papers/ACIS2019_PaperFIN_141.pdf

Brotsky, S. R., & Giles, D. (2007). Inside the "Pro-ana" community: A covert online participant observation. *Eating Disorders*, *5*(2), 93–109. https://doi.org/10.1080/10640260701190600

Bruns, A., & Burgess, J. (2015). Twitter Hashtags from Ad Hoc to Calculated Publics. In *Hashtag Publics: The Power and Politics of Discursive Networks* (pp. 13–28). https://doi.org/10.3726/978-1-4539-1672-8

Carpenter, K. J., & Dittrich, D. (2011). Bridging the Distance: Removing the Technology Buffer and Seeking Consistent Ethical Analysis in Computer Security Research. In *1st International Digital Ethics Symposium. Loyola University Chicago Center for Digital Ethics and Policy*.

Cohen, J. E. (2012). *Configuring the networked self: Law, code, and the play of everyday practice*. Yale University Press.

Conway, M. (2014). Ethical issues in using twitter for public health surveillance and research: Developing a taxonomy of ethical concepts from the research literature. *Journal of Medical Internet Research*, *16*(12), 1–9. https://doi.org/10.2196/jmir.3617

Conway, M., & O'Connor, D. (2016). Social media, big data, and mental health: Current advances and ethical implications. *Current Opinion in Psychology*, *9*(2), 77–82. https://doi.org/10.1016/j.copsyc.2016.01.004

Couldry, N. (2010). Why voice matters: Culture and politics after neoliberalism. In *Why Voice Matters: Culture and Politics after Neoliberalism*. SAGE Publications. https://doi.org/10.4135/9781446269114

Couldry, N., & Powell, A. (2014). Big Data from the bottom up. *Big Data and Society*, *1*(2), 1–5. https://doi.org/10.1177/2053951714539277

Crawford, K., & Finn, M. (2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, *80*(1), 491–502. https://doi.org/10.1007/s10708-014-9597-z

Curran, J., Fenton, N., & Freedman, D. (2012). *Misunderstanding the internet*. Routledge. https://doi.org/10.4324/9780203146484

danah boyd. (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In *Networked Self: Identity, Community, and Culture on Social Network Sites*. Routledge.

Dijck, J. Van, & Poell, T. (2013). Understanding Social Media Logic. In *Media and Communication* (pp. 2–14). https://doi.org/10.12924/mac2013.01010002

Dubois, E., Gruzd, A., & Jacobson, J. (2020). Journalists' Use of Social Media to Infer Public Opinion: The Citizens' Perspective. *Social Science Computer Review*, *38*(1), 57–74. https://doi.org/10.1177/0894439318791527

Elgesem, D. (2002). What is special about the ethical issues in online research? *Ethics and Information Technology*, *4*(1), 195–20. https://doi.org/10.1023/A:1021320510186

Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *British Medical Journal*, *323*, 1103–1105. https://doi.org/10.1136/bmj.323.7321.1103

Fiesler, C., Lampe, C., & Bruckman, A. S. (2016). Reality and perception of copyright terms of service for online content creation. *ACM Conference on Computer Supported Cooperative Work (CSCW)*, 1450–1461. https://doi.org/10.1145/2818048.2819931

Fiesler, C., & Proferes, N. (2018). "Participant" Perceptions of Twitter Research Ethics. *Social Media and Society*, *4*(1), 1–14. https://doi.org/10.1177/2056305118763366

Foot, P. (1983). Moral Realism and Moral Dilemma. *The Journal of Philosophy*, *80*(7), 379–398.

Franz, D., Marsh, H. E., Chen, J. I., & Teo, A. R. (2019). Using facebook for qualitative research: A brief primer. *Journal of Medical Internet Research*, *21*(8), 1–12. https://doi.org/10.2196/13544

Fuchs, C. (2014). *Social Media: A Critical Introduction*. SAGE Publications. https://doi.org/10.4135/9781446270066

Gillespie, T. (2014). The Relevance of Algorithms. In *Media Technologies* (pp. 167–194). https://doi.org/10.7551/mitpress/9780262525374.003.0009

Golder, S., Ahmed, S., Norman, G., & Booth, A. (2017). Attitudes toward the ethics of research using social media: A systematic review. *Journal of Medical Internet Research*, *19*(6), 1–19. https://doi.org/10.2196/jmir.7082

Golder, S., Scantlebury, A., & Christmas, H. (2019). Understanding public attitudes toward researchers using social media for detecting and monitoring adverse events data: Multi methods study. *Journal of Medical Internet Research*, *21*(8), 1–13. https://doi.org/10.2196/JMIR.7081

Gruzd, A., Abul-Fottouh, D., & Mashatan, A. (2020). Who is Influencing the #GDPR Discussion on Twitter: Implications for Public Relations. *Hawaii International Conference on System Sciences (HICSS)*, 2619–2629. https://doi.org/10.24251/hicss.2020.319

Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and "Ethically important moments" in research. *Qualitative Inquiry*, *10*(2), 261–280. https://doi.org/10.1177/1077800403262360

Hair, N., & Clark, M. (2007). The ethical dilemmas and challenges of ethnographic research in electronic communities. *International Journal of Market Research*, *49*(6), 1–13. https://doi.org/10.1177/147078530704900609

Heilferty, C. M. (2011). Ethical considerations in the study of online illness narratives: A qualitative review. *Journal of Advanced Nursing*, *67*(5), 945–953. https://doi.org/10.1111/j.1365-2648.2010.05563.x

Heimer, C. A., & Petty, J. (2010). Bureaucratic ethics: IRBS and the legal regulation of human subjects research. *Annual Review of Law and Social Science*, *6*(1), 601–626. https://doi.org/10.1146/annurev.lawsocsci.093008.131454

Ioannidis, J. P. A. (2013). Informed Consent, Big Data, and the Oxymoron of Research That Is Not Research. *American Journal of Bioethics*, *13*(4), 40–42. https://doi.org/10.1080/15265161.2013.768864

Johnson, B. (2010). Privacy no longer a social norm, says Facebook founder. *The Guardian*.

Kaschesky, M., Sobkowicz, P., Lobato, J. M. H., Bouchard, G., Archambeau, C., Scharioth, N., Manchin, R., Gschwend, A., & Riedl, R. (2013). Bringing representativeness into social media monitoring and analysis. *Hawaii International Conference on System Sciences (HICSS)*, 2003–2012. https://doi.org/10.1109/HICSS.2013.120

Kennedy, H., Elgesem, D., & Miguel, C. (2017). On fairness: User perspectives on social media data mining. *Convergence*, *23*(3), 270–288. https://doi.org/10.1177/1354856515592507

Kennedy, H., & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data and Society*, *2*(2), 1–11. https://doi.org/10.1177/2053951715611145

Kerr, I. (2001). The Legal Relationship Between Online Service Providers and Users. *Canadian Business Law Journal*, *35*, 1–40.

Levac, D., Colquhooun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation Science*, *5*(69), 1–18. https://doi.org/10.1017/cbo9780511814563.003

Libert, T. (2015). Privacy implications of health information seeking on the web. *Communications of the ACM*, *58*(3), 68–77. https://doi.org/10.1145/2658983

Lomborg, S., & Bechmann, A. (2014). Using APIs for Data Collection on Social Media. *Information Society*, *30*(4), 256–265. https://doi.org/10.1080/01972243.2014.915276

Malsbender, A., Voigt, M., Beverungen, D., & Rosemann, M. (2013). Digital social signal processing - Theoretical underpinning and research agenda. *Pacific Asia Conference on Information Systems (PACIS)*, 1–12.

Mavroudis, J., & Milne, E. (2016). Researching microcelebrity: Methods, access and labour. *First Monday*, *21*(7), 1–20. https://doi.org/10.5210/fm.v21i7.6401

Mayer, M., & Till, J. E. (1996). The Internet: A modern Pandora's Box? *Quality of Life Research*, *5*(1), 568–571. https://doi.org/10.1007/BF00439230

Mayring, P. (2014). *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. SSOAR. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

Mays, N., Roberts, E., & Popay, J. (2001). Synthesising research evidence. In *Studying the Organisation and Delivery of Health Services* (pp. 1–33). Routledge.

Mikal, J., Hurst, S., & Conway, M. (2016). Ethical issues in using Twitter for population-level depression monitoring: A qualitative study. *BMC Medical Ethics*, *17*(1), 1–11. https://doi.org/10.1186/s12910-016-0105-5

Mirbabaie, M., Bunker, D., Stieglitz, S., Marx, J., & Ehnis, C. (2020). Social Media in Times of Crisis: Learning from Hurricane Harvey for the COVID-19 Pandemic Response. *Journal of Information Technology*, *35*(3), 195–213. https://doi.org/10.1177/0268396220929258

Moreno, M. A., Fost, N. C., & Christakis, D. A. (2008). Research ethics in the MySpace era. *Pediatrics*, *121*(1), 157–161. https://doi.org/10.1542/peds.2007-3015

Moreno, M. A., Grant, A., Kacvinsky, L., Moreno, P., & Fleming, M. (2012). Older adolescents' views regarding participation in facebook research. *Journal of Adolescent Health*, *51*(5), 439–444. https://doi.org/10.1016/j.jadohealth.2012.02.001

Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, *140*(4), 32–48. https://doi.org/10.1162/DAED_a_00113

Ohm, P. (2015). Sensitive information. *Southern California Law Review*, *88*(5), 1125–1196.

Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2014). Synthesizing information systems knowledge: A typology of literature reviews. *Information and Management*, *52*(2), 183–199. https://doi.org/10.1016/j.im.2014.08.008

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & Mcewen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, *5*(4), 371–385. https://doi.org/10.1002/jrsm.1123

Randolph, J. J. (2005). Free-Marginal Multirater Kappa. *Joensuu Learning and Instruction Symposium*, 1–20.

Reuter, C., Hughes, A. L., & Kaufhold, M. A. (2018). Social Media in Crisis Management: An Evaluation and Analysis of Crisis Informatics Research. *International Journal of Human-Computer Interaction*, *34*(4), 280–294. https://doi.org/10.1080/10447318.2018.1427832

Richter, D., Riemer, K., & vom Brocke, J. (2011). Internet Social Networking. *Business & Information Systems Engineering*, *3*(2), 89–101. https://doi.org/10.1007/s12599-011-0151-y

Roberts, L. D. (2015). Ethical Issues in Conducting Qualitative Research in Online Communities. *Qualitative Research in Psychology*, *12*(3), 314–325. https://doi.org/10.1080/14780887.2015.1008909

Rothstein, M. A., & Shoben, A. B. (2013). Does Consent Bias Research? *American Journal of Bioethics*, *13*(4), 27–37. https://doi.org/10.1080/15265161.2013.767955

Shropshire, J. (2014). Extending the cloud with fog: Security challenges & opportunities. *Americas Conference on Information Systems (AMCIS)*, 1–10.

Sidorova, A., Evangelopoulos, N., Valacich, J. S., & Thiagarajan, R. (2008). Uncovering the Intellectual Core of the Information Systems Discipline. *Management Information Systems Quarterly (MISQ)*, *32*(3), 467–482.

Solove, D. J. (2013). Introduction: Privacy self-management and the consent dilemma. *Harvard Law Review*, *126*(1), 1–24.

Stevinson, C., & Lawlor, D. A. (2004). Searching multiple databases for systematic reviews: Added value or diminishing returns? *Complementary Therapies in Medicine*, *12*(4), 228–232. https://doi.org/10.1016/j.ctim.2004.09.003

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and Information Diffusion in Social Media - An Investigation of Sentiment of Microblogs and Sharing Behavior. *An Investigation of Sentiment of Microblogs and Sharing Behavior. Journal of Management Information Systems (JMIS)*, *29*(4), 217–248.

Stieglitz, Stefan, Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social Media Analytics: An Interdisciplinary Approach and Its Implications for Information Systems. *Wirtschaftsinformatik*, *56*(2), 101–109. https://doi.org/10.1007/s11576-014-0407-5

Stieglitz, Stefan, Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, *39*(10), 156–168. https://doi.org/10.1016/j.ijinfomgt.2017.12.002

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology and Behavior*, *7*(3), 321–326. https://doi.org/10.1089/1094931041291295

Swirsky, E. S., Hoop, J. G., & Labott, S. (2014). Using Social Media in Research: New Ethics for a New Meme? *American Journal of Bioethics*, *14*(10), 60–61. https://doi.org/10.1080/15265161.2014.948302

Trottier, D. (2012). *Social media as surveillance: Rethinking visibility in a converging world*. Routledge.

Vayena, E., Salathé, M., Madoff, L. C., & Brownstein, J. S. (2015). Ethical Challenges of Big Data in Public Health. *PLoS Computational Biology*, *11*(2), 1–7. https://doi.org/10.1371/journal.pcbi.1003904

Vitak, J., Proferes, N., Shilton, K., & Ashktorab, Z. (2017). Ethics Regulation in Social Computing Research: Examining the Role of Institutional Review Boards. *Journal of Empirical Research on Human Research Ethics*, *12*(5), 372–382. https://doi.org/10.1177/1556264617725200

Vitak, J., Shilton, K., & Ashktorab, Z. (2016). Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. *ACM Conference on Computer Supported Cooperative Work (CSCW)*, 941–953. https://doi.org/10.1145/2818048.2820078

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & 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*, 205–224. https://doi.org/10.17705/1cais.03709

Yeung, D. (2018). Social media as a catalyst for policy action and social change for health and

well-being: Viewpoint. *Journal of Medical Internet Research*, *20*(3), 1–12. https://doi.org/10.2196/jmir.8508

Zafeiropoulou, S., Sarker, S., & Carlsson, S. A. (2015). What's Trending in Social Media Analytics Area? A Retrospective. *Americas Conference on Information Systems (AMCIS)*, 1–15.

Zeng, D., Chen, H., Lusch, R., & Li, S. H. (2010). Social media analytics and intelligence. *IEEE Intelligent Systems*, *25*(6), 13–16. https://doi.org/10.1109/MIS.2010.151

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(1), 313–325. https://doi.org/10.1007/s10676-010-9227-5

Zimmer, M., & Proferes, N. J. (2014). A topology of twitter research: Disciplines, methods, and ethics. *Aslib Journal of Information Management*, *66*(3), 250–261. https://doi.org/10.1108/AJIM-09-2013-0083

doi: https://doi.org/10.3127/ajis.v26i0.3287