Media and Communication
Open Access Journal | ISSN: 2183-2439
Volume 4, Issue 4 (2016)

Successes and Failures in Studying Social Media: Issues of Methods and Ethics

Editors

Epp Lauk and Niina Sormanen

COGITATIO

Media and Communication, 2016, Volume 4, Issue 4
Issue: Successes and Failures in Studying Social Media: Issues of Methods and Ethics

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Academic Editors
Epp Lauk, University of Jyväskylä, Finland
Niina Sormanen, University of Jyväskylä, Finland

Available online at: www.cogitatiopress.com/mediaandcommunication

This issue is licensed under a Creative Commons Attribution 4.0 International License (CC BY). Articles may be reproduced provided that credit is given to the original and *Media and Communication* is acknowledged as the original venue of publication.

Table of Contents

Media and Communication (ISSN: 2183-2439)
2016, Volume 4, Issue 4, Pages 63-65
doi: 10.17645/mac.v4i4.793

Editorial

Issues of Ethics and Methods in Studying Social Media

Niina Sormanen and Epp Lauk *

Department of Communication, University of Jyväskylä, 40014 Jyväskylä, Finland; E-Mails: niina.sormanen@jyu.fi (N.S.), epp.lauk@jyu.fi (E.L.)

* Corresponding author

Submitted: 3 October 2016 | Published: 10 October 2016

Abstract
The Editorial raises some challenging ethical and methodological aspects of Internet based research (such as protection of informational privacy, informed consent, general ethical guidelines vs case-based approach), which are further discussed in the five articles of this special issue.

Keywords
ethics; Facebook; Internet based research; participation; social media

Issue
This editorial is part of the issue "Successes and Failures in Studying Social Media: Issues of Methods and Ethics", edited by Epp Lauk and Niina Sormanen (University of Jyväskylä, Finland).

© 2016 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Along with the exponential growth of both online information and digital platforms, the emergence of the social web has increased interest towards Internet based research. Simultaneously, the debate on the issues related to the ethics and methods of this research has been expanding (e.g. Markham & Buchanan, 2012; Moreno, Goniou, & Moreno, 2013; Walther, 2002). Questions about ethical protection of research subjects, the ethical ways of gathering and using the data as well as the validity of the Internet based data are of the utmost importance, especially in those fields, in which human subjects are central to research. From the perspectives of research ethics and methods, online and social media seem to be the most problematic fields, as they enable a vast quantity of qualitative and quantitative approaches. Also, no general ethical rules of thumb as yet exist to guide research across all digital platforms and in accordance with the laws of particular countries.

It has been argued that the ethical problems that arise in Internet based research are basically similar to problems typical to the research within humanities and social sciences, but they still have some special aspects (Elgesem, 2002). Elgesem refers primarily to the private/public distinction, which in online environment obtains different dimensions than offline, and makes protection of informational privacy more difficult. The social media era also raises new questions, for example about the meaning of 'informed consent' in 'big data' studies; about the application of copyright principles; about anonymity and confidentiality of data; should social media be seen as a source or tool for research, etc.

The idea for this special issue originated from various unanswered questions: should the intertwined methodological and ethical choices of research be guided by, for instance, law, the cultural context, specific social media platform regulations, general ethical association guidelines such as The Association of Internet Researchers (AoIR) (see Markham & Buchanan, 2012) or specific university ethical boards' guidelines, or perhaps scholars' own ethical decisions and common sense?

The Internet and social media platforms and sites have also defined their own specific rules of peoples' personal information and communication data usage and, for example, big data mining. For example, Facebook's user data can generally be considered public as users can individually determine the information they are willing to share publicly. However, Facebook has determined its own terms of what and how data can be

mined for research purposes (e.g. Facebook, 2014; see also Sormanen et al., 2016). Public data, available on any social media site, does not automatically mean it has any unproblematic availability for research (e.g. Zimmer, 2010). In the user agreements of social media platforms users (information owners) have usually agreed with the use of their data for platform improvements, content optimization and marketing purposes, but not for research purposes.

In addition to ethical guidelines and any site-specific terms and conditions, there are general codes of ethics and laws to be considered when conducting research online. For instance, for scholars operating in Finland, the two most prominent laws that define research online are the Copyright Act and the Personal Data Act. A large part of the data available on the Internet can be considered as copyrighted material, such as blog posts and photographs. The researcher should therefore use appropriate references when copying or using material or ideas available online. The Personal Data Act is a crucial guideline in conducting research online, since most of the problems and ambiguities concerning the ethics of online research tend to involve issues about informational privacy. The legal approaches, indeed, vary country by country and the way they are followed, by cultural context. Therefore, also "ethical considerations should be more case-sensitive instead of relying on one model for all solutions", as the authors of the article in this issue—"Hazy Boundaries: Virtual Communities and Research Ethics"—Helena Kantanen and Jyri Manninen (2016) argue. Furthermore, online research is most often multidisciplinary and researchers have different disciplinary backgrounds. Thus, their experiences with research ethics also vary according to their specific backgrounds. Sari Östman and Riikka Turtiainen (2016), therefore, in their article "From Research Ethics to Researching Ethics in an Online Specific Context" suggest that the focus should be moved from defining general ethical guidelines to studying research ethics.

The major ethical concerns regarding social media big data research are the possible misuse and abuse of the information gathered. The risks include violations of personal privacy, civil rights and consumer freedoms (Bollier, 2010). When conducting massive data mining processes and projects, it is nearly impossible to receive consent from every individual from whom the data are collected. In these cases the researcher needs to secure the privacy of the individuals and ascertain that the information does not end up in the wrong hands and is not used for criminal purposes (see Acquisti & Gross, 2009).

When looking at social media from big data gathering and usage perspectives, informed consent should be requested when possible even in observational research, and should not obstruct the results of the study. Social media experimental research is always more risky from the ethical perspective, and demands more reflection on its impact on the research subjects. These aspects will be further discussed in one article of the issue, "Facebook's Emotional Contagion Experiment as a Challenge to Research Ethics" by Jukka Jouhki, Epp Lauk, Maija Penttinen, Niina Sormanen and Turo Uskali (2016).

Two articles of this issue concentrate on the methodology of Internet based research. Johanna Sumiala, Minttu Tikka, Jukka Huhtamäki and Katja Valaskivi (2016), (in "#JeSuisCharlie: Towards a Multi-Method Study of Hybrid Media Events") introduce a three-phase multi-method approach for the analysis of hybrid media events. The authors outline a model, in which the research process moves from preliminary digital ethnography to quantitative social network analysis and lastly to in-depth interpretation, demonstrating how links and connections in the hybrid media landscape can be disclosed. Mikko Villi's and Janne Matikainen's (2016) article discusses a methodologically challenging issue in studying social media: "Participation in Social Media: Studying Explicit and Implicit Forms of Participation in Communicative Social Networks". They argue that too little attention has been paid to what constitutes participation when the users create connections rather than content. Unlike explicit participation, implicit participation does not involve any conscious participation, but combining different accounts (e.g. Facebook with Spotify) the users' information automatically becomes available on the other platform as well. This implicit user participation produces data that is valuable to marketers. The authors call for more attention to the research of implicit participation than has so far been employed.

The articles published in this special issue are selected from among the papers presented at the international symposium "Successes and Failures in Studying Social Media: Issues of Methods and Ethics" held on 20 November 2015 in the University of Jyväskylä, Finland.

Acknowledgments

The editors of this special issue appreciate the financial support from the University of Jyväskylä and the invaluable input of the reviewers.

Conflict of Interests

The authors declare no conflict of interests.

References

Acquisti, A., & Gross, R. (2009). Predicting social security numbers from public data. *Proceedings of the National Academy of Science*, *106*(27), 10975–10980. doi:10.1073/pnas.0904891106

Bollier, D. (2010). *The promise and peril of big data*.

Washington DC: The Aspen Institute. Retrieved from http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf

Elgesem, D. (2002). What is special about the ethical issues in online research? *Ethics and Information Technology*, *4*(3), 195-203.

Facebook. (2014). *Statement of rights and responsibilities*. Retrieved from https://www.facebook.com/legal/terms

Jouhki, J., Lauk, E., Penttinen, M., Sormanen, N., & Uskali, T. (2016). Facebook's emotional contagion experiment as a challenge to research ethics. *Media and Communication*, *4*(4), 75-85.

Kantanen, H., & Manninen, J. (2016). Hazy boundaries: Virtual communities and research ethics. *Media and Communication*, *4*(4), 86-96.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making and internet research. Recommendations from the AoIR ethics working committee (Version 2.0)*. Chicago: Association of Internet Researchers. Retrieved from http://aoir.org/reports/ethics2.pdf

Moreno, M. A., Goniou, N., & Moreno, P. S. (2013). Ethics of social media research: Common concerns and practical considerations. *Cyberpsychology, Behavior, and Social Networking*, *16*(9), 708-713.

Östman, S., & Turtiainen, R. (2016). From research ethics to researching ethics in an online specific context. *Media and Communication*, *4*(4), 66-74.

Sormanen, N., Rohila, J., Lauk, E., Uskali, T., Jouhki, J., & Penttinen, M. (2016). Chances and challenges of computational data gathering and analysis: The case of issue-attention cycles on Facebook. *Digital Journalism*, *4*(1), 55-74. doi:10.1080/21670811.2015.1096614

Sumiala, J., Tikka, M., Huhtamäki, J., & Valaskivi, K. (2016). #JeSuisCharlie: Towards a multi-method study of hybrid media events. *Media and Communication*, *4*(4), 97-108.

Villi, M., & Matikainen, J. (2016). Participation in social media: Studying explicit and implicit forms of participation in communicative social networks. *Media and Communication*, *4*(4), 109-117.

Walther, J. B. (2002). Research ethics in internet-enabled research: Human subjects issues and methodological myopia. *Ethics and Information Technology*, *4*(3), 205-216.

Zimmer, M. (2010). But the data is already public. On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313-325.

About the Authors

Niina Sormanen (MA) is a PhD candidate of Organizational Communication and Public Relation (PR) at the University of Jyväskylä, Department of Communication. Her research interests include communicative behavior and power relations in the social media context. Her PhD thesis is focused on the interplay of organizational and media professionals and individuals in the social media context and uses of social media in building their communicative power.

Epp Lauk (PhD) is Professor of Journalism and Head of the Department of Communication at the University of Jyväskylä, Finland. Her research and publications focus on journalism cultures and history, media and journalism in Central and East European countries, media self-regulation and innovations in journalism.

Media and Communication (ISSN: 2183-2439)
2016, Volume 4, Issue 4, Pages 66-74
doi: 10.17645/mac.v4i4.571

Review

From Research Ethics to Researching Ethics in an Online Specific Context

Sari Östman * and Riikka Turtiainen

School of History, Culture and Arts Studies, Faculty of Humanities, University of Turku, 28100 Pori, Finland;
E-Mails: sari.ostman@utu.fi (S.Ö.), riikka.turtiainen@utu.fi (R.T.)

* Corresponding author

Submitted: 28 January 2016 | Accepted: 31 August 2016 | Published: 10 October 2016

Abstract
Along with the rise of a research field called digital humanities, online specific research ethics plays an especially significant role. Research on the same (Internet related) topic is usually multidisciplinary, and understanding research ethics even inside the same research community may vary essentially. It is important to recognise and pay attention to online specific contexts as well as the researcher's own disciplinary background. In this empirical research paper, we will first sum up our previous work. Currently, we are working on a model which will help in positioning multidisciplinary researchers as ethical actors based on their research topics and backgrounds. In this article, we will present this model with a demonstration of the empirical data collected as part of a Finnish research project called *Citizen Mindscapes*, which concerns the cultures and history of Finnish discussion forums. We argue that in Finland, and probably also worldwide, online research ethics is in a phase where the focus should be moving from defining the ethical guidelines to studying research ethics as such. We will also discuss how the model will be further developed in an in-depth empirical process.

Keywords
digital culture; digital humanities; Internet; online specific research ethics; online studies; research ethics

Issue
This review is part of the issue "Successes and Failures in Studying Social Media: Issues of Methods and Ethics", edited by Epp Lauk and Niina Sormanen (University of Jyväskylä, Finland).

© 2016 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

As an online researcher, how do you understand the ethical research process? Do you work in a field where research ethics is being taught starting the first years of your university studies? Or did you, as a junior researcher, have to stumble your way through your first online specific research process trying to recognise and solve the ethical questions by yourself? Or, maybe you are from a discipline in which ethical questions are not a major concern?

We have worked with research ethics for years; teaching, publishing, developing Finnish understanding about *online specific research ethics* (see Turtiainen & Östman, 2013). In Finland, research ethics and online studies are both quite novel objects to academic interest (Pekkala, 2000). Until now, the ethical interest in Finland has mainly concerned creating guidelines for conducting online-related research. However, such guidelines have existed internationally for at least 15 years (Ess, 2002; Baym & Markham, 2009; Markham & Buchanan, 2012) and the Finnish ones have not differed from those significantly (Kuula, 2006; Turtiainen & Östman, 2009).

In the rise of *digital humanities,* online specific research ethics has also come to a situation where these basic principles become as important as ever. Simultaneously, we need to pay attention to new ethical aspects: digital humanities bring together researchers from multiple various disciplinary backgrounds, which

means multiple ethical backgrounds. We work in the project *Citizen Mindscapes* (*CM24*)[1], which brings together researchers, among others, from the following disciplines: sociology, psychology, digital culture, computer science, language technologies, statistics, anthropology and history. All of these researchers have rather varied opinions about which kind of ethics should be followed when studying *Suomi24*, Finland's oldest and largest online discussion forum. Östman is the responsible researcher in a work package, which concerns research ethics. Among the project, we have noticed that in order to create new, multidisciplinary understanding about online specific research ethics, we need to make it so that ethics is the actual object of research. Therefore, we are moving on from our previous work with developing online specific research ethics to *researching online research ethics*. In this report, we question: *Which kind of variety of ethical viewpoints exists in a multidisciplinary, online specific research project?* With this question, we will start mapping the field out for *how the ethics can be studied.*

Our current data is preliminary and comprises 16 answers to an open-answer survey in a project consortium seminar in May 2016. We asked the participants about their disciplinary background, their understanding and experiences of research ethics and the ethical challenges and questions in their current research situation.

While teaching and publishing about research ethics, we have developed certain self-reflexive methods for (novice) online researchers. We understand ethics as passing through the entire research process from choosing the subject to reporting and societal discussion. Therefore, it is also combined with every stage of the process and, in our opinion, especially with source criticism. In this report, we will present one of these methods which combines ethics with source criticism. This is an important aspect in qualitative online research (Baym & Markham, 2009, pp. xii–xv). At the end of the report, we will tie the benefits of this method to our current findings.

First, however, we will discuss the understanding of research ethics among digital humanities. We will consider the linkage between ethics and methodology, especially source criticism. Then, we will present the story-method and 'room board' which reminds one of certain basic principles in online specific research ethics. At the end, we present a model for organising and pre-analysing the data collected from the *CM24* participants. Finally, we will briefly discuss how these findings seem to lead us towards a new ethical paradigm—studying the ethical processes themselves.

[1] The project is funded by the Academy of Finland's Digital Humanities Programme. About this, see: *Academy of Finland grants funding for Citizen Mindscapes in a Social Media project*: https://www.uef.fi/en/-/suomen-akatemia-rahoittaa-kansakunnan-mielenliikkeet-sosiaalisessa-mediassa-tutkimusta

2. Digital Humanities

First, we will briefly introduce the concept of *digital humanities*. The term has gained extensive visibility in the academic societies in the Northern countries during the last couple for years. It was coined in the early 2000s, although it has not been possible to pinpoint it to any specific researcher or discipline. In 2010, however, the Professor of English Literature, William Pannapacker, announced digital humanities to be 'the next big thing' in research, since the digital technology had affected every possible discipline. A year later, he continued this argument on his blog in *The Chronicle of Higher Education* by writing that actually there is no 'next' to that; and, that very soon digital humanities will only be 'the humanities'. (Pannapacker, 2011; see also Spiro, 2012, p. 16). A mere year after that, media and communication researcher David Parry (2010) blogged about how digital humanities would fundamentally challenge all of the current humanistic research.

According to the researchers of digital culture, Jaakko Suominen and Anna Haverinen, holistic change was strongly linked to the discourse about the digital revolution at the end of the 1990s. At that time, several projects were started which, for example, aimed at digitalising humanistic research material. In addition, tools for computer-assisted production and handling of data were developed. (Suominen & Haverinen, 2015).

Digital humanities was evolved in a situation where Western societies have been changing rapidly. It does not have centuries or even decades of tradition, which would have given it a stabilised definition. As Suominen and Haverinen write, it is a discipline—or an era of research—which has been actively developed by researchers. Digital humanities lives and further evolves along with the technologies applied in the current societies. (Suominen & Haverinen, 2015).

There are at least two separate ways of defining digital humanities: for example media theorist Gary Hall as well as media and technology researchers Leighton Evans and Sian Rees see it as humanistic research appropriating information technological methods, which in their opinion is not necessarily a good thing (Evans & Rees, 2012, p. 29; Hall, 2013, pp. 2-3, 133-134; see also Suominen & Haverinen, 2015). The other orientation emphasises a wider understanding about the current, digitalised culture; it includes theorising the digital world more than just applying computer-assisted methods in producing and analysing the materials. (Suominen & Haverinen, 2015). Our understanding follows the latter orientation. In positioning ourselves in the field of digital humanities, we have utilised Suominen and Haverinen's model (Figure 1) for evaluating the relationship of the researcher/research project with digitality. The model is based on the two orientations mentioned above, and on whether the research in question emphasises digitality itself as mak-

ing and developing things or as a theoretical understanding about the digitalised world.

According to Suominen and Haverinen, this coordination might help to define any researcher's or project's, even networks' or consortiums' (e.g. *CM24*), internal relationships concerning digitality. This would be of significant help especially in multi-/interdisciplinary research. (Suominen & Haverinen, 2015). As researchers of digital culture, we would position ourselves quite far on the right, more in the top than the lower corner; however, the vertical position depends extensively on every current research topic and may, therefore, vary a lot. It is also possible for us to move up and down as well as left and right on the axels according to the stage of research, even inside an individual project. After all, the division between the practical and theoretical digital humanities is not biased as much as a moving line.

With our main focus more on the understanding end, we are conducting research which quite often involves human beings either as subjects or informants of the study. In addition, we see the Internet and digitality as a *research environment* that ethnologists might call *the field*. Ethnographical methods are often applied to these kinds of studies. It is especially in these cases, but also with other kinds of methods, that ethics is extensively intertwined with methodology. Next, we will discuss what we mean by this.

3. Self-Reflexive Ethical Tools for Digital Humanities

In the English-speaking academic world, online studies have existed since the early 1990s, and less than a decade later, online specific research ethics was also taken into account (see Baym & Markham 2009, pp. viii-ix). In Finland, however, research ethics in general only became an object of interest in the 1980s, but it did not get much attention until the 1990s (Pekkala, 2000). The Internet generalised in Finland since c. 1995, and only started to appear in research more widely a decade later. Since it is becoming increasingly more common to Finnish scholars to publish internationally, we need to update our late-born understanding of the online research ethics and further develop it. This is what we are aiming at in the *CM24*'s ethics-related work-package. This further work, reaching for studying ethical processes as such, will also be generalisable more widely; we find this approach beneficial for multiple disciplinary backgrounds in international academic cultures.

The challenges that the traditional Finnish disciplines had to face due to this novelty, the Internet where people shared and lived their lives, became clearly visible to us around 2008. At that time, we were working on our online-related doctoral theses in Digital Culture. In 2008 and the coming years, we got several

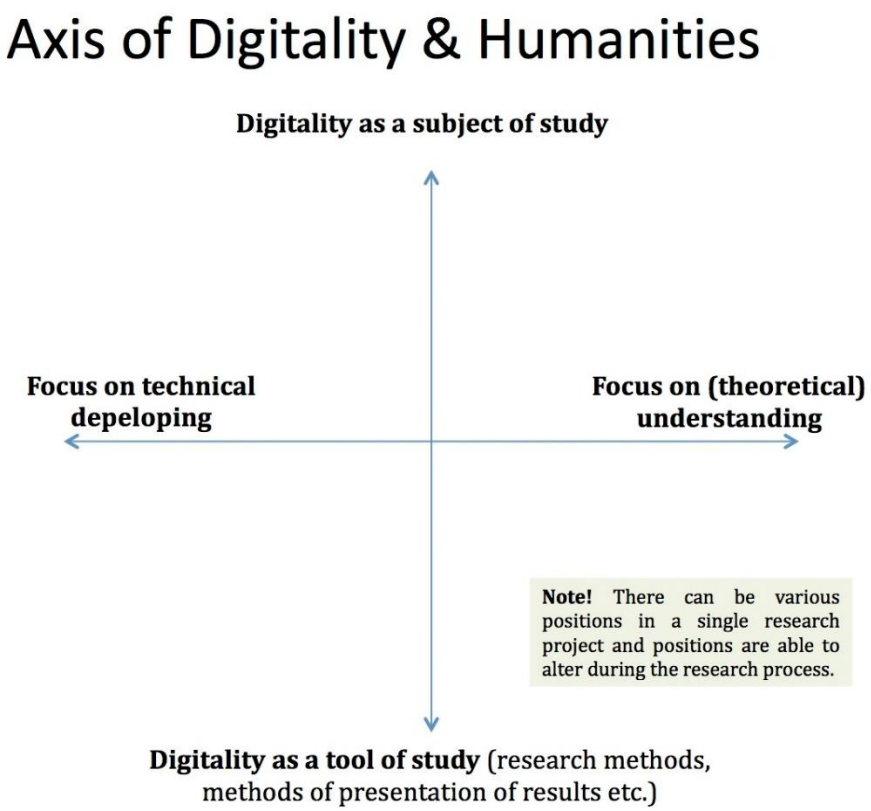

Figure 1. Model for positioning yourself among Digital Humanities research. Source: Suominen & Haverinen, 2015 (S. Östman & R. Turtiainen, Transl.).

questions from doctoral students, lecturers and researchers from other disciplines (for example ethnology, anthropology etc.) concerning online studies. We were asked, for example, on what conditions we could trust our informants to be what they claimed to be; we were confronted about asking people's permission for studying them. It seemed to be a common idea that when people voluntarily published pieces of their lives online, these lives could be used for research without questioning their publicity and authors' right for the contents they had created.

On the Web, everything builds to the context: how the sites are built, what they are used for, who uses them, how the users communicate, how they feel about their contents…these elements vary a lot. Not knowing enough might lead to misinterpretations or (unintended) disrespectful usage of material (see Aull Davies, 2008; Ess, 2002). We have created a tool with certain basic principles for online researchers: 'a Room Board of Ethics' (Figure 2): in it, we ask whether the Internet is your *tool, source* or *subject*; are you conducting research *with, on* or *about* the Internet. Depending on your relationship to the Internet as a researcher, the Internet might be your *research environment* in a way which in ethnographic studies might be called *the field*. In these cases, it should also be seen as any other ethnographical field: you need to get hands-on experience in order to understand its mechanisms.

There are some tools for essential ethical consideration. The question of whether something can be used freely or more carefully lies at the bottom. Asking for permission, rights for citing, identities of informants/subjects to research and so on seem to be important ethical matters even still in 2016. Malin Sveningsson and others have presented in *Att fånga nätet. Kvalitativa metoder för Internetforsking* (*To Study the Web. Qualitative Methods for Internet Studies*, Sveningsson, Lövheim, & Bergquist, 2003) a coordination system (Figure 3) to which researchers may position their subject, informants, material and so on,

BOARD OF ETHICS FOR ONLINE RESEARCHER
(R. Turtiainen & S. Östman 2013,
see also partial transl. A. Haverinen 2014, 61)

Is the internet your	**in other words, is your research**
tool	*with the internet*
source ***or***	*on the internet* ***or***
subject	*about the internet*

To an internet researcher, the web is often a combination of all the abovementioned, in which case the internet (or a piece of it) is ***the research environment***, where the researcher must inter alia:

* remember that ethics is related to **source criticism**
* appropriate the **cultural conventions** of the particular research environment
* understand the **contextualities** of how the research material is being created in the environment
* consider the **intimacy or publicity** of the content from the point of view of the informants (or the members of the community, if you just examine the texts)
* **respect** the people being researched and the contents they produce
* remember that online environments are not static, but constantly changing, which is why all research ethics are to be considered **case-by-case**

The researcher should also remember that the responsibility and freedom of science always walk hand in hand. You need to be able to stand for your choices: you need to **reason your every choice.**

Figure 2. The Roomboard of Ethics: Some guidelines and basic principles for online research (See also Turtiainen & Östman, 2013).

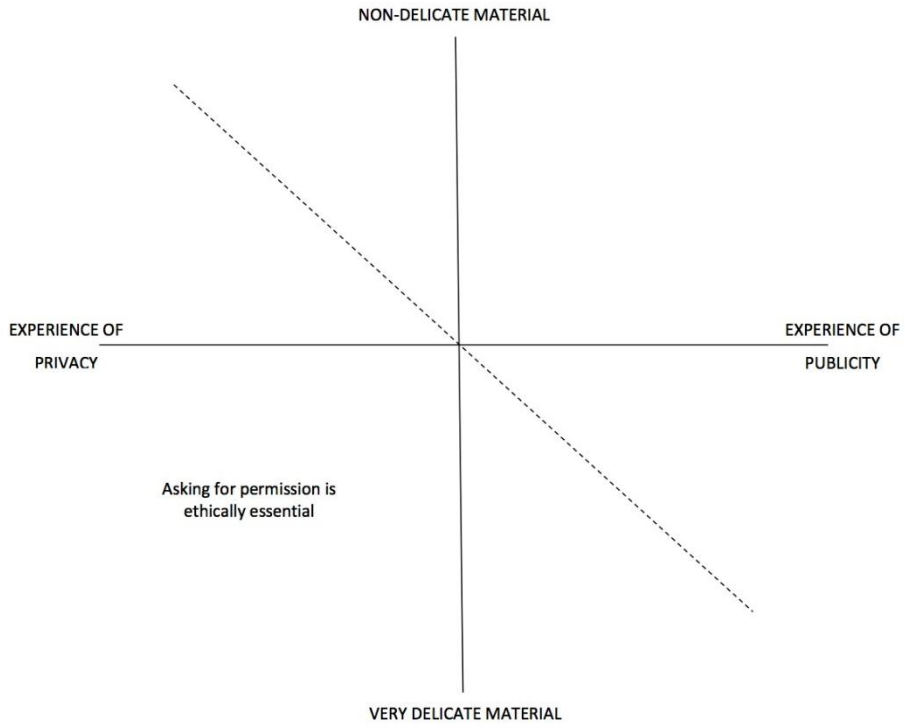

Figure 3. Delicacy of the material. Source: Sveningsson, Lövheim and Bergquist, 2003, p. 186 (the figure has been applied for example by: McKee & Porter, 2008, p. 73, 2009, p. 11; Turtiainen, 2012, p. 52; Östman, 2015, p. 74).

according to how delicate the material is (to a reader) and how private or public the creators consider it. This most probably requires contact with the authors—or at least an extensively thorough reading by the researcher in order to learn the authors' understanding about their actions and culture: this understanding might only exist implicitly in their media texts. Either way, Sveningsson's method is very useful.

We recommend Sveningsson's model for any researcher whose data consist of users' personal online contents. This tool, however, does not offer a practical walk-through of the ethical decision-making process as an entity. For that, we have developed a story-method which has proved useful among Digital Culture undergraduates as well as graduated, but still novice online researchers.

4. 'Mary's Mistakes'—A Method for Recognising Online Specific Ethical Matters

'Mary's mistakes' is a method that we have built for recognising ethical challenges in online specific studies. It is a story with as many ethical dilemmas as possible, for which Östman invented a prototype some years ago for teaching. We developed this idea further in our Finnish article in 2013, and Turtiainen has been using the final version of the story in her teaching ever since. We find Mary's story internationally useful, since it can be applied to any individual case of research, at least in digitally oriented cultural studies. It is our further aim to find out whether or not a similar tool would work in other contexts among Digital Humanities as well (such as extremely theoretical or strictly quantitative/statistical disciplines).

Mary[2] is writing her master's thesis in psychology. She wants to study eating disorders with girls. Her research question is: 'Which body images do 13–15 year old girls with an eating disorder have?' The subject is delicate, and Mary fears that volunteers are hard to find. She studies Digital Culture as a minor, and by linking her study to that, she finds a solution: Mary decides to use some pro-ana[3] discussions and blogs she found on a random discussion site. Mary has had an eating disorder, but has never participated in any Internet activity about it. Actually, she does not use social media almost at all. For her study, she creates a fake profile to the site, with the nickname 'Thinspy-98'. 98 in a nickname usually refers to the birth year—Mary was born in 1991. With this nickname, she starts discussions, asking and commenting, wishing that this way she would get material for her study.

'Mary' has built for herself a great number of traps in the story: for example, her question does not quite

[2] Mary is a fictional character to whom and whose study we have gathered as many ethical problems as we could think of on the basis of the questions we obtained from researchers from other disciplines.

[3] 'Pro-ana' refers to culture in which people with anorexia nervosa consider themselves not ill, but rather willingly choose this life-threatening lifestyle.

match her chosen material; the subject is very delicate as well as the informants, who are teenagers with a severe illness; she has not thought her methods through thoroughly enough (online material is not necessarily the best choice for studying experience); she plans on cheating her informants with the fake profile; she is not going to ask for their permission; she is studying a subject very close to her; yet, she does not know anything about the environment in which she is going to produce her material.

An experienced online researcher might easily recognise all of Mary's traps. For example, in online environments, it is essential to know something about the context: who is using the site, what was it built for, which kind of conventions and habits are there (e.g. are old users and 'newbies'[4] equally allowed to speak up—usually they are not), which kind of tone of voice is usual there, which aims and needs do the users have (e.g. peer support, encouraging others, asking for help, sharing own experiences…) and so on. All these aspects affect the discourse, which is unique on every site, even if they cover the same subject. Knowing the field you are working; knowing who is speaking, what they are saying, to whom, how and why they say it the way they do is basic source criticism, especially in ethnographically emphasised research.

Moreover, every case of research is unique and includes multiple different ethical dilemmas. Writing down your intended research process—for example into a story like this—may help to recognise the case-bounded ethical matters. Mary's story has proved useful for the students in Turtiainen's yearly ethics class for digital culture students, and we also used it in a workshop for an ethnography course at the University of Turku in 2015. The learning process proceeds as followed:

1. Students read Mary's story;
2. The class analyses the story in the lecturer's guidance piece by piece until the class has recognised every ethical risk in the story;
3. The students write their own stories about their bachelor's/master's/doctoral theses;
4. In the course exam, they are given back their own stories for analysis, similar to which the class did together earlier with Mary's story.

The story method has several benefits. It forces the researchers to be self-reflective, which helps them in positioning themselves in relation to the subject as well as practising sustainable source criticism. The doctoral students in our ethnography workshop found this method very useful: in their feedback, they described it as practical, illustrative and eye-opening. Some of the students had not recognised the current challenges, but also understood why certain earlier processes or phases had not succeeded. A doctoral student wrote:

'I found the workshop really useful considering my doctoral studies and future career as a researcher. Especially, I'm thankful for the concrete examples with which both the lecture and the workshop demonstrated the ethical questions that we might face in the research process.'

Next, we will move on from online specific research ethical tools for researchers; we will proceed to discussing the possible tools for actually studying these ethics. By applying our model to preliminary *CM24* data, we will aim at finding out how ethical viewpoints vary among a multidisciplinary project. The next steps after this pilot study will lead us towards creating some commonly applicable guidelines for such widely multidisciplinary projects.

5. Towards Studying Research Ethics

The coordinate system that we presented in the beginning (Suominen & Haverinen, 2015), was suggested to be utilised for scholars to position themselves according to their research focus and motives. This model is a great start for positioning yourself on the field of online studies. It also has inspired us in developing a tool for studying various ethical approaches (Figure 4). We collected some preliminary data in a *CM24* consortium seminar in May 2016. Östman asked the seminar participants to answer the following questions:

1. Disciplinary background and current research;
2. Describe your understanding about research ethics and/or ethically conducted research; which matters do you find the most essential; what kinds of things have affected your understanding?
3. Which ethical questions and/or practices do you expect to face during your current research; how do you plan to proceed with them?

We obtained 16 answers, one of which could not be integrated to our model. We have coded the answers as P1–P15 (P = Participant). The participants came from multiple disciplinary backgrounds, and they integrate into our model (see Figure 4).

The basic idea of this fourfold table is based on Suominen and Haverinen's figure. We have further developed it on the basis of our roomboard: in it, we defined a researchers' relationship with the online environment as being threefold: it may be a tool, a subject or a source for the study. In this example (Figure 4), we have specified 'online environment' as the Suomi24 discussion forum, which is the focus of the whole *CM24* project. Our former studies have shown that the more focal the online environment is and the more understanding the manner of research, the more

[4] New users.

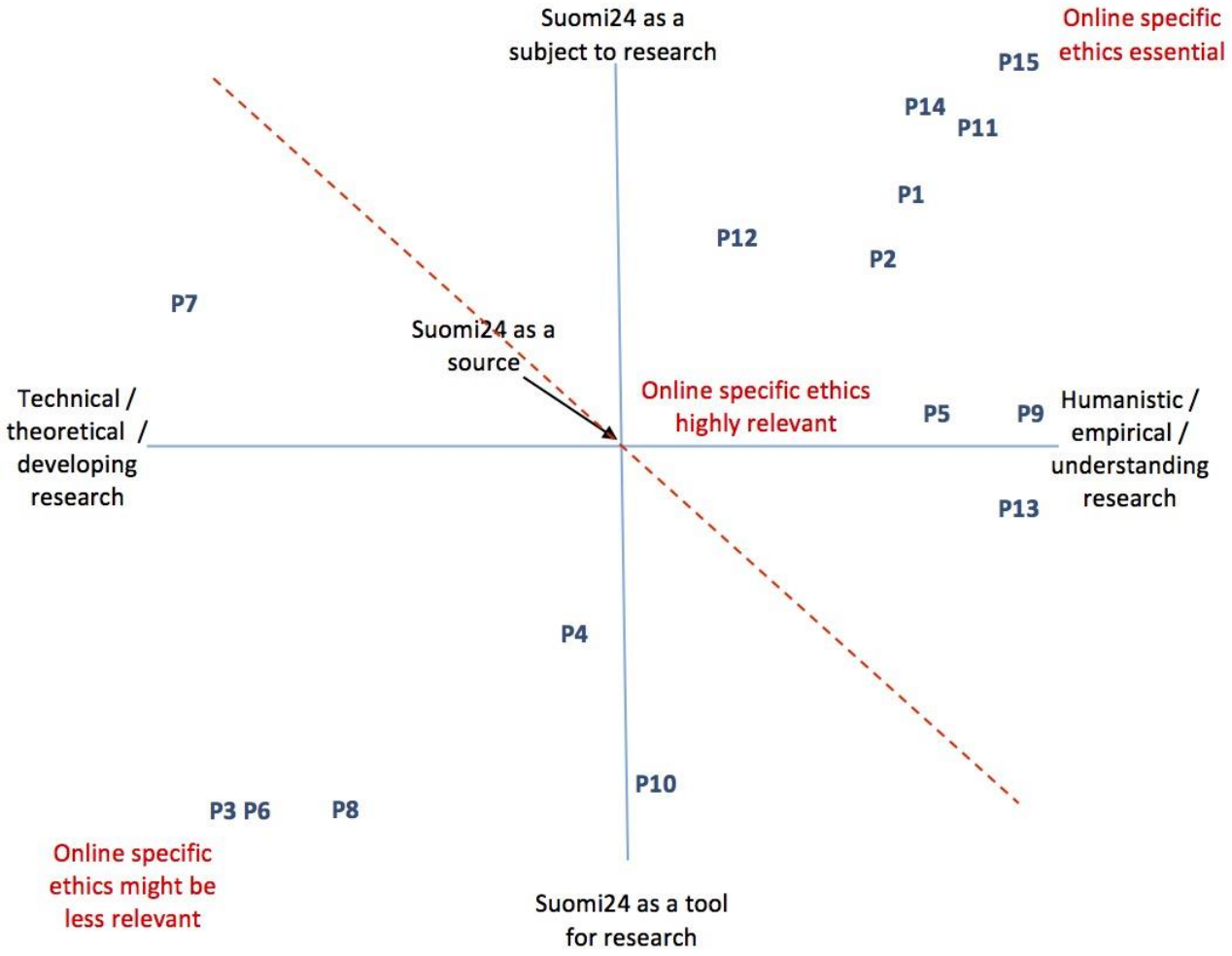

Figure 4. Online environment (Suomi24 discussion forum) in multidisciplinary research.

relevant, even essential, it is to take online specific ethics into account (Östman & Turtiainen, in press).

In Figure 4, the participants P3, P4, P6, P7, P8 and P10 represent quantitatively and statistically oriented disciplines and backgrounds[5]. According to their survey answers, they are situated below the line, which defines the relevance of online specific ethics. Among those, an online specific take to ethics might not be that necessary, which was well articulated in their answers. With their background, ethics mostly have to do with such matters as copyright laws (P3, P7) and open access to the data: 'Privacy protection might be in contradiction with what could be learned from the data' (P4).

Participants P1, P2, P5, P9, P11–P15 are situated above the relevance-defining line. They come from humanistic backgrounds and tend to have human-related, understanding and in several cases, culturally oriented research take: 'Will my analysis do justice to the data; that is, will it give a right kind of a voice to the participants?' (P5). They answered the survey more thoroughly and seemed to have a wider understanding about the research ethical processes; some of them had been taught ethics in their university studies, whereas the first group (below the line) had not been educated about ethics and tended to see it more as legal points rather than decision-processes. However, multidisciplinary research environments helped the participants to appropriate a more reflexive, process-oriented ethical take. P12 had their background in information technology. However, he/she saw their multidisciplinary experience and discussions as an ethically evolving actor, which had led to a reflexive ethical approach.

To be able to consider all the facts needed for ethical decision-making in a multidisciplinary process, the researcher needs to understand at least these three things; the contexts, the research environment and their own background (of importance of recognising the contextual specificities see also Nielsen, Paasonen, & Spisak, 2015, p. 11). Doing this, we suggest that indi-

[5] In order to keep answers anonymised, we cannot specify the individual disciplines.

vidual researchers, even those situated in the low left corner of our tool, might benefit from a case-based self-reflexive ethical analysis such as 'Mary's mistakes' method. The next step of our study will include engaging these participants to apply Mary-method in their own case studies. This will be done in the next consortium seminar in fall 2016. By analysing the results of that step, we will find out whether the ethical co-understanding of a multidisciplinary project will benefit from this kind of case-based learning. Simultaneously, we will further proceed in this new field of studying online specific research ethics as its own entity. Based on this process, we would like to suggest that in Finland, possibly also worldwide, a shift of paradigm from research ethics to researching ethics is about to appear.

6. Summary

In this article, we have presented a) a tool to researchers for positioning themselves in the field of Digital Humanities (Suominen & Haverinen, 2015), b) a method for contextually reflexive ethical decision-making and certain basic principles for that, and finally c) a tool for studying the actual ethical decision-making processes. By leading researchers to actively consider their ethical approaches, we will be able to study research ethics and its practical applying-processes in multiple disciplines. By doing this, we aim at creating a commonly functional and effective guideline for online specific, widely multidisciplinary research. The follow-up for this preliminary study will be carried out in fall 2016 (Step 2: Mary-method) and 2017 (Steps 3 and 4: undefined[6]).

Acknowledgements

The article is part of the *Citizen Mindscapes* project which is funded by the Academy of Finland.

Conflict of Interests

The authors declare no conflict of interests.

References

Aull Davies, C. (2008). *Reflexive ethnography. A guide to researching selves and others*. London and New York: Routledge. (Original work published 1998)

Baym, N. K., & Markham, A. N. (2009). Introduction: Making smart choices on shifting ground. In A. N. Markham & N. K. Baym (Eds.), *Internet inquiry. Conversations about method* (pp. viii-xix). Los Angeles: Sage Publications.

Ess, C. (2002). *Ethical decision-making and Internet research. Recommendations from the AoIR ethics working committee*. Retrieved from http://aoir.org/reports/ethics.pdf

Evans, L., & Rees, S. (2012). An interpretation of digital humanities. In M. B. David (Ed.), *Understanding digital humanities* (pp. 21-41). New York: Palgrave Macmillan.

Hall, G. (2013). Towards a postdigital humanities: Cultural analytics and the computational turn to data-driven scholarship. *American Literature*, *85*(4), 781-809.

Kuula, A. (2006). *Tutkimusetiikka: aineistojen hankinta, käyttö ja säilytys*. Tampere: Vastapaino.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making ad internet research*. Retrieved from aoir.org/reports/ethics2.pdf

McKee, H., & Porter, J. (2008). The ethics of digital writing research: A rhetorical approach. *College Composition and Communication*, *59*(4), 711-749.

McKee, H., & Porter, J. (2009). Playing a good game: Ethical issues in researching MMOGs and virtual worlds. *International Journal of Internet Research Ethics*, *2*(1), 5-37.

Nielsen, S., Paasonen, S., & Spisak, S. (2015). 'Pervy role-play and such': Girls' experiences of sexual messaging online. *Sex Education: Sexuality, Society and Learning*, *15*(5), 472-485.

Östman, S. (2015). *'Millasen päivityksen tästä sais?' Elämäjulkaisijuuden kulttuurinen omaksuminen* (Published Doctoral dissertation). University of Turku, Pori, Finland.

Östman, S., & Turtiainen, R. (in press). How to position yourself as an ethical researcher? Analytical viewpoints to online research ethics. In P. Hakamies & A. Heimo (Eds.), *Doing folkloristics in the digital age.* Turku: Folklore Fellows Communications.

Pannapacker, W. (2011, January 8). Pannapacker at MLA: Digital humanities triumphant? *The Chronicle of Higher Education*. Retrieved from http://www.chronicle.com/blogs/brainstorm/pannapacker-at-mla-digital-humanities-triumphant/30915

Parry, D. (2010). The MLA, @briancroxall, and the non-rise of the Digital Humanities. *AcademicHack*. Retrieved from http://academhack.outsidethetext.com/home/2010/the-mla-briancroxall-and-the-non-rise-of-the-digital-humanities

Pekkala, A. (2000). Tutkijan eettiset ongelmat. In P. Laaksonen, S. Knuuttila, & U. Piela (Eds.), *Tutkijat kentällä* (pp. 85-110). Helsinki: SKS.

Spiro, L. (2012). "This is why we fight": Defining the values of digital humanities. In M. K. Gold (Ed.), *Debates in the digital humanities* (pp. 16-34). Minnesota: University of Minnesota Press.

Suominen, J., & Haverinen, A. (2015). Koodaamisen ja kirjoittamisen vuoropuhelu?—Mitä on digitaalinen humanistinen tutkimus. *Ennen ja nyt.* Retrieved from http://www.ennenjanyt.net/2015/02/koodaamisen-ja-kirjoittamisen-vuoropuhelu-mita-on-digitaalinen-humanistinen-tutkimus/#identifier_14_1502

[6] In order to protect the novelty of our research.

Sveningsson, M., Lövheim, M., & Bergquist, M. (2003). *Att fånga Nätet. Kvalitativa metoder för Internet-forskning*. Lund: Studentlitteratur.

Turtiainen, R. (2012). *Nopeammin, laajemmalle, monipuolisemmin. Digitalisoituminen mediaurheilun seuraamisen muutoksessa*. (Published Doctoral dissertation). University of Turku, Pori, Finland.

Turtiainen, R., & Östman, S. (2009). Tavistaidetta ja verkkoviihdettä. Omaehtoisten verkkosisältöjen tutkimusetiikkaa. In M. Grahn & M. Häyrynen (Eds.), *Kulttuurituotanto. Kehykset, käytäntö ja prosessit* (pp. 336-358). Helsinki: SKS.

Turtiainen, R., & Östman, S. (2013). Verkkotutkimuksen eettiset haasteet: Armi ja anoreksia. In J. Matikainen & M. Tikka (Eds.), *Otteita verkosta. Verkon ja sosiaalisen median tutkimusmenetelmät* (pp. 49-67). Helsinki: Gaudeamus.

About the Authors

Sari Östman (PhD) is a Digital Culture researcher currently working as a researcher in the Citizen Mindscapes project at the University of Turku. She wrote her doctoral thesis (2015) about online life-publishing. Östman has been teaching and studying online research ethics from 2009. Her other interests include online cultures and histories.

Riikka Turtiainen (PhD) is working as a University Lecturer of Digital Culture at the University of Turku. She teaches research skills: research ethics, material collecting methods, research-based thinking and creative writing. Her doctoral thesis (2012) considered digital media sports. Her current research interests focus on new aspects of sports cultures, social media and online research ethics.

COGITATIO

Media and Communication (ISSN: 2183-2439)
2016, Volume 4, Issue 4, Pages 75-85
doi: 10.17645/mac.v4i4.579

Article

Facebook's Emotional Contagion Experiment as a Challenge to Research Ethics

Jukka Jouhki [1,*], Epp Lauk [2], Maija Penttinen [1], Niina Sormanen [2] and Turo Uskali [2]

[1] Department of History and Ethnology, University of Jyväskylä, 40014 Jyväskylä, Finland; E-Mails: jukka.jouhki@jyu.fi (J.J.), maija.s.penttinen@student.jyu.fi (M.P.)
[2] Department of Communication, University of Jyväskylä, 40014 Jyväskylä, Finland; E-Mails: epp.lauk@jyu.fi (E.L.), niina.sormanen@jyu.fi (N.S.), turo.uskali@jyu.fi (T.U.)

* Corresponding author

Submitted: 31 January 2016 | Accepted: 12 April 2016 | Published: 10 October 2016

Abstract
This article analyzes the ethical discussion focusing on the Facebook emotional contagion experiment published by the *Proceedings of the National Academy of Sciences* in 2014. The massive-scale experiment manipulated the News Feeds of a large amount of Facebook users and was successful in proving that emotional contagion happens also in online environments. However, the experiment caused ethical concerns within and outside academia mainly for two intertwined reasons, the first revolving around the idea of research as manipulation, and the second focusing on the problematic definition of informed consent. The article concurs with recent research that the era of social media and big data research are posing a significant challenge to research ethics, the practice and views of which are grounded in the pre social media era, and reflect the classical ethical stances of utilitarianism and deontology.

Keywords
Big data; emotional contagion; Facebook; informed consent; manipulation; methodology; privacy; research ethics; social media; user data

Issue
This article is part of the issue "Successes and Failures in Studying Social Media: Issues of Methods and Ethics", edited by Epp Lauk and Niina Sormanen (University of Jyväskylä, Finland).

© 2016 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

In June 2014 the *Proceedings of the National Academy of Sciences* (*PNAS*) published an article entitled "Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks". It was about an experiment[1] conducted by Adam D. I. Kramer from Facebook's Core Data Science Team together with Jamie E. Guillory and Jeffrey T. Hancock from Cornell University. The article provided experimental evidence about emotional contagion, a phenomenon that has been widely studied before but mostly in offline environments. In January 2012, the research team manipulated the News Feeds of a massive number (N = 689,003) of Facebook users for a week, reducing the amount of emotional content in their feeds. After analyzing over three million posts and over 122 million words, the results showed that when the amount of positive status updates published in their News Feed was reduced, users published more negative status updates and fewer positive updates. Conversely, when the amount of negative status updates was reduced, users published more positive status updates and fewer negative updates. Moreover, the less emotional content the users were exposed to, the fewer words they used in their status updates. (Kramer, Guillory, & Hancock, 2014).

The research suggested that emotional states "can be transferred to others via emotional contagion, lead-

[1] Henceforth, "the Facebook experiment" or "the experiment".

ing people to experience the same emotions without their awareness" (Kramer et al., 2014, p. 8788). Emotional contagion had been proved earlier (e.g. Barsade, 2002; Huntsinger, Lun, Sinclair, & Clore, 2009; Kramer et al., 2014, p. 8788 also refer to several other studies), but proving that it happens "outside of in-person interaction" and particularly in the increasingly popular social media was new (see e.g. Ferrara & Yang, 2015 for a similar but more recent study). Moreover, as there are common conceptions about positive social media postings making people sad or envious (e.g. Copeland, 2011), the experiment produced valuable information to the contrary. The experiment suggested that people's "hearts and minds", as Schroeder (2014, p. 3) puts it, can be manipulated online, for good or ill. (See also Shah, Capella, & Neuman, 2015; Summers-Effler, Van Ness, & Hausmann, 2015, p. 472; cf. Parkinson & Manstead 2015, p. 377.)

Academic and non-academic reactions to the study—defined as ethically controversial (Ananny, 2015, p. 101; Harriman & Patel, 2014; Pejovic & Musolesi, 2015, p. 18; Simon, 2014; Thorson & Wells, 2015, p. 10)—were mixed. On a broader view, the heterogeneity of the views on the ethics of the experiment is a sign of how contested and fluid the concept of privacy is (e.g. Ess, 2013, p. 260). Moreover, as Facebook cooperates with several universities such as Cornell, Stanford and Harvard (see e.g. Cheng, Adamic, Dow, Kleinberg, & Leskovec, 2014; Friggeri, Adamic, Eckles, & Cheng, 2014; Sun, Rosenn, Marlow, & Lento, 2009)[2] the experiment has raised debate about whose research ethics prevail in such joint ventures—those of a private company or those of an academic research institution. In this article, we focus on the academic but also look to some extent at the non-academic ethical commentary on the Facebook experiment, and ask what it tells us about ethical research issues in the current era of social media research.

The ethical discussion presented in this article is founded on an integrative literature review (see e.g. Card, 2010; Torraco, 2005) that we conducted by searching major journal databases such as Science Direct, Google Scholar, Sage Journals, and Ebsco Academic Search Elite for articles covering the experiment. As a result we obtained articles from journals such as *Research Ethics*; *Big Data & Society*; *Media, Culture & Society*; *Nature*; and *Information, Communication & Society*. In addition to journal articles, we searched for conference proceedings on the experiment as well as scholarly analyses of the issue published in blogs and other internet sites. Some news and magazine articles as well as blog posts were also included in order to offer some non-academic views on the issue.

Overall, our approach to the ethical discussion revolving around the Facebook experiment is essayistic in nature (see e.g. Ceserani, 2010; Cornelissen, Gajewskade, Piekkari, & Welch, 2012, pp. 198-199), which means that we prefer exploring and discussing the topic in a heuristic manner: we tend to concentrate on raising questions rather than put forward any definite results based on empirical research. However, we do argue that there are two crucial themes of debate which sum up the ethical discussion revolving around the experiment: *research as manipulation* (discussed in Section 3) and the related *informed consent* (discussed in Section 4). Moreover, we suggest that the debates about the ethics of human-subject big data research, while demanding a rethink of research ethics, still reflect the classical divide between the utilitarian and the deontological points of view. In the next section we will introduce some key questions of research ethics in the era of social media. Then we move on to present the Facebook experiment and the ensuing ethical discussion.[3]

2. Research Ethics and the Human Subject

The views on research ethics generally put into practice in any academic research can be seen as balancing between two classic moral philosophical stances. Utilitarianism attempts to calculate the morality of an act by estimating the total amount of happiness or suffering produced by the act, while deontology views certain actions as immoral or moral per se, regardless of their consequences. Both these stances are applied, for example, in social media research when scholars contemplate the effect of their study on the subjects' privacy: the utilitarian view of privacy might allow certain incursions into privacy if the result is the greater good, whereas from the deontological point of view, a certain level of privacy is a right that should not be violated, for example, by conducting a study without receiving the informed consent of the subjects of the study (Ess, 2013, pp. 256-262; Shrader-Frechette, 2000). Both stances are problematic, and neither of them is applied in research without any consideration of the other—or in moral decision-making outside of academia, for that matter. At any rate, the utilitarian emphasis on avoidance of harm and the more deontological value of receiving informed consent from research subjects are considered the two most significant imperatives of research ethics in studies with human participants (e.g. the British Psychological Society, 2010). Actual policies as to how exactly the imperatives are defined and in what situations they apply (e.g. in big data research) vary significantly.

[2] For Stanford's recent collaboration with Facebook, see www.sserg.org/new-collaboration-with-stanford-university-and-facebook. More about Facebook's partnerships at https://research.facebook.com. On Facebook's exclusive cooperation with a few universities that have been granted access to Facebook's data, see e.g. Paolillo, 2005, p. 50.

[3] This article is based on an unpublished conference paper by Jouhki et al. (2015).

One of the key ethical principles of the Association of Internet Researchers (AoIR)—that the greater the vulnerability of the subject of study, the greater the obligation of the researcher to protect the subject—is a good example of how challenging it is to formulate specific rules of ethical research (Markham & Buchanan, 2012, pp. 4-5). Obviously, protecting the research subject depends on how one defines both the harm that might be inflicted on the unprotected person and also a research subject. The context of any research setting means that ethical codes are not so much strict rules as incentives to individual researchers to reflect on the moral ground of their research and make ethical decisions using their own judgment of what is in fact practicable in the circumstances. Especially when informed consent cannot be obtained in human-subject research, the benefits of the study should outweigh the harm of any invasion of privacy.

Often anonymity is seen as enough to ensure the no-harm rule in cases of non-experimental (e.g. purely observational) research. In experiments that affect the participants' behavior, the rules are stricter (See e.g. Vainio, 2012; Vanderpool, 1996.). The level of sensitivity required for the decision-making to be ethically sufficient is a constant topic of debate. For example, a research institution or a commercial company engaging in research might hold the view that obeying the law is enough to make the research ethical (Hudson & Bruckman, 2004, pp. 132-133). If the participants are not harmed in any way during the data gathering, an ethically sensitive researcher—whether working in a private company or a university—might still take into account the hypothetical situation that a research participant at some point learns about his or her role in the research and is offended (i.e. harmed) by having been a participant without having given consent (e.g. Hudson & Bruckman, 2004, pp. 136-138). Moreover, an ethically sensitive researcher might treat public content on the internet (e.g. tweets, blog posts) as intimate parts of their creator's personhood. Most researchers, however, use this content without securing informed consent (Hesse, Moser, & Riley, 2015, p. 27).

The fact that data are accessible and public does not necessarily mean that using them is not jeopardizing privacy and is thus ethically justified (see e.g. boyd, 2010;[4] Marx, 2013; Tinati, Halford, Carr, & Pope, 2014, p. 673; Zimmer, 2010). The boundaries between private and public information—especially on the internet—are frustratingly ambiguous, contested and changing (Markham & Buchanan, 2012, p. 6; Ess, 2007, p. 499; see also Rooke, 2013; Rosenberg 2010; Weeden, 2012, pp. 42-43). Even when a researcher wants to have participants' informed consent to take part in a study, it might be impossible for him or her to obtain it if the research in question concerns, for example, massive data mining processes and projects. Moreover, big data researchers often ignore the whole question of informed consent because they define their data as either public or proprietary (Paolillo, 2015, p. 49). Also, when there is no direct contact between the researchers and their human subjects it is questionable whether the subjects should even be called participants. Besides, when experiments are made on them, it is unclear whether they are to be subject to the same ethical research scrutiny as human-subject study participants normally are (Hutton & Henderson, 2015, p. 178; Kahn, Vayena, & Mastroianni, 2014, p. 13677.). Even if a researcher did in such cases manage to obtain the participants' consent, there would be no real guarantee that it was indeed informed (Flick, 2016, p. 15-17).

To problematize the issue further, even if informed consent was verified and the researcher was allowed to use the participants' personal data, the data might also include information about people (e.g. contacts of the users) who had not given their informed consent (Phillips, 2011, p. 32). Thus it is no surprise that a large number of extensive data mining projects are carried out without informing the groups or individuals targeted by the researchers; the only measure taken to promote the ethicality of the research is making sure that the participants are anonymous, thus ensuring confidentiality (Lindsay & Goldring, 2010; Zwitter, 2014, p. 5; see also Sormanen et al., 2016).

In contrast, when conducting qualitative research like virtual ethnography or, more specifically, participant observation, in smaller internet forums, obtaining the consent of participants is technically relatively easy. However, it is rarely done because of the possibility that knowing that they are being observed might cause participants to act differently from usual, which would skew the data. Then again, in practice, many scholars do not seek informed consent because they are afraid it would be denied (e.g. Hine, 2000, pp. 23-24.). Sometimes participant observation even without consent is impossible (e.g. in the case of private discussion groups), so the researcher might engage in deception (e.g. an invented alias) in order to gain access to the group of participants. As Brotsky and Giles (2007, pp. 95-96) put what is indeed rather obvious, covert participant observation is "highly controversial from an ethical position", but as in most completed research projects with ethical research challenges, it is ultimately justified by reference to the benefits brought by the results. Sometimes even informed consent does not create an authentic consensual atmosphere, for example if the subjects of the research do not feel they have been treated fairly or if the purpose of the research is not felt to be morally valuable enough (Kennedy, Elgesem, & Miguel, 2015). Lastly, even if informed consent is received, there is the problem of the level of informedness. How can a researcher be sure that the

[4] The author danah boyd wants her name to be written in lower case.

research subject has sufficiently understood the purpose and the consequences of the research? (E.g. Escobedo, Guerrero, Lujan, Ramirez, & Serrano, 2007; Svanteson, 2007, p. 72.)

3. The Facebook Experiment as Manipulation

On Facebook, the News Feed is practically a list of status updates of the contacts in a user's network. The updates shown in or omitted from the News Feed depend on "a ranking algorithm that Facebook continually develops and tests in the interest of showing viewers the content they will find most relevant and engaging". Facebook is thus like any traditional media as it provides content to its users selectively, but where it differs from the old media is that the content is modified individually according to what the medium evaluates to be the optimally engaging experience. (Kramer et al., 2014, p. 8788.) Users accept this practice when signing up for Facebook.

In their massive-scale experiment, Kramer et al. (2014) tested the emotional engagement of Facebook users by modifying their News Feed. The "experiment on the manipulative power of Facebook feeds", as Peacock (2014, p. 8) described it, was criticized almost immediately upon publication of the article. Bloggers claimed Facebook made users "sad for a psych experiment" (Grimmelmann, 2014) or the company was using people as "lab rats" (a blogger quoted by Rushe, 2014). According to *The Guardian*'s poll (Fishwick, 2014), most people who read about the experiment were not surprised that Facebook would experiment on user data the way it did but, at the same time, they declared they had now "lost trust" in Facebook and were considering closing their account. The "secret" experiment, as *The Guardian* called it, "sparked outrage from people who felt manipulated by the company". It can be speculated that had Facebook known what the public reaction to their experiment was going to be, they would not have published it. danah boyd (2014; see also Paolillo, 2015, p. 49) suggests that the intended PR outcome of the experiment from Facebook's point of view was to show that Facebook can downplay negative content in their service and thus make customers happier. Presumably this was seen as better for users and better for Facebook, as experimentation is how websites make their services better (Halavais, 2015, pp. 689-690; Kahn et al., 2014, p. 13677).

It is possible that many people missed the benevolent intention of the research team and concentrated on the contestable ethics of their method. The criticism about the experiment reached such levels that Facebook's researcher and the first author of the article, Adam Kramer, defended the experiment in his own Facebook page, pointing to the minimal "actual impact on people". During the week of the experiment, he explained, the users who were affected "produced an average of one fewer emotional word, per thousand words". (Kramer, 2014.) The magnitude of the impact was perhaps unknown to many critics of the experiment, as many objected to it on the grounds that Facebook was "controlling the emotions" of its users. Moreover, regardless of the magnitude of the impact of the experiment, the user agreement of Facebook can be interpreted to mean that users of Facebook allow researchers to experiment on them.

Thus, many ethicists would agree with Meyer (2014), who published a statement with five co-authors and on behalf of 27 other ethicists "to disagree with these sweeping condemnations" of Facebook's ethics in the experiment. She wrote that "the experiment was controversial, but it was not an egregious breach of either ethics or law." If Facebook is permitted to mine user data and study users for personal profit but academics are not permitted to use that information and learn from it, it "makes no one better off" (Meyer, 2014). However, for many critics it was more a matter of ethical principle than actual impact. For example, Kleinsman and Buckley (2015, p. 180) rejected Meyer's statement and claimed that "[i]f an experiment is in 'breach of either ethics or law,' then whether it is an 'egregious' breach or not is irrelevant." In this view, there is no grey area in research ethics, and consequently, a person as a subject of research is—in a binary way—either harmed or not harmed.

Many scholars were even more critical than Kleinsman and Buckley (2015). Recuber (2016), for example, noted how quick scholars were to draw analogies between the Facebook experiment and the infamous Milgram's (1963) experiment analyzing obedience to authority, as well as to the Stanford Prison experiment, also known as the Zimbardo experiment (Haney, Banks, & Zimbardo, 1973; Zimbardo, 1973), that studied the psychological effects of becoming a prisoner or a guard. According to Recuber, there were indeed some similarities between the Facebook experiment and the two notorious experiments from the 1960s, one being the fact that all three studied the researchers' ability to manipulate change in the participants' behavior. However, the Facebook experiment was different in its failure to reflect on this aspect (Recuber, 2016, pp. 46-47). The user reactions studied in the Facebook experiment were caused by the observers but the power relations between the experimenters and the experimentees were downplayed or normalized, and not at all problematized. This, at least to Recuber, is a typical and insidious element of contemporary big data research. (Recuber, 2016.) When the number of research subjects is so high, individually they tend to vanish in the haze of the overarching term "big data". However, the "power" exerted per capita over the participants in the Facebook experiment can be viewed as rather minimal (albeit massive in scale). The experiments carried out by Milgram and Zimbardo, on the other hand, caused

their participants to suffer severe physical and psychological stress.

The ethics of human-subject research is mainly about protecting the subject. In this sense, the Facebook experiment was found ethically questionable. Strict assessments of the experiment conclude that the study indeed "harmed" its participants (albeit almost unnoticeably), because it changed the participants' mood (e.g. Bryman & Bell, 2015, p. 141; Grimmelmann, 2014; Kleisman & Buckley, 2015, p. 181). However, if harming is defined as changing a participant's mood, then a vast quantity of empirical research on humans is harmful, especially research that requires face-to-face interaction. In general, big data studies or techniques to test or predict personality or actions might not be legally problematic but they do undermine a "sense of individuality on a personal level", claims Schroeder (2014, p. 7).

Facebook has experimented on its users before, and has published research about it (see e.g. Bond et al., 2012; Chan, 2015, p. 1081; Simonite, 2012). However, these experiments were explicit in their intention to influence users. For example, in 2010 on the day of the US congressional elections, Facebook encouraged randomly assigned users to vote, managed to increase voting activity, and afterwards published an article about it in *Nature* (Bond et al., 2012). Moreover, in 2012 Mark Zuckerberg, the CEO of Facebook, used Facebook to encourage people to register as organ donors, after which organ donor enrollment increased significantly in the US (Simonite, 2012). These forms of "manipulation" did not raise as much ethical debate as the experiment we discuss here did. The reason for this might be that people see explicit forms of intended manipulation as more acceptable than covert forms, even if the explicit manipulation attempts to elicit significantly greater change in the subject than the covert form.

Research ethics are often implemented more strictly in the academic world than in the corporate research environment. Then again, the ethical views of social media users might be quite flexible, and a lot of how users relate to being studied and experimented on by researchers depends on the application of the results (Kennedy et al. 2015, pp. 8-10). It seems like people do not want to be experimented on for the sake of an experiment but they are more likely to accept it if the experiment might result in some kind of benefit for themselves or others. Many people also do not mind commercials or other manipulations—even outright propaganda—as they are often part of the deal between users and service providers (cf. Searls 2015 on ad blockers). In the case of the Facebook experiment, even though scholars did not read any status updates, some people still felt that their privacy was violated. The problem in these kinds of cases is often the fact that one has a feeling of being private while actually being public (Kennedy et al., 2015, p. 13). According to Chan (2015, p. 1080), the fact that neither Facebook nor Cornell University—the two parties involved in conducting the study—apparently anticipated the public backlash they would face for the data manipulation shows "the vast disconnect between the research culture of big data (whether based in corporate or academic institutions) and the general public's cultural expectations."

4. The Problem of Informed Consent

It is the "informed" in informed consent that is the other major ethical research issue in the experiment that worried both the general public and academia (see e.g. Kahn et al., 2014, p. 13677). Cornell University researchers (Guillory and Hancock) analyzed the data after Facebook (Kramer) had collected them. The study therefore did not go through an ethical review at Cornell University, which might have been critical of how the informed consent of the participants was going to be secured (Paolillo, 2015, p. 50). In the article, research ethics is discussed in two sentences (Kramer et al., 2014, p. 8789). The first sentence states that the researchers themselves did not read any of the texts analyzed for the experiment as a linguistic software program was used to analyze the data. The other sentence declares that the data collection "was consistent with Facebook's Data Use Policy, to which all users agree prior to creating an account on Facebook, constituting informed consent for this research." In other words, the authors interpreted Facebook's user agreement to mean informed consent.

In that case, the level of informedness is highly debatable, as most users of Facebook do not read or completely understand the data use policy (Flick, 2016, p. 17; see also Kennedy et al., 2015, pp. 10-15). When a user accepts the terms and signs up for Facebook, he or she is informed that the service provider will use the personal data for all sorts of things (Facebook, 2015a). The user might give their consent but is most likely not well informed, since the description of the data use policy is not very precise (see e.g. Grady, 2015, p. 885; Sloan, Morgan, Burnap, & Williams, 2014, p. 16.). For example, at the time of the experiment, the research use of personal data was not mentioned although, following the wide publicity the experiment received, it has subsequently been added to the policy.

Kleisman and Buckley (2015; see also Bail, 2015, p. 23) hold the view that because the authors of the Facebook experiment could have asked for proper informed consent from the users, they should have done so. It does not matter whether the research is unlikely to cause harm or if it is beneficial or otherwise important: consent is always essential if it can be obtained. The scholars should at least have informed those users who were affected afterwards (Recuber, 2016, p. 54; see also McKelvey, Tiessen, & Simcoe, 2015, pp. 580-581). A month after the publication of

the experiment, *PNAS*'s Editor-in-Chief, Inder M. Verma (2014), added a foreword to the contested article. It was entitled "Editorial Expression of Concern and Correction" and it defended the authors' ethical choices by separating Facebook's data collection process from the actions of Cornell University. Readers were reminded that it was a non-academic private company (= Facebook's Kramer) that gathered the data, and the academics (= Cornell's Guillory and Hancock) only analyzed them. However, as the responsibility fell partly on the journal (see e.g. Kahn et al., 2014, p. 13679), Verma (2014, p. 10779; see also Schroeder, 2014, pp. 2-3) did concede that perhaps everything was "not fully consistent with the principles of obtaining informed consent and allowing people to opt out."

Many human-subject big data scientists know that a strict interpretation of the opting-out option makes their research extremely difficult. The problem is further complicated by the fact that in many cases the data are not collected by academics but by third parties such as Facebook. Should the data collectors abide by the ethical research norms of academia? If they did, there would be a lot of ethical problems, particularly with data produced by third parties, such as filmed footage, photographs, Google Street View data, televised rock concert recordings, and so on. Even if participant anonymity was secured, the human subjects in these cases could not opt out. If scholars did not have to worry about opting out as an ethical norm, they could team up with someone outside of academia to do their "dirty work" (see e.g. Kahn et al., 2014, p. 13677; Wrzus & Mehl, 2015, p. 264; cf. boyd, 2014.). On the other hand, one could say that a person can opt out of any Facebook experiment by not signing up for Facebook in the first place—just like a potential participant in a psychology experiment can decide not to attend the experiment if he or she does not want to be manipulated.

In general, an ethically pragmatic social media user's informed consent is more like meta-informedness, or "implicit informed consent" (Bryman & Bell, 2015, p. 139), where the user knows that for example Facebook will do various known and unknown things with its user data but is unlikely to do anything that is morally too dubious—although it has been observed that users tend to underestimate the level of their privacy when they are excited about a social media application (Kehr, Kowatsch, Wentzel, & Fleisch, 2015). For most users, Facebook's data policy is thus a reasonably informed and fair trade-off between the user who gets to use the service without a fee, and the service provider who gets to sell the data to third parties such as advertisers (Kennedy et al., 2015, p. 12; see also Hutton & Henderson, 2015, p. 178). This is actually the common logic of commercial media, and the "ethical fig leaf" (O'Hara, Nguyen, & Haynes, 2014, p. 4) of a social media researcher.

As Chan (2015, p. 1080; see also Aiken & Mahon, 2014, p. 4) notes, Facebook's data use policy "enables any user to potentially become an experiment subject without need for prior consent". In the end, a scholar interested in research ethics might ask if there is anything ethically new in the Facebook experiment. People were studied without their knowing about it but they had allowed it by signing up for Facebook. (Schroeder, 2014, p. 3; see also Zwitter, 2014, p. 1.) Certainly companies have been doing experiments with only vaguely informed consent before, as have psychologists, so many people think the Facebook experiment is merely a recent example of an old ethical research issue (Schroeder, 2014, pp. 1-2; cf. Selinger & Hartzhog, 2016).

In a way, the Facebook user agreement is similar to the informed consent form the participants in most psychological experiments have to fill out. Participants are informed that they (and the data they will produce) will be used for scientific purposes but the participant might not know exactly what those purposes are. He or she might even be deceived about the real purpose of the study to which they have consented. The message of informed consent is, "I trust you. Do what is needed." Perhaps the only new aspect in this case is that there are over a billion people on Facebook every day. It is an essential networking tool for a large amount of people, many of whom are dependent (to a greater or lesser extent) on the service. This means that its user agreement is not necessarily an ethical act between two equal parties: opting out of an experiment becomes equal to opting out of a significant part of one's social life (see e.g. Gertz, 2016). One might therefore suggest that a participant might be sufficiently informed but the question of consent is more controversial.

After multiple critical reviews of the experiment, Mike Schroepfer, the Chief Technology Officer for Facebook, wrote an apologetic post for Facebook's Newsroom. According to him, they should have "considered non-experimental ways" to do the research. Also, the research would have "benefited from more extensive review by a wider and more senior group of people". Schroepher also notes that they did not inform the public about the experiment well enough (Schroepfer, 2014). Schroepher also introduced a new framework of research that Facebook is going to implement. It included clearer guidelines for researchers, a more extensive review stage, and training (including on privacy and security matters), as well as the establishment of a special research website (Facebook, 2015b).

Describing the new guidelines section, Schroepher announced that a more enhanced review process would be conducted prior to research if the intended research focused on "studying particular groups or populations (such as people of a certain age) or if it related to content that may be considered deeply personal (such as emotions)." Also, a further review would be conducted if there was any collaboration with the academic community. The statement ends with trying to convince the reader—supposedly a daily Facebook

user—that Facebook wants to do research "in a way that honors the trust you put in us by using Facebook every day." (Schroepher, 2014.) This seems to be Facebook's way of admitting that the experiment lacked informed consent. Perhaps for PR reasons as well as due to potential legal issues, Schroepher could not say outright that the experiment failed to obtain informed consent (cf. Verma, 2014).

5. Discussion

In this article we have shown how the debate around the Facebook experiment brings up two crucial and interrelated themes of research ethics: research as manipulation, and the problem of informed consent. The debate around the experiment shows that the era of big data research demands some rethinking of research ethics. Although the two key issues presented here are not unique to contemporary research but had been debated for decades before big data research came in (see e.g. Faden & Beauchamp, 1986; Roelcke, 2004), the unprecedentedly large amount of human subjects that are called for in such research has led to a need for special scrutiny. At the same time, it seems that the ethical evaluation of such experiments is based on the classical ethical stances of utilitarianism or deontology. The proponent of the former sees little or no harm done in such an experiment and no loss of happiness caused by it, while the proponent of the latter considers that, regardless of the degree of actual harm, human integrity has been violated (see e.g. Ess, 2013, pp. 256-262; Harman & Cornelius, 2015, p. 58; Shrader-Frechette, 2000).

Reaching any ethical consensus about the Facebook experiment is further impeded by disagreements over the definition of key concepts such as the "harm" done to human subjects, and their "informed consent". When academic research ethics is so vague, it might seem simpler for scholars to leave it to the law and user agreements to define the ethics of the research. However, according to Chan (2015, p. 1082; see also Paolillo, 2015, p. 50; Burgess & Bruns, 2015, p. 99), commercial companies' ethical research standards should not be allowed to spread to the academic world. Flick (2016; see also Halavais, 2015, p. 592) agrees and thinks that the commercial and academic sectors should negotiate and agree on standards, but without making any concessions in the commercial companies' favor. However, as universities' opportunities to cooperate with private companies working with big data increase, the opportunities to leave the problematic ethics of data collection to companies increase likewise.

Mike Schroepher, the Chief Technology Officer of Facebook, stated that Facebook should have communicated "clearly why and how" they did the experiment (Schroepfer, 2014). The statement implies that a person is deprived of optimal well-being if the reasons and methods of any actions carried out on him or her are not properly communicated. On the other hand, one could easily claim the opposite: a person suffers less when he or she does not know or notice anything about such actions. As Stilgoe (2015, pp. 46-47) observes, the Facebook experiment was rare in being openly published and publicly scrutinized, since most such experiments are conducted in secret. We can wonder if people were outraged about the experiment because Facebook altered its users' states of mind or because it reminded them that their states of mind are being altered all the time by all kinds of things, people and organizations (see e.g. boyd, 2016; see also Kehr et al., 2015).

At the same time, Kennedy et al. (2015, p. 2) observe that there has been little research about what social media users themselves actually think about being observed, studied and—we would add—experimented on. This is rather disconcerting, given the massive number of people that use social media and are in some form or other observed and experimented on by researchers. Perhaps surprisingly, the social media users Kennedy et al. (2015, pp. 3-4) studied seemed to be concerned about privacy, but mainly about social privacy. That is, they wanted to be sure that they could choose which individuals in their network have access to their personal information. They were not so worried about institutional privacy, or "the mining of personal information by social media platforms, commercial companies and governments". Although we are talking about only one study, there is reason to suggest that the ethical criticism of the Facebook experiment made by academics might not reflect users' worries. This is a topic that should be further studied, as it would be relevant for research ethics in the era of social media to be more grounded in the user level. A more holistic and inclusive ethical research study would ensure that researchers do more than define what is morally optimal in big data research; or, as Tama Leaver (2013) states, "Big Data needs Big Ethics, and we don't have them yet."

If we go further into the ethical implications of social media experiments that aim to enhance user experience, we are faced with a more profound ethical challenge than a discussion of manipulation and informed consent reveals. If in Facebook we are fed imagery that further filters our experiences of the "real" world, then what are the ethical ramifications of researchers teaming up with companies that aim to give people "the experience they want" (Simonite, 2012)? Would the companies be in charge of the "hard ethical choice…of what content to show…without oversight, transparency, or informed consent" (boyd, 2014)? The way media and new media influence our perceptions of reality has already been widely studied (e.g. Fairclough, 1995; Macey, Ryan, & Springer, 2014) but there has been little consideration so far of the ethics of academics taking part in these kinds of studies.

The way big data is "all at once essential, valuable, difficult to control, and ubiquitous" seems to be reflected in our complex, context-dependent attitudes toward it (Pushcmann & Burgess, 2014, p. 1695). Gertz (2016, p. 56) notes that despite the Facebook controversy, the number of Facebook users is still growing. At the same time, users' autonomy seems to be diminishing. From this it can be concluded that many users do not mind the asymmetrical relationship they have with the service provider. As Ess (2013, p. 254) notes, "our engagements with new digital media appear to bring in their wake important transformations in our sense of self and identity." Our "foundational conception of autonomous self" that has legitimated concepts of privacy that "modern liberal-democratic" states respect seems to be changing. Perhaps the question we should ask is primarily existential rather than ethical, as Gertz (2016, p. 61) suggests. According to him, we should first think about the increasingly significant role technology plays in our lives. If we accept it, then we can have a more meaningful discussion on the ethics of scholars experimenting with it.

Conflict of Interests

The authors declare no conflict of interests.

References

Aiken, M., & McMahon, C. (2014). A primer of research in mediated environments: Reflections on cybermethodology. *SSRN Working Papers*. Retrieved from http://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2462700

Ananny, M. (2015). Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Science, Technology, and Human Values*, *4*(1), 93-117.

Bail, C. A. (2015). Taming big data: Using app technology to study organizational behavior on social media. *Sociological Methods & Research*. Retrieved from http://smr.sagepub.com/content/early/2015/05/15/0049124115587825.full

Barsade, S. (2002). The ripple effect. Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47*(4), 644-675.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295-298.

boyd, d. (2010, April). *Publicity and privacy in web 2.0.* Keynote speech at WWW2010, Raleigh, USA. Retrieved from www.danah.org/papers/talks/2010/SXSW2010.html

boyd, d. (2014). What does the Facebook experiment teach us? Growing anxiety about data manipulation. *The Message*. Retrieved from https://medium.com/message/what-does-the-facebook-experiment-teach-us-c858c08e287f

boyd, d. (2016). Untangling research and practice: What Facebook's "emotional contagion" study teaches us. *Research Ethics*, *12*(1), 4-13.

The British Psychological Society (2010). *Code of human research ethics.* Leicester: The British Psychological Society. Retrieved from www.bps.org.uk/sites/default/files/documents/code_of_human_research_ethics.pdf

Brotsky, S. R., & Giles, D. (2007). Inside the "Pro-ana" community: A covert online participant observation. *Eating Disorders: The Journal of Treatment & Prevention*, *15*(2), 93-109.

Bryman, A., & Bell, E. (2015). *Business research methods*. Oxford: Oxford University Press.

Burgess, J., & Bruns, A. (2015). Easy data, hard data: The politics and pragmatics of Twitter research after the computational turn. In G. Langlois, J. Redden, & G. Elmer (Eds.), *Compromised data: From social media to big data* (pp. 93-111). New York: Bloomsbury Academic.

Card, N. A. (2010). Literature review. In N. J. Salkind (Ed.), *Encyclopedia of research design* (pp. 726-729). Thousand Oaks: Sage.

Ceserani, R. (2010). The essayistic style of Walter Benjamin. *Primerjalna književnost*, *33*(1), 83-92.

Chan, A. (2015). Big data interfaces and the problem of inclusion. *Media, Culture & Society*, *37*(7), 1078-1083.

Cheng, J., Adamic, L. A., Dow, P. A., Kleinberg, J., & Leskovec, J. (2014). Can cascades be predicted? *Proceedings of the 23rd International Conference on World Wide Web* (pp. 925-936). New York: ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2567997

Copeland, L. (2011). The anti-social network. *Slate.* Retrieved from www.slate.com/articles/double_x/doublex/2011/01/the_antisocial_network.html

Cornelissen, J., Gajewskade, M., Piekkari, R., & Welch, C. (2012). Writing up as a legitimacy seeking process: Alternative publishing recipes for qualitative research. In S. Gillian & C. Cassel (Eds.), *Qualitative organizational research: Core methods and current challenges* (pp. 185-203). London: Sage.

Escobedo, C., Guerrero, J., Lujan, G., Ramirez, A., & Serrano, D. (2007). Ethical issues with informed consent. *Bio-Ethics*, *1*, 1-8. Retrieved from http://cstep.cs.utep.edu/research/ezine/Ezine-ethicalIssueswithInformedConsent.pdf

Ess, C. M. (2007). Internet research ethics. In A. Joinson (Ed.), *Oxford handbook of internet psychology* (pp. 481-502). Oxford: Oxford University Press.

Ess, C. M. (2013). Global media ethics? Issues, requirements, challenges, resolutions. In S. J. A. Ward (Ed.), *Global media ethics: Problems and perspectives* (pp. 253-271). West Sussex: Wiley-Blackwell.

Facebook (2015a). *Terms of service.* Retrieved from www.facebook.com/terms.php

Facebook (2015b). *Research at Facebook.* Retrieved from https://research.facebook.com

Faden, R. R., & Beauchamp, T. L. (1986). *A History and theory of informed consent.* Oxford: Oxford University Press.

Fairclough, N. 1995. *Media discourse.* London: Bloomsbury.

Ferrara, E., & Yang, Z. (2015). Measuring emotional contagion in social media. *PloS ONE*, *10*(11). Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0142390

Fishwick, C. (2014, June 30). Facebook's secret mood experiment: Have you lost trust in the social network? *The Guardian.* Retrieved from www.theguardian.com/technology/poll/2014/jun/30/facebook-secret-mood-experiment-social-network

Flick, C. (2016). Informed consent and the Facebook emotional manipulation study. *Research Ethics*, *12*(1), 14-28.

Friggeri, A., Adamic, L. A., Eckles, D., & Cheng, J. (2014). Rumor cascades. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media (ICWSM)* (pp. 101-110). Palo Alto, CA: AAAI Press. Retrieved from www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8122

Gertz, N. (2016). Autonomy online: Jacques Ellul and the Facebook emotional manipulation study. *Research Ethics*, *12*(1), 55-61.

Grady, C. (2015). Enduring and emerging challenges of informed consent. *The New England Journal of Medicine*, *372*(9), 855-862.

Grimmelmann, J. (2014). As flies to wanton boys. *The Laboratorium.* Retrieved from http://laboratorium.net/archive/2014/06/28/as_flies_to_wanton_boys

Halavais, A. (2015). Bigger sociological imaginations: Framing big social data theory and methods. *Information*, *Communication & Society*, *18*(5), 583-594.

Haney, C., Banks, C., & Zimbardo, P. (1973). A study of prisoners and guards in a simulated prison. *Naval Research Reviews*, *30*(9), 4-17.

Harman, L. B., & Cornelius, F. (2015). *Ethical health informatics.* Burlington, MA: Jones & Bartlett.

Harriman, S., & Patel, J. (2014). The ethics and editorial challenges of internet-based research. *MBC Medicine*, *12*, 124-127.

Hesse, B. W., Moser, R. P., & Riley, W. T. (2015). From big data to knowledge in the social sciences. *The Annals of American Academy of Political and Social Science*, *659*(1), 16-32.

Hine, C. (2000). *Virtual ethnography.* London: Sage.

Hudson, J. M., & Bruckman, A. (2004). Go away: Participant objections to being studied and the ethics of chatroom research. *The Information Society*, *20*(2), 127-139.

Huntsinger, J. R., Lun, J., Sinclair, S., & Clore, G. L. (2009). Contagion without contact: Anticipatory mood matching in response to affiliative motivation. *Personality and Social Psychology Bulletin*, *35*(7), 909-922.

Hutton, L., & Henderson, T. (2015). "I didn't sign up for this!": Informed consent in social network research. *Proceedings of the Ninth International AAAI Conference on Web and Social Media (ICWSM)* (pp. 178-187). Palo Alto, CA: AAAI Publications. Retrieved from www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10493

Jouhki, J., Lauk, E., Penttinen, M., Rohila, J., Sormanen, N., & Uskali, T. (2015, November). *Social media personhood as a challenge to research ethics: Exploring the case of the Facebook experiment.* Paper presented at the Social Media Research Symposium, Jyväskylä, Finland.

Kahn, J. P., Vayena, E., & Mastroianni, A. C. (2014). Opinion: Learning as we go: Lessons from the publication of Facebook's social-computing research. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(38), 13677-13679.

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal*, *25*, 607-635.

Kennedy, H., Elgesem, D., & Miguel, C. (2015). On fairness: User perspectives on social media data mining. *Convergence: The International Journal of Research into New Media Technologies.* Retrieved from http://con.sagepub.com/content/early/2015/06/26/1354856515592507.full

Kleinsman, J., & Buckley, S. (2015). Facebook study. A little bit unethical but worth it? *Journal of Bioethical Inquiry*, *12*(2), 179-182.

Kramer, A. (2014). A post on 29.6.2014. *Facebook.* Retrieved from www.facebook.com/akramer/posts/10152987150867796

Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(24), 8788-8790.

Leaver, T. (2013, September). *Birth, death and Facebook.* Paper presented at Adventures in Culture in Technology (ACAT) Seminar Series, Perth. Retrieved from www.tamaleaver.net/2013/10/03/birth-death-and-facebook

Lindsay, S., & Goldring, J. (2010). Anonymizing data for secondary use. In A. J. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of case study research* (pp. 25-27). London: Sage.

Macey, D. A., Ryan, K. M., & Springer, N. J. (Eds.). (2014). *How television shapes our worldview: Media representations of social trends and change.* New York: Lexington Books.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making and Internet research. Recommendations from the AoIR ethics working committee (version 2.0).* Chicago: Association of Internet Researchers. Retrieved from http://aoir.org/reports/ethics2.pdf

Marx, G. T. (2013). An ethics for the new (and old) surveillance. In F. Flammini, R. Stola, & G. Franceschetti (Eds.), *Effective surveillance for homeland security: Balancing technology and social issues* (pp. 2-20). Boca Raton, FL: Taylor & Francis.

McKelvey, F., Tiessen, M., & Simcoe, L. (2015). A consensual hallucination no more? The Internet as simulation machine. *European Journal of Cultural Studies, 18*(4-5), 577-594.

Meyer, M. N. (2014). Misjudgments will drive social trials underground. *Nature*, *511*(7509), 265.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, *67*(4), 371-378.

O'Hara, K., Nguyen, M.-H. C., & Haynes, P. (2014). Introduction. In K. O'Hara, M.-H. C. Nguyen, & P. Haynes (Eds.), *Digital enlightenment yearbook 2014: Social networks and social machines, surveillance and empowerment* (pp. 3-24). Amsterdam: IOS Press.

Paolillo, J. C. (2015). Network analysis. In A. Georgakopoulou & T. Spilioti (Eds.), *The Routledge handbook of language and digital communication* (pp. 36-54). London: Routledge.

Parkinson, B., & Manstead, A. S. R. (2015). Current emotion research in social psychology: Thinking about emotions and other people. *Emotion Review*, *7*(4), 371-380.

Peacock, S. E. (2014). How web tracking changes user agency in the age of Big Data: The used user. *Big Data & Society*, *1*(2), 1-11. Retrieved from http://bds.sagepub.com/content/1/2/2053951714564228

Pejovic, V., & Musolesi, M. (2015). Anticipatory mobile computing. A survey of the state of the art and research challenges. *ACM Computing Surveys*, *47*(3), 1-29.

Phillips, M. L. (2011). Using social media in your research. Experts explore the practicalities of observing human behavior through Facebook and Twitter. *gradPSYCH*, *9*(4), 32.

Recuber, T. (2016). From obedience to contagion: Discourses of power in Milgram, Zimbardo, and the Facebook experiment. *Research Ethics*, *12*(1), 44-54.

Roelcke, V. (2004). Introduction: Historical perspectives on human subjects research during the 20th century, and some implications for present day issues in bioethics. In V. Roelcke & G. Maio (Eds.), *Twentieth century ethics of human subjects research: Historical perspectives on values, practices, and regulations* (pp. 11-18). Stuttgart: Franz Steiner Verlag.

Rooke, B. (2013). Four pillars of internet research ethics with Web 2.0. *Journal of Academic Ethics*, *11*(4), 265-268.

Rosenberg, Å. (2010). Virtual world research ethics and the private/public distinction. *International Journal of Internet Research Ethics*, *3*(1), 23-37.

Rushe, D. (2014, October 2) Facebook sorry—almost—for secret psychological experiment on users. *The Guardian.* Retrieved from www.theguardian.com/technology/2014/oct/02/facebook-sorry-secret-psychological-experiment-users

Schroeder, R. (2014). Big Data and the brave new world of social media research. *Big Data & Society*, *1*(2), 1-11.

Searls, D. (2015, November 6). Ad blockers and the next chapter of the Internet. *Harvard Business Review*. Retrieved from https://hbr.org/2015/11/ad-blockers-and-the-next-chapter-of-the-internet

Selinger, E., & Hartzog, W. (2016). Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control. *Research Ethics*, *12*(1), 35-43.

Shah, D. V., Cappella, J. N., & Neuman, W. R. (2015). Big data, digital media, and computational social science: Possibilities and perils. *The Annals of the American Academy of Political and Social Science*, *659*(1), 6-13.

Shrader-Frechette, K. (2000). *Ethics of scientific research.* London: Rowman & Littlefield.

Simon, J. R. (2014). Corporate research ethics: Whose responsibility? *Annals of Internal Medicine*, *161*(12), 917-918.

Simonite, T. (2012, June 13). What Facebook knows. *MIT Technology Review.* Retrieved from www.technologyreview.com/s/428150/what-facebook-knows

Sloan, L., Morgan, J., Burnap, P., & Williams, M. (2014). Who tweets? Deriving the demographic characteristics of age, occupation and social class from Twitter user meta-data. *LloS ONE*, 10(3), 1-20.

Sormanen, N., Rohila J., Lauk, E., Uskali T., Jouhki J., & Penttinen M. (2016). Chances and challenges of computational data gathering and analysis: The case of issue-attention cycles on Facebook. *Digital Journalism*, *4*(1), 55-74.

Stilgoe, J. (2015). *Experiment earth: Responsible innovation in Geoengineering.* London: Routledge.

Summers-Effler, E., Van Ness, J., & Hausmann, C. (2015). Peeking in the black box: Studying, theorizing, and representing the micro-foundations of day-to-day interactions. *Journal of Contemporary Ethnography*, *44*(4), 450-479.

Sun, E., Rosenn, I., Marlow, C., & Lento, T. (2009). Gesundheit! Modeling contagion through Facebook news feed. *Proceedings of the Third International ICWSM Conference* (pp. 146-153). Palo Alto, CA: AAAI Publications. Retrieved from http://aaai.org/ocs/index.php/ICWSM/09/paper/view/185

Svanteson, D. J. B. (2007). *Private international law and the Internet.* Alphen aan den Rijn: Kluwer Law International.

Thorson, K., & Wells, C. (2015). Curated flows: A frame-

work for mapping media exposure in the digital age. *Communication Theory*. Retrieved from http://ssc.sagepub.com/content/early/2015/10/19/0894439315609528.refs

Tinati, R., Halford, S., Carr, L., & Pope, C. 2014. Big data: Methodological challenges and approaches for sociological analysis. *Sociology*, *48*(4), 663-681.

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, *4*(3), 356-367.

Vainio, A. (2012). Beyond research ethics: Anonymity as 'ontology', 'analysis' and 'independence'. *Qualitative Research*, *13*(6), 685-698.

Vanderpool, H. Y. (1996). *The ethics of research involving human subjects: Facing the 21st century.* Frederick: University Publishing Group.

Verma, I. M. (2014). Editorial expression of concern and correction. *Proceedings of the National Academy of Sciences*, *111*(29), 10779.

Weeden, M. R. (2012). Ethics and on-line research methodology. *Journal of Social Work Values and Ethics*, *9*(1), 40-51.

Wrzus, C., & Mehl, M. R. (2015). Lab and/or field? Measuring personality processes and their social consequences. *European Journal of Personality*, *29*(2), 250-271.

Zimbardo, P. (1973). On the ethics of intervention in human psychological research: With special reference to the Stanford prison experiment. *Cognition*, *2*(2), 243-256.

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313-325.

Zwitter, A. (2014). Big Data ethics. *Big Data & Society*, *1*(2), 1-6.

About the Authors

Jukka Jouhki (PhD, Docent) is a Cultural Anthropologist working as a Senior Lecturer of Ethnology at the Department of History and Ethnology, and a member of the Social Media Research Institute at University of Jyväskylä, Finland. Jouhki's research interests include democracy, nationalism, imagined communities, online gambling, old and new media, as well as various issues in human-technology relations, and cultural phenomena related to them.

Epp Lauk (PhD) is Professor of Journalism and Head of the Department of Communication at the University of Jyväskylä, Finland. Her research and publications focus on journalism cultures and history, media and journalism in Central and East European countries, media self-regulation and innovations in journalism.

Maija Penttinen is an undergraduate student at the Department of History and Ethnology, University of Jyväskylä. Her research interests include both political participation and civic action on social media, in addition to studying the integration of social networking sites as platforms for everyday activities. She is currently working on her Master's thesis on everyday activities and experiences manifested in research literature concerning the social networking site Facebook.

Niina Sormanen (MA) is a PhD candidate of Organizational Communication and Public Relation (PR) at the University of Jyväskylä, Department of Communication. Her research interests include communicative behavior and power relations in the social media context. Her PhD thesis is focused on the interplay of organizational and media professionals and individuals in the social media context and uses of social media in building their communicative power.

Turo Uskali (PhD) is the Head of Journalism and Senior Research Scholar at the Department of Communication, University of Jyväskylä, Finland. He leads several research projects focusing on innovations in journalism. The most recent ones focus on mobile data journalism, and wearables. Uskali is also an Associate Professor at the University of Bergen, Norway, and he has authored or co-authored seven books about the evolution of global journalism and the changes in media industries.

Media and Communication (ISSN: 2183-2439)
2016, Volume 4, Issue 4, Pages 86-96
doi: 10.17645/mac.v4i4.576

Article

Hazy Boundaries: Virtual Communities and Research Ethics

Helena Kantanen [1,*] and Jyri Manninen [2]

[1] Business School, University of Eastern Finland, 70211 Kuopio, Finland; E-Mail: helena.kantanen@uef.fi
[2] School of Educational Sciences and Psychology, University of Eastern Finland, 80101 Joensuu, Finland; E-Mail: jyri.manninen@uef.fi

* Corresponding author

Submitted: 30 January 2016 | Accepted: 19 April 2016 | Published: 10 October 2016

Abstract
This paper examines ethical issues specific to research into virtual communities. Drawing on an empirical case with online forums of education experts, we identify the following key issues: publicity versus privacy of the community; the definition of human subjects research; participant recruitment; informed consent; and ethical questions associated with observing virtual communities, and with reporting and disseminating research results. We maintain that different research cultures in different countries can present challenges when studying global forums. Acknowledging the ephemeral characteristics of Internet contexts, this paper argues that ethical considerations should be more case-based, instead of relying on one model for all solutions. We suggest that local ethics committees or institutional review boards could, with their expert knowledge of ethics, provide valuable support for researchers operating in the complex and dynamic terrain of Internet research, as well as in fields and research settings where an ethical review is not a standard part of the research process.

Keywords
digital ethics; ethics; ethics committee; institutional review board; Internet research ethics; netnography; online communities; virtual communities

Issue
This article is part of the issue "Successes and Failures in Studying Social Media: Issues of Methods and Ethics", edited by Epp Lauk and Niina Sormanen (University of Jyväskylä, Finland).

© 2016 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

The ever-increasing popularity of social media platforms has made them alluring from the point of view of academic research. Researchers now have a unique and cost-efficient opportunity to examine and understand behaviours and beliefs in naturalistic contexts, and to reach populations which may be hard to reach otherwise (Madge, 2007; Moreno, Goniou, & Moreno, 2013). At the same time, the easy availability of research data made possible by social media raises new ethical questions such as what is public and what is private. Moreover, the prevailing ethical guidelines may not be applicable in all research settings. Therefore, the basic principles of ethical research conduct associated with respecting the autonomy of research subjects, avoiding harm, and protecting privacy and data, are extremely topical, and more discussion and understanding is needed about their implications in the context of relatively new electronic environments (Bruckman, 2006; Dennen, 2012; Finnish Advisory Board on Research Integrity, 2009; Hine, 2000; Kozinets, 2010; Markham & Buchanan, 2012; Zimmer, 2010).

This paper, written from the Finnish perspective, focuses on ethical issues related to the study of discussions on Internet forums—or, more specifically, virtual communities. Virtual communities, or online communities, are Internet-based communication forums or social networks where interaction is based on computer–mediated communication (Lakkala, 2010; Manninen & Nevgi, 2000). In these communities, participants typically have a common interest, like product develop-

ment or the use of social media tools at work (Kosonen, 2009; Wasko & Faraj, 2005). These communities are also characterised by "public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace" (Rheingold, 1993, p. 193).

During our ongoing research project (Kantanen, Manninen & Kontkanen, 2014), we have explored learning, communication and innovation generation, in both a national and global online community. During the different phases of our project, we have tackled several ethical issues. The questions we have asked have included, for example: Do we need an official ethical review when studying a global LinkedIn community, administered from the United States? Are we studying a public or private forum? Are community postings exempt? Does our study qualify as human subjects research? What would be an appropriate way to recruit research participants? How would we gain informed consent, and is it always necessary? How should we behave in an online forum when observing participants? How can we safeguard the confidentiality of participants' contributions when reporting the study?

Ethical questions are complicated in the real world but they are even more complex in the virtual environment. Such questions are related to the concepts of public and private, confidentiality, the integrity of data, reputational risks, intellectual property issues, and whether the research qualifies as human subjects research or not. Other issues include participant recruitment, disclosure of presence, and citing, anonymising and crediting when reporting and disseminating research results (Bassett & O'Riordan, 2002; Bruckman, 2006; Buchanan & Zimmer, 2013; Hine, 2000; Kozinets, 2010; Markham & Buchanan, 2012; Turtiainen & Östman, 2013; Walther, 2002).

The remainder of this paper is organised as follows: Section 2 clarifies different ethical questions concerning Internet research. Section 3 describes our study, the empirical case, and the ethical issues involved with it. Section 4 discusses these issues, and our conclusions follow in Section 5.

2. Internet Research Ethics

The Finnish Advisory Board on Research Integrity (2009) defines the ethical principles of research in the humanities and social and behavioural sciences in three regards: (1) respecting the autonomy of research subjects; (2) avoiding harm; and (3) privacy and data protection. Respecting the autonomy of research subjects means, before all, voluntary participation based on informed consent. Harm can be mental, financial or social, and can occur during the collecting of data, retaining of data, or during the publishing of research results. Privacy and data protection involve issues related to protecting research data and confidentiality, storing and destroying data, and publishing research results. The Advisory Board guidelines emphasise a balance between confidentiality and the openness of research.

Rosenberg (2010) shows how boundary work has challenged Internet researchers and revitalised the discussion about research ethics. For example, boundaries between private and public are often blurred. Rosenberg summarises the general principles of ethical matters: researchers should maximise benefits and minimise harm, people must be treated fairly and equally, research subjects should be treated as autonomous individuals, and those with diminished autonomy must be entitled to protection (Rosenberg, 2010).

One of the key issues in Internet research is whether to use the Internet as a practical research tool (e.g., online questionnaires, Internet–mediated research or online research practice), as a research data source (i.e., studies in the web), or as a research object (i.e., studies about the web) (Turtiainen & Östman, 2013). Turtiainen and Östman (2013) remind us that, often, all of these elements create a research environment that requires, for example, awareness of source criticisms, knowledge of cultural practices associated with the specific research environment, knowledge about how the data has been created, and an understanding of how research subjects perceive the publicity or privacy of their web presence.

Specific ethical questions related to virtual environments have become topical since 2000 (Turtiainen & Östman, 2013). Madge (2007) identifies five key issues discussed in the literature on online research ethics: informed consent, confidentiality, privacy, debriefing and netiquette. All of these issues will be discussed later from the point of view of our study. By netiquette, Madge refers to the often flexible codes of conduct or guidelines applied in Internet communication, including phenomena like flaming and online harassment. Tavani (2006) discusses ethical issues related to the increasing use of cybertechnology, which refers to a wide range of computing and communication systems. He calls for cyberethics as "a field of applied ethics that examines moral, legal, and social issues involving cybertechnology" (p. 19). According to Tavani, professional codes of conduct can often help to resolve professional ethics issues, which is also what the different ethical guidelines are meant for. To our knowledge, the most advanced recommendations and guidelines concerning Internet research come from the Association of Internet Researchers (AoIR) and its Ethics Working Committee. They outline general principles that are meant to guide decision-making in Internet research, regardless of rapidly changing technological contexts (Markham & Buchanan, 2012). Markham and Buchanan call for guidelines rather than a code of practice to ensure the flexibility of research and the usefulness of those guidelines in different research contexts. The major issues or considerations identified in the AoIR guidelines are: (1) questions of human subjects research; (2) publicity and privacy; (3) are we studying

text or persons; and (4) top–down versus bottom–up approaches to ethics (Markham & Buchanan, 2012). Regarding the last tension, Markham and Buchanan refer to the need for researchers to balance between contextual, case–based requirements, and disciplinary, institutional, legal and cultural constraints.

McKee and Porter (2009a) developed an inquiry strategy to guide decisions about research ethics. They "believe that ethical decision-making for research must be systematic, deliberative, collaborative, and multidisciplinary in order to be valid" (p. 7), and offer a framework for those criteria. According to them, researchers must recognise the special circumstances of each case and be able to situate their case in proximity to other parallel cases and to community expectations in particular (McKee & Porter, 2009a). Instead of codes of practice or guidelines, they emphasise case-based processes that would help researchers with their decision-making in the ever-changing Internet environment (McKee & Porter, 2009b).

Ethnographic research has an interest in cultures and cultural meanings with an emphasis on the insider view, as well as in language and rhetoric (Eriksson & Kovalainen, 2016). Robert V. Kozinets coined the concept of *netnography*, which refers to ethnography conducted online. According to Kozinets (2010, p. 60): "Netnography is participant-observational research based in online fieldwork. It uses computer-mediated communications as a source of data to arrive at the ethnographic understanding and representation of a culture or communal phenomenon."

Kozinets (2010) maintains that "pure" ethnography studies communities or cultures, without important online elements, through face-to-face interaction and data collection, while "pure" netnography does the same, without important in-person elements, through entirely online interaction and data collection. Therefore, Kozinets challenges another pioneer of virtual ethnography, Christine Hine (2000, 2005), who maintains that online ethnography is always partial because the online experience is only one aspect of the social experience, and also because the culture or community studied does not have the field site needed for a holistic description of the culture.

The nature of our study in relation to netnography is examined further in the following section.

3. The Empirical Case and Its Ethical Challenges

In our project, our aim is to analyse learning, communication and innovation generation in virtual communities. Our data is derived from discussion threads published in online forums, as explained in Table 1.

In our pilot study (Kantanen et al., 2014), we examined a Finnish web-based community intended for those interested in using social media tools in teaching and learning (*Sometu*; http://sometu.ning.com). As of September 2013, that community, *Sometu*, had 4,828 members. The research question in our pilot study was: *What are the prerequisites of learning and innovation development in virtual communities of practice used by professionals?* In the more recent study (Kantanen & Manninen, 2014), the empirical data were collected from the international LinkedIn group Higher Education Teaching and Learning (HETL), founded in February 2010. As of January 2016, the group has 63,124 members. The HETL LinkedIn group is a network of professionals who participate by using their own names, job titles and photos. The research question of this study is: *How can discussions in virtual communities contribute to professional learning and development?* In addition to these studies, our project included Master's theses writers who use the HETL discussion threads as data. One completed thesis analysed the sense-making process in a discussion thread concerning the use of electronic devices in class (Tiiliharju, 2015).

This paper focuses on the ethical issues related to our HETL LinkedIn group study because its global nature allows more diverse ethical issues to be discussed than national forums. In 2010, at the advent of social media expansion, there was an exciting discussion thread in the HETL LinkedIn community. The opening question was: "Do you accept your students' invitations to connect on Facebook and other social networks?" The discussion continued over seven months and included about 280 discussants from over 190 organisations. Altogether, there were 508 replies. When printed out, the data were 135 pages long. We are interested in

Table 1. Virtual communities of practice project.

	Online Community	Practitioners Studied	Data	Approach/Methods
Pilot study	Finnish community Social Media Supporting Web-Learning (*Sometu*)	Higher education and business experts involved in the use of social media in education	Discussion threads	Qualitative content analysis (ATLAS.ti)
Current study	Global LinkedIn community, Higher Education Teaching and Learning (HETL)	Higher education and business experts involved in higher education and learning	Discussion threads, online observations, interviews	Qualitative content analysis (ATLAS.ti), Virtual ethnography (netnography)

both the process and the content of idea and innovation development in the forum. We have already studied *how* new ideas and innovations are developed on virtual forums, and *what* kinds of ideas and innovations are developed. In this paper, the focus is not on our research results, but on different ethical issues related to the study of this global virtual community of education experts.

Our study qualifies as Internet research because we utilise the Internet to collect data, study how people use the Internet through observing participation in virtual forums, utilise datasets available via the Internet, and employ content analysis to study the web (Markham & Buchanan, 2012). Drawing on Kozinets (2010), we can consider our study a netnography because we enter discussion forums as researchers, use discussion threads as natural data, observe participants on the forum, and also participate in the discussions ourselves. We are both members of the HETL group, and one of us also participated in the Facebook friending discussion thread. According to Kozinets (2010), netnographers should spend significant time interacting within and becoming a part of an online community. We see this participation as an advantage in our attempt to understand the innovation development process and communication of the group. Moreover, we claim that our thorough understanding of the context, culture and dynamics of the target forums contributed to the quality of our research because we were able to interpret our data from the point of view of our informants (Turtiainen & Östman, 2013).

3.1. Challenge 1: Ethical Review

We completed our literature review on virtual communities in early 2013 (Kantanen et al., 2014) and were ready to proceed to the data collection phase of our study. We thought that it would be a good idea to have the Committee on Research Ethics (equivalent to institutional review boards, or IRBs, in the U.S.) of our home university evaluate and approve our research plan. In our country, and in the field of the humanities and social and behavioural sciences, an ethical review is needed if the study involves an intervention in the physical integrity of subjects, deviates from the principle of informed consent, involves children under the age of 15, exposes research subjects to strong stimuli and evaluating possible harm requires expertise, has the potential to cause long-term mental harm, or poses a safety risk to subjects (Finnish Advisory Board on Research Integrity, 2009). None of these elements were involved in our planned project; however, we intended to study the global LinkedIn forum, HETL, operated from within the United States, and concluded that obtaining ethical approval would be advisable. The Chair of the Committee assured us that our national procedure is very cautious, and that the evaluation of ethical aspects of the research plan would be valid anywhere in the world.

Neither of us had previous experience with ethical reviews because, in our fields of business and adult education, the standard research process does not include ethical reviews—as it often does, for example, in the field of health research. Moreover, Finland has often been characterised as a high-trust society (e.g., Korhonen & Seppälä, 2005), which includes trust towards institutions, such as universities and their researchers. Therefore, many research environments have been accessible to our researchers without complicated application procedures.

Our first application was returned for revision, as was the second application. At this stage, however, we had come far enough that no Committee meeting was needed; but we nonetheless had to make final changes, which would then be evaluated and accepted—or rejected—by the Chair and the Secretary. In total, the process from the first application to the supporting statement took 1.5 years. There were several reasons for the delay. First, the instructions given and the structure and themes of the application form were poorly applicable to the type of Internet research we were planning. For this reason, we found it difficult to answer the questions as specifically as was required. Moreover, we were quite frustrated and not able to rewrite the application because of other duties. It was, indeed, very difficult to make decisions about possible ethical issues in advance of our study (Markham & Buchanan, 2012).

The HETL forum founder, who is the central gatekeeper when accessing the forum, accepted the ethical review from our home university, but it was a long process to gain the ethical approval. We obviously applied for the ethical review without fully thinking through the different aspects of Internet research ethics. Had we known then what we know now, our application would have been accepted in the first round. Our struggle with the ethics committee triggered a more profound understanding of the different and often complex dimensions of Internet research ethics. Because we have realised that both our students and fellow researchers sometimes have quite a nonchalant attitude towards ethical issues and Internet data, we wish to share what we have learned.

3.2. Challenge 2: Public or Private Forum

Like many other social media researchers, we pondered whether the HETL LinkedIn group was a public or private forum, or if it was a text more than it was a place. If it was a place, was it a public place? If it was a text, could it be quoted like a book? Or, could it be treated like the Letters to the Editor section in newspapers, where readers submit their input knowing that it will be made publicly available to all readers?

If the discussions can be considered a text, then the focus is totally different from the point of view of ethics than when dealing with human subjects. Could we observe LinkedIn discussants just as we would observe people in a public place? Or could we consider their inputs in the forum as texts? Does it make a difference if the discussion thread studied is several years old, as compared to a synchronous discussion?

McKee and Porter (2009b) make a distinction between a space and a place. If an Internet site is seen as a space, it is primarily a medium of publication, and the focus of research is mainly on what is published. If it is a place, people gather there to discuss and to share, and the focus of research is on people instead. The dominant view has long been that the Internet is a social domain, and this has led to "an imperative to apply restrictions from the human subjects model" (Bassett & O'Riordan, 2002, p. 234). The Internet is neither public nor private. It is neither a place (residence) nor a space (publication medium). As Kozinets (2010) explains, it is actually many types of social interaction, including chats, blogs, soundclips, and videos.

Private information is such that an individual can expect it to not be monitored or collected, or made publicly available. However, questions of publicity and privacy are complicated in virtual environments. Even on public forums, people may have expectations of privacy, or find it inappropriate that their inputs are read, collected or analysed by external parties (Dennen, 2012; Markham & Buchanan, 2012; Walther, 2002; Zimmer, 2010). Even with postings on public discussion forums, people do not expect that their inputs will be analysed. For example, in a public virtual world like Second Life, users consider their virtual homes to be private (Rosenberg, 2010). Also, public blogs have been considered to be a part of their writers' identity and should not be treated as publicly available data (Dennen, 2012; Markham & Buchanan, 2012). Dennen (2012) maintains that the only reliable way to evaluate a research subject's desire for privacy is to request his or her consent to participate in research.

In our case, our data, the Facebook friending thread in the HETL LinkedIn forum, can be considered as text because it was active several years earlier and is, therefore, archived material. Moreover, the question of publicity was clarified in the HETL Policy, which was published after we started our project. The Policy states that "The HETL LinkedIn discussion forum (i.e., global online community or practice) is considered a quasi-public group for the purpose of academic research and existing postings are therefore exempt. However, researchers must get approval from the HETL IRB before starting any research project involving this discussion forum data" (https://www.hetl.org/hetl-research/). The Institutional Review Board (IRB) was founded after we started our study, but the use of forum postings was approved by the HETL forum's founder, Dr. Patrick Blessinger, and we have committed to keeping him, as well as the Chair of the HETL Review Board, informed about the different phases of our project.

3.3. Challenge 3: Human Subjects Research—Or Not?

The definition of human subjects research matters because in many countries, like in the U.S., approval by the Institutional Review Board (IRB) is always required when researching human subjects (Bruckman, 2006). A study becomes human subjects research when the researcher deals and interacts with a living person (Bruckman, 2002). But do we interact with living persons when studying the LinkedIn group discussion threads?

"Human subjects research is research in which there is an intervention or interaction with another person for the purpose of gathering information, or in which information is recorded by a researcher in such a way that a person can be identified through it directly or indirectly" (U.S. Department of Health & Human Services, 2009).

The concept of human subjects research, which was originally related to the treatment of persons in medical experiments, still defines ethical research considerations (Markham & Buchanan, 2012). The concept has also been criticised by, for example, Bassett and O'Riordan (2002, p. 244), who claim that "research that positions the Internet as a social space containing cultural activity ripe for observation ignores the range of textual applications that the Internet supports." Bruckman (2002) suggests that, instead of human subjects, we should talk about "amateur artists" who use the Internet as a playground to create semi-published work.

According to the definition above, ethnographic Internet research (netnography) *is* human subjects research. However, research use of spontaneous conversations, gathered in a publicly accessible venue, is not human subjects research, according to the Code of Federal Regulations (2009) that governs Institutional Review Boards (IRBs) in the United States (Kozinets, 2010). The Association of Internet Researchers reminds us that because there are individual persons involved in digital information, researchers might need to consider principles related to research on human subjects (Markham & Buchanan, 2012).

Terms such as harmful, vulnerable, or personally identifiable information may be more relevant than the human subjects model, at least outside of the regulatory framework of research ethics (Markham & Buchanan, 2012). Instead of the spatial models behind the human subjects view, Bassett and O'Riordan (2002) suggest a hybrid model of relational ethics that would incorporate text, space and bodies and thus extend the limited application of the human subjects model. They

argue that the human subjects research model is too narrow and does not consider the Internet as a cultural production of texts.

The conclusion is that research of publicly accessible conversations—or even semi-public conversations, like in our HETL case—is not human subjects research, but this does not exclude the need to consider ethical principles of avoiding harm and protecting privacy.

3.4. Challenge 4: Participant Recruitment and Informed Consent

When we discuss recruiting participants for research, we address two separate issues: first, under which conditions can we use the HETL discussions as research material; and second, what kind of process should be applied when recruiting HETL members for additional interviews.

In our study, we have thus far only been dealing with forum discussions. Because the HETL policy considers group postings exempt for academic research, we have not yet been involved in participant recruitment when studying forum postings. Also, Kozinets (2010) states that a netnographer's normal, asynchronous actions in online communities do not require informed consent. However, we are aware that there are different points of view about this. In the U.S., a utilitarian stance may prevail, meaning that benefits to society are weighed against potential risks; whereas in Central and Northern Europe, a deontological or communitarian stance that does not compromise confidentiality and anonymity may be taken (Markham, 2006, p. 48). Despite our Finnish background, we have adopted the Anglo-American view because the focal group for our study is an online forum administrated from the U.S.

The question of informed consent will become relevant at the next stage of our project because informed consent is always required for interviews. Also, the HETL Policy states that when collecting research data directly from research participants, researchers must first get approval from the HETL Institutional Review Board and receive informed consent from the research participants (https://www.hetl.org/hetl-research/). In the ethical review that we went through, the consent form that we made on the basis of Kozinets (2010, p. 194), as well as the process of electronic consent, was approved for our research purposes by our local ethics committee.

Enrolling research participants requires contact between the researchers and the study participants. Several sources consider informed consent to be a cornerstone of ethical research conduct (e.g., Buchanan & Zimmer, 2013). The process of informed consent involves the participants becoming aware, through conversations, of the purpose of the study, what rights and responsibilities participation involves, the risks and benefits of the study, possible compensation or costs, confidentiality and participant rights (Flicker, Haans, & Skinner, 2004; Kozinets, 2010). When studying virtual communities, all of this can be done online, and several authors advise gathering all relevant research information into a project website, and obtaining informed consent through a registration page (Bruckman, 2002, 2006; Kozinets, 2010). This process is recommended for adults and non-vulnerable populations, as well as when the research is not high risk.

According to our experience, online groups' Terms of Service and policies are very helpful when pondering the question of participant recruitment (see also Markham & Buchanan, 2012, p. 8). In our case, the HETL Policy requires HETL Institutional Review Board approval for any research project, as well as informed consent when collecting data directly from forum participants. The Policy also helped shape the public–private division and text–space discussion. Moreover, we advise researchers to identify different gatekeepers and to discuss their intentions openly with forum founders, administrators and group moderators. This is what we did with both our national and global group. Then, a statement can be added to the informed consent form that specifies that the study plan has been accepted by the forum administrator or owner.

One could also consider the timing of informed consent from a different point of view. Several authors (Markham & Buchanan, 2012; McKee & Porter, 2009b) suggest a process approach to research ethics. One may not need consent when collecting the data, but rather at the reporting and dissemination stage, if one wishes, for example, to quote a forum participant (Dennen, 2012). Many decisions at this stage pose ethical concerns; for example, which (if any) details to reveal about the study site and participants.

3.5. Challenge 5: Observing the Community

Much debate exists concerning how researchers should behave when observing a virtual community. The general rule should be to work as transparently as possible because, as Rutter and Smith (2005, p. 90) argue: "the ramifications of unethical disclosure are real and inescapable." However, not even the question of disclosing your presence is simple in the online environment. In our case, it was easy because we did not study synchronous discussions of the HETL LinkedIn group, but rather those that had become active several years earlier, and were already archived. However, during potential new phases of our study, these questions may become more important. We now have enough knowledge to face these questions in an ethical manner; that is, by disclosing our presence as researchers to the forum participants or to the discussants of a specific thread.

Kozinets (2010) advises netnographers to always disclose their presence, affiliations and intentions. Sev-

eral authors are unequivocal about identity deception, such as Kozinets (2010): "Netnographers should *never, under any circumstances*, engage in identity deception" (p. 147). Identity deception may occur on the sites studied, but researchers should always identify themselves. There are different ways to do this, such as putting information about ongoing research on one's own personal profile (Bruckman, 2006).

There may be projects that would be unsuccessful if the researcher were to disclose his or her presence. Hudson and Bruckman (2004) studied how chatroom discussants responded to the researchers' presence and/or attendance in their forum and found that they reacted with hostility when they became aware of being studied. Therefore, the researchers considered the informed consent process to be impracticable. Sanders (2005) observed online communities of sex workers without disclosing her presence, as a so-called "lurker". She did not want the participants to alter their behaviour, nor did she want to harm the shared community or provoke hostility. Sometimes, this kind of anonymity may also protect the researcher. There may be some unwanted consequences if the researcher were to reveal his or her identity and affiliations to certain online groups (Sanders, 2005). It is possible that because hostile online behaviour is becoming more common (Jane, 2015), researchers might need to consider their own protection, even when studying relatively "safe" online forums—especially if discussed issues are sensitive, such as topical discussions revolving around asylum seekers. Bruckman (2006) gave an example of a student researcher who became the victim of racism on the site that he was studying. We know from our home university that hostile messages have been targeted at researchers working on issues related to immigration (Mikkonen, 2015).

3.6. Challenge 6: Reporting and Dissemination

Reporting and dissemination ethics apply to the various media in and through which research results are presented: academic journals and reports, conference presentations, websites, videos, press releases, interviews and information leaflets. Each of these reporting and dissemination mediums have their own distinct consequences. For example, writing about research methods transparently and in detail increases the credibility of a study.

It is obvious that there has been a delay in our research publication and reporting efforts because of the different ethical concerns we have experienced. The ethical questions at this phase of the study include, in our experience, questions of citing, anonymising, and crediting. Can we mention the HETL LinkedIn group and the specific discussion thread studied in our articles? Do the participants own copyright over their input in the discussion forum? If so, do we need to add author references for all cited input? And, if we delete all identifiers (like names and affiliations) to protect the informants' privacy, how do we give credit to those who contributed innovative ideas in the discussion thread?

Bruckman (2006, p. 91) states that "one of the thorniest problems concerns how to disguise names of people and sites." She asks how the researcher should balance the need to protect research participants with the need to give credit for their work (Bruckman, 2002). She uses the metaphor of "amateur artists", meaning that "all user-generated content on the Internet can be viewed as various forms of amateur art and authorship" (p. 229); consequently, these amateur artists deserve credit for what they have produced.

Bruckman (2002) proposes different levels of disguising, on a continuum from no disguise, to light and moderate disguise, to heavy disguise. In the case of no disguise, the report would have real names or pseudonyms of research subjects and the researcher would, therefore, respect the individual's copyright over his or her input. At the other end of the continuum, in the case of heavy disguise, the group studied would not be named and all identifying details would be changed. The report would include no verbatim quotes, and some false details might be introduced. It is worth noting that one cannot be careful enough with personal data. Zimmer (2010) explains a case where the researchers failed to anonymise their data with information about 1,700 college students' Facebook profiles. Even though all of the identifiers were deleted or encoded, the university in question was discovered and the privacy of research subjects' sensitive personal information was endangered.

Our case falls into the category of "light disguise" (Bruckman, 2002), where the community studied is named but participant names and some other identifying details like organisational affiliations are removed. We are also going to use verbatim quotes in our papers, even if they can be used to identify a group member with the help of search engines. For example, Moreno et al. (2013) advise against direct quotations to protect confidentiality. We do not consider this to be a problem because, as explained above, our study is not of high risk and, therefore, light disguise should be enough. Some authors have adopted the utilitarian approach to research ethics (Markham, 2006). For instance, Bruckman (2002) advises balancing the degree of risk against the benefits of the study, while Markham and Buchanan (2012) advise balancing the rights of subjects with the social benefits of the research and researchers' rights to conduct research.

McKee and Porter (2009b, p. 88) illustrate this issue of disguising by constructing a figure within which the necessity of informed consent can be assessed according to the following variables: Public vs. private, Topic sensitivity, Degree of interaction and Subject vulnerability. Each variable should be evaluated along the pri-

vate–public or high–low scale. At the "low end" of the scale, consent may not be necessary; while at the "high end" of the scale, consent will most likely be required. The authors emphasise that these variables can be hard to determine and can also be culture–specific; in other words, a sensitive topic for one culture may not be sensitive at all for another. That said, our study occupies the "low end" of the scale in regards to all of these variables.

Importantly, our research question specifically examines verbatim quotes, not who said them, which thus eliminates the need to give credit to discussion participants. The main purpose of our study is neither to analyse what kind of innovations are generated in a discussion nor to examine the behaviour of those who generate them. Instead, the objective is to analyse online discussions and the internal logic of how text-based, delayed, asynchronous communication develops and proceeds in online communities (see Kantanen et al., 2014). In other words, we are not interested in *who* wrote *what*; rather, we are interested in the types of input (e.g., questions, experience sharing, reflections, etc.) generated in online discussions, as well as whether such discussions follow the five cycles of value creation (Wenger, Trayner & de Laat, 2011). As such, in this kind of analysis, verbatim quotes are important—even mandatory—because the research focus is on *how* text is written.

Bruckman (2002) suggests that if the study is low risk, the researcher could ask the research subjects if they want a pseudonym, real name, both, or neither to be used. If the study is high risk, it is not appropriate to list names or pseudonyms. Minimal risk means, by definition, "that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests" (U.S. Department of Health & Human Services, 2009, Section 46.102). Our study falls into this category of minimal risk.

Ethical research conduct could and should also include the sharing of research results with the community studied so that corrections in the analysis and interpretation of the results can be made when needed (Madge, 2007). This kind of conduct might contribute to a more positive online environment where the researchers would be met with trust rather than with suspicion. We shared the results of our pilot study (Kantanen et al., 2014) with the national *Sometu* community and received positive feedback from the forum administrators.

4. Discussion

In the following, we discuss our case in light of the basic ethical principles of research defined in our country: respecting the autonomy of research subjects, avoiding harm, and protecting privacy and data (Finnish Advisory Board on Research Integrity, 2009).

Participation in research must be voluntary and based on informed consent. So far, we have not necessarily needed consent because we have been studying archived online discussion threads. When interviewing participants, we will need their consent, which we can gain electronically by using a form accepted by the Committee on Research Ethics of our university. The form includes the following information for the research subjects: project title, research team members, purpose of the study, research funding, information about what participation will involve, risks, benefits, costs and compensations, confidentiality, participant rights, and contact information (see also Kozinets, 2010, p. 194).

Our study has not at any stage included particularly vulnerable groups like minors, cognitively impaired people, survivors of abuse, or support groups for those with serious diseases. Groups of that kind are, of course, very vulnerable and deserve extra protection and careful ethical consideration.

In our country, some exemptions from informed consent are possible "if advance information would distort the results of the study" (Finnish Advisory Board on Research Integrity, 2009, p. 8). For example, power relationships can be studied without the consent of those in power, or there may be groups that can be a risk to the researcher's safety if he or she reveals his or her identity. In these cases, ethical review is always necessary. Research cultures differ from country to country and, therefore, global forums present challenges. For instance, many virtual communities and social media platforms are owned by and administered from the United States; therefore, it is not enough to take national guidelines of research ethics into account, but U.S. standards as well. Moreover, if the research results are to be published in one of the "international" languages, other nations' standards may also need to be taken into account.

We consider the question of protecting Internet researchers a very topical one, even though this has not been a problem in our study. What if a senior investigator sends a junior to study a hostile online forum and the research subjects trace his or her personal information and start verbal attacks with threats of off-line violence? Is this a question of research ethics? Or is it unethical to withdraw from studying risky issues and/or forums because of fear? Both Bruckman (2006) and Jane (2015) give terrifying examples of what impudent online hostility—Jane calls this "e-bile"—can involve. Therefore, we wish to include the aspect of protecting researchers, particularly inexperienced junior researchers, in our ethical considerations.

We have attempted to treat our research subjects with respect and dignity, which is not difficult because

they (like us) are mostly experts within the field of education. Our study does not include elements that might cause mental, financial or social harm to the research subjects. If, however, we had studied synchronous interaction in online communities, our presence might have had negative consequences to the natural flow of conversations and could have provoked irritation or hostility—mental harm, that is.

The Terms of Service of different virtual platforms may have different views on privacy than those of their users (Markham & Buchanan, 2012). Moreno, Goniou and Moreno (2013) give examples of cases where courts have concluded that a person cannot expect privacy when posting on social networking sites. Often, ethical demands can be harder than legal ones. Extant laws set the ultimate limits, but do not necessarily determine what is ethical. In this article, our focus has been on ethical, not legal considerations (for a possible basis for legal liability of Internet researchers, see Lipinski, 2006).

Protecting privacy in research publications may be challenging because we use direct quotes from our data. However, as explained above, the issues discussed in our community are not very sensitive; therefore, a "light disguise" (Bruckman, 2002), where the names and other identifiers are hidden, should be enough. It is of course possible that if a quotation appears in a journal article, someone could track the original source of the quotation with the help of a search engine. Therefore, there is a slight risk that our study might cause some reputational harm to forum participants if, for example, original quotations with hostile and aggressive comments, "flames" can be tracked. However, this risk is low and, if we ever publish any flaming-related quotes, the focus will be on showing how moderators diffuse heated situations in a specific virtual community, not on the identities of those involved in them.

In our project, each researcher is responsible for safeguarding the research materials in his or her office and office computer. For ethical reasons, the materials may not be handed over to third parties—not even after the research project has been finished. After the project, the research data will be stored at the university for ten years according to the prevailing archive's regulations. Our data will not be available for secondary research, not even without identifiers, because this is the procedure accepted by our ethics committee. In this regard, the ethics committee may need to change their view because of increasing demand for openly accessible research data.

5. Conclusion

"Internet research ethics is complex, not impossible" (McKee & Porter, 2009b, p. 141). However, Bruckman (2006) maintains that studying online communities in an ethical fashion is a challenge, even for experienced researchers, and that the Internet continually raises novel ethical issues. If we are too lax in our ethics, we may violate the rights of individuals, or disrupt the communities we study; if we are too strict, we may not gain the knowledge needed to understand these communities (Bruckman, 2002, p. 218). She gives an example of how one student of hers was denied access to study an online group because someone else had studied the community earlier, "and left members feeling like their activities were disrupted and their privacy violated" (Bruckman, 2006, p. 217). Therefore, forgetting ethical conduct harms the whole academia.

Working with international forums can prove problematic. Researchers, forum owners and moderators, and research subjects may come from very different cultures, including research cultures with different views on research ethics. Therefore, we recommend consulting the ethical review board of one's own university. First, it is much easier to approach an international community with ethical approval, because in many countries it is always a part of any research process. Moreover, the process makes researchers seriously consider different ethical aspects of their projects. That said, we would like to add that, of course, it helps if the local ethics committee or review board has a supportive, advisory attitude towards its researcher clients. By delving into the complexities of Internet research, we have also tried to pay attention to the need for ethical considerations in fields where official reviews are not necessary. For us, the process that started with our local ethics committee has been very useful in the long run.

Acknowledging the ephemeral characteristics of Internet contexts, we argue that, in line with other authors (Markham & Buchanan, 2012; McKee & Porter, 2009b), ethical considerations should be more case-based and processual, rather than relying on one model for all solutions. We suggest that local ethics committees or institutional review boards could, with their expert knowledge of ethics, provide valuable support for researchers operating in the complex and dynamic terrain of Internet research, as well as in fields and research settings where an ethical review is not a standard part of the research process. There may also be a need for these review boards to revise their instructions and forms to better respond to the volatile research environments studied by Internet researchers (Dennen, 2012).

Acknowledgements

This article is one of the fruits of the Innovation Practice Project of the Business School of the University of Eastern Finland, funded as one of the Innovative Research Initiatives of the University during 2011–2014. It originated in a workshop on Ethical and Unethical Re-

search of Internet Data, held by Helena Kantanen, at the International Doctoral Consortium on Management and Organization Studies at St. Mary's University, Halifax, in June 2014. The key points of this paper were presented at the Social Media Research Symposium at the University of Jyväskylä, Finland, on November 20, 2015. We thank our colleagues, Professor Päivi Eriksson and Professor Albert J. Mills, for their valuable insights on improving the manuscript.

Conflict of Interests

The authors declare no conflict of interests.

References

Bassett, E. H., & O'Riordan, K. (2002). Ethics of Internet research: Contesting the human subjects research model. *Ethics and Information Technology*, *4*(3), 233-247.

Bruckman, A. (2002). Studying the amateur artist: A perspective on disguising data collected in human subjects research on the Internet. *Ethics and Information Technology*, *4*(3), 217-231.

Bruckman, A. (2006). Teaching students to study online communities ethically. *Journal of Information Ethics*, *Fall 2006,* 82-95.

Buchanan, E., & Zimmer, M. (2013). Internet research ethics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2013 Edition). Retrieved from http://plato.stanford.edu/archives/fall2013/entries/ethics-Internet-research

Dennen, V. P. (2012). When public words are not data. Online authorship, consent, and reasonable expectations of privacy. In D. Heider & A. L. Massarini (Eds.), *Digital ethics* (pp. 21-38). New York: Peter Lang.

Eriksson, P., & Kovalainen, A. (2016). *Qualitative methods in business research*. *A practical guide to social research*. London: Sage.

Finnish Advisory Board on Research Integrity. (2009). *Ethical principles of research in the humanities and social and behavioural sciences and proposals for ethical review*. Retrieved from http://www.tenk.fi/sites/tenk.fi/files/ethicalprinciples.pdf

Flicker, S., Haans, D., & Skinner, H. (2004). Ethical dilemmas in research on Internet communities. *Qualitative Health Research*, *14*(1), 124-134.

Hine, C. (2000). *Virtual ethnography*. London: Sage.

Hine, C. (2005). Virtual methods and the sociology of cyber-social-scientific knowledge. In C. Hine (Ed.), *Virtual methods. Issues in social research on the Internet* (pp. 1-13). Oxford: Berg.

Hudson, J. M., & Bruckman, A. (2004). "Go away": Participant objections to being studied and the ethics of chatroom research. *The Information Society*, *20*(2), 127-139.

Jane, E. A. (2015). Flaming? What flaming? The pitfalls and potentials of researching online hostility. *Ethics and Information Technology*, *17*(1), 65-87.

Kantanen, H., & Manninen, J. (August, 2014). *Seeds, sprouts, weeds and blooms of innovation: Learning in professional virtual communities*. Paper presented at the 7th EARLI SIG 14. Learning and Professional Development Conference. Transformations in Knowledge and Professional Learning, Oslo, Norway.

Kantanen, H., Manninen, J., & Kontkanen, J. (2014). Emergent dialogue as a prerequisite of learning and innovation in professional virtual communities. *International Journal of Web Based Communities*, *10*(2), 211-231.

Korhonen, J., & Seppälä, N. (2005). The strength of a high–trust society. In A. Habisch, J. Jonker, M. Wegner, & R. Schmidpeter (Eds.), *Corporate social responsibility across Europe* (pp. 13-22). Berlin: Springer.

Kosonen, M. (2009). Knowledge sharing in virtual communities—A review of the empirical research. *International Journal of Web Based Communities*, *5*(2), 144-163.

Kozinets, R. V. (2010). *Netnography. Doing ethnographic research online*. Los Angeles: Sage.

Lakkala, M. (2010). How to design educational settings to promote collaborative inquiry: Pedagogical infrastructures for technology–enhanced progressive inquiry. (Doctoral dissertation). University of Helsinki, Helsinki, Finland. Retrieved from http://urn.fi/URN:ISBN:978-952-10-6178-3

Lipinski, T. A. (2006). Emerging tort issues in the collection and dissemination of Internet-based research data. *Journal of Information Ethics*, *Fall 2006*, 55-81.

Madge, C. (2007). Developing a geographers' agenda for online research ethics. *Progress in Human Geography*, *31*(5), 654-674.

Manninen, J., & Nevgi, A. (2000). Opetus verkossa [Teaching in the Web]. In J. Matikainen & J. Manninen (Eds.), *Aikuiskoulutus verkossa: Verkkopohjaisten oppimisympäristöjen teoriaa ja käytäntöä*. [Adult Education in the Web—Theory and Practice of Web Based Learning Environments] (pp. 93-108). Helsinki: Palmenia-kustannus.

Markham, A. N. (2006). Ethic as method, method as ethic: A case for reflexivity in qualitative ICT research. *Journal of Information Ethics*, *Fall 2006*, pp. 37-54.

Markham, A., & Buchanan, E. (2012). *Ethical decision-making and Internet research. Recommendations from the AoIR ethics working committee (version 2.0).* Retrieved from http://aoir.org/reports/ethics2.pdf

McKee, H. A., & Porter, J. E. (2009a). Playing a good game: Ethical issues in researching MMOGs and virtual worlds. *International Journal of Internet Research Ethics*, *2*(1), 5-37.

McKee, H. A., & Porter, J. E. (2009b). *The ethics of internet research: A rhetorical, case-based process*. New York: Peter Lang.

Mikkonen, J. (2015, November 7). Tutkija sai tarpeekseen vihapuheesta—harkitsee tutkintapyyntöä [Researcher Fed Up with Hostile Messages—Considers request for investigation]. *Savon Sanomat*. Retrieved from http://www.savonsanomat.fi/uutiset/kotimaa/ita-suomen-yliopiston-tutkija-sai-tarpeekseen-vihapuheesta-harkitsee-tutkintapyyntoa/2164266

Moreno, M. A., Goniou, N., & Moreno, P. S. (2013). Ethics of social media research: Common concerns and practical considerations. *Cyberpsychology, Behavior, and Social Networking*, *16*(9), 708-713.

Rosenberg, Å. (2010). Virtual world research ethics and the private/public distinction. *International Journal of Internet Research Ethics*, *3*(12), 23-35.

Rutter, J., & Smith, G. W. H. (2005). Ethnographic presence in a nebulous setting. In C. Hine (Ed.), *Virtual methods. Issues in social research on the Internet* (pp. 81-92). Oxford: Berg.

Sanders, T. (2005). Researching the online sex work community. In C. Hine (Ed.), *Virtual methods. Issues in social research on the Internet* (pp. 67-79). Oxford: Berg.

Tavani, H. T. (2006). Cyberethics as an interdisciplinary field of applied ethics. Key concepts, perspectives, and methodological frameworks. *Journal of Information Ethics*, *Fall 2006*, 18-36.

Tiiliharju, P. (2015). *Yhteisöllinen oppiminen ja järkeistäminen virtuaalisissa käytäntöyhteisöissä* [Collective learning and sensemaking in virtual communities of practice] (Master Thesis). University of Eastern Finland, Joensuu, Finland.

Turtiainen, R., & Östman, S. (2013). Verkkotutkimuksen eettiset haasteet: Armi ja anoreksia [Ethical challenges of internet research: armi and anorexia]. In S. Laaksonen, J. Matikainen, & M. Tikka (Eds.), *Otteita verkosta. Verkon ja sosiaalisen median tutkimusmenetelmät* [Grips on the net. Internet and social media research methods] (pp. 49-67). Tampere: Vastapaino.

U.S. Department of Health & Human Services. (2009). *Code of federal regulations. Title 45. Public welfare department of health and human services*. Retrieved from http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html

Walther, J. B. (2002). Research ethics in Internet-enabled research: Human subjects issues and methodological myopia. *Ethics and Information Technology*, *4*(3), 205-216.

Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, *29*(1), 35-57.

Wenger, E., Trayner, B., & de Laat, M. (2011) *Promoting and assessing value creation in communities and networks: A conceptual framework*. Retrieved from http://wenger-trayner.com/wp-content/uploads/2011/12/11-04-Wenger_Trayner_DeLaat_Value_creation.pdf

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313-325.

About the Authors

Helena Kantanen (PhD) is a Senior Lecturer with the Innovation Management Team of the Business School of the University of Eastern Finland. She joined the faculty in 2010, after 20 years as a PR practitioner, manager, and consultant. Her research focuses on virtual expert communities and on business renewal through participative management, leadership and communication. Helena's teaching involves courses concerned with marketing, communication and leading and managing change. She is a regular Erasmus visitor of ESTICE Business School, in Lille, France.

Jyri Manninen (EdD) is a Professor of Adult Education at the University of Eastern Finland. His research interests include adult learning processes and human interaction in various face-to-face and online learning situations. One research topic has been computer-mediated communication in web-based learning environments and emergent dialogue in virtual expert communities.

Media and Communication (ISSN: 2183-2439)
2016, Volume 4, Issue 4, Pages 97-108
doi: 10.17645/mac.v4i4.593

Article

#JeSuisCharlie: Towards a Multi-Method Study of Hybrid Media Events

Johanna Sumiala [1], Minttu Tikka [1,*], Jukka Huhtamäki [2] and Katja Valaskivi [1]

[1] School of Communication, Media and Theatre, University of Tampere, 33014 Tampere, Finland; E-Mails: johanna.sumiala@helsinki.fi (J.S.), minttu.tikka@uta.fi (M.T.), katja.valaskivi@uta.fi (K.V.)
[2] Intelligent Information Systems Laboratory, Tampere University of Technology, 33720 Tampere, Finland; E-Mail: jukka.huhtamaki@tut.fi

* Corresponding author

Submitted: 9 February 2016 | Accepted: 6 May 2016 | Published: 10 October 2016

Abstract
This article suggests a new methodological model for the study of hybrid media events with global appeal. This model, developed in the project on the 2015 *Charlie Hebdo* attacks in Paris, was created specifically for researching digital media—and in particular, Twitter. The article is structured as follows. Firstly, the methodological scope is discussed against the theoretical context, e.g. the theory of media events. In the theoretical discussion, special emphasis is given to i) disruptive, upsetting, or disintegrative media events and hybrid media events and ii) the conditions of today's heterogeneous and globalised media communication landscape. Secondly, the article introduces a multi-method approach developed for the analysis of hybrid media events. In this model, computational social science—namely, automated content analysis (ACA) and social network analytics (SNA)—are combined with a qualitative approach—specifically, digital ethnography. The article outlines three key phases for research in which the interplay between quantitative and qualitative approaches is played out. In the first phase, preliminary digital ethnography is applied to provide the outline of the event. In the second phase, quantitative social network analytics are applied to construct the digital field for research. In this phase, it is necessary to map a) what is circulating on the websites and b) where this circulation takes place. The third and final phase applies a qualitative approach and digital ethnography to provide a more nuanced, in-depth interpretation of what (substance/content) is circulating and how this material connects with the 'where' in the digital landscape, hence constituting links and connections in the hybrid media landscape. In conclusion, the article reflects on how this multi-method approach contributes to understanding the workings of today's hybrid media events: how they create and maintain symbolic battles over certain imagined constructs of social imaginaries of solidarity, belonging, contestation, and exclusion, a topic of core value for the theory of media events.

Keywords
automated content analysis; *Charlie Hebdo*; digital ethnography; hybrid media event; social network analytics; Twitter

Issue
This article is part of the issue "Successes and Failures in Studying Social Media: Issues of Methods and Ethics", edited by Epp Lauk and Niina Sormanen (University of Jyväskylä, Finland).

© 2016 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction: *Charlie Hebdo* 2015, a Hybrid Media Event

On Wednesday, 7 January 2015, at 11:30 a.m., French–Algerian brothers Saïd and Chérif Kouachi attacked the headquarters of *Charlie Hebdo*. Eleven people were killed in the rampage. After the attack, the Kouachi brothers returned to their car and exchanged fire with the police officers blocking their escape route. A few minutes later, they executed an injured police officer named Ahmed Merabet at point-blank range. The perpetrators escaped from the building, and the shooting of the police officer was filmed from a nearby apartment. The event instantly exploded into a transnational media event, and the amateur video material that was filmed began to circulate rapidly. Newsrooms all over

the world followed the massive security operation as the Kouachi brothers hijacked another car and fled north out of Paris. In the evening, tens of thousands of people took to the streets around Europe to show their solidarity with those killed by the gunmen. The next day, 8 January 2015, the attackers continued their escape, and thousands of security personnel were deployed to comb the area approximately 90 kilometres from Paris, where the two men were last seen. Meanwhile, in Paris, reports emerged that a policewoman had been shot and killed; however, the link with the *Charlie Hebdo* attack was not immediately apparent. As night fell, the Eiffel Tower's lights were switched off in memory of the victims. On Friday, 9 January 2015, the police located the attackers in the Dammartin-en-Goële area. The brothers were chased to an industrial complex 35 kilometres from Paris, where they seized a printworks and took a hostage. In east Paris, at around 12:30 p.m., a third gunman named Amedy Coulibaly seized a Jewish supermarket, killed four people, and took hostages. It emerged that Coulibaly was responsible for the killing of the Parisian policewoman, Clarissa Jean-Philippe, the day before. In his phone call to the French TV station BFM-TV, Coulibaly stated that his attack was synchronised with the attacks of the Kouachi brothers, and that they belonged to the same group of terrorists. He also threatened to kill his hostages unless the Kouachi brothers were allowed to go free. After several hours of this hostage situation, police special forces stormed the market and killed Coulibaly. The Kouachi brothers were killed by the special forces on the same day.

Over the course of these three days, new updates constantly appeared on websites, on YouTube, and on news broadcasts. Social media websites were inundated with comments, links, and images connected to the event, and these were shared and commented on by both journalists and ordinary citizens. The course of the events, as presented by professional journalists and international and national media houses, was intermixed with memes and comments that citizens from different countries shared via social media. In addition, various strategic and spontaneous (both political and religious) interest groups made use of the situation and competed for attention, tailoring and recycling details about the events with content aimed at different audiences.

One of the prominent features of the *Charlie Hebdo* event was the use of the slogan *Je Suis Charlie* ("I Am Charlie"), which became a symbol of solidarity and freedom of expression. The volume of communication around the event is well illustrated by the fact that the hashtag *#JeSuisCharlie* was—at least at the time of the event—the most popular tweet in the history of Twitter. The tag was tweeted 6,500 times per minute at its height and was featured in 3.4 million tweets in one 24-hour period (Whitehead, 2015). In addition to *#JeSuisCharlie*, there were many other expressions articulated and shared via Twitter. The slogan *Je ne suis pas Charlie* ("I am not Charlie") came to represent myriads of opinions opposed to or critical of the mass *Je Suis Charlie* declaration. Another perspective was highlighted by the slogan *Je Suis Ahmed* ("I am Ahmed"), which referred to the French police officer Ahmed Merabet, who was Muslim and who was shot on the street by terrorists shouting, "Allahu Akbar" and "We have avenged the prophet". The slogan *Je Suis Ahmed* brought forth the perspective of French Muslims, who opposed the association between Islam and terrorism, as victims of the terror attack. The slogans were used on social media, in the news, and in demonstrations, and they were also circulated in images and caricature drawings emblematic of the case of *Charlie Hebdo*.

This brief illustration of the media workings in the *Charlie Hebdo* attacks is given here to demonstrate the hybrid nature of communication around the events and how an event can be transformed into a hybrid media event (Vaccari, Chadwick, & O'Loughlin, 2015). The term hybrid refers to a complex intermedia dynamic between mainstream news media and social media, as well as the complex circulations between messages and actors and the recombination of media on a variety of media platforms (Chadwick, 2013; Kraidy, 2002). Vaccari, Chadwick and O'Loughlin (2015, p. 1044) describe hybrid media events as "media events whose significance for media professionals, politicians, and non-elites is being reconfigured by the growth of social media". When thinking about the *Charlie Hebdo* attacks as a hybrid media event, we may approach it as a constellation of fluid social intensifications that are most typically created in a complex network of Internet-based and mobile communication technologies. The *Charlie Hebdo* attacks comprise elements of ceremonial mass media communication, but these also converge with contemporary forms of vernacular mass self-communication (cf. Castells, 2009), occasionally also thought of as a form of citizen journalism (cf. Allan & Thorsen, 2009). The element of "liveness" in the Charlie Hebdo attacks as a hybrid media event is intensified in the real-time circulation of texts and images and the dispersion of the event in several locations simultaneously. The level of connectivity between the official and viral narratives of the event may vary greatly, depending on the nature of a message in circulation. Hence, the concept of the "whole world" watching Paris needs to be analysed as an experience that is scattered onto a multiplicity of screens. While people may be sharing *Charlie Hebdo* as a collective spectacle—to use Kellner's (2003) terminology—they are connected to it in different ways. That is, they use different communication media to follow the event, associate with different—and even conflicted—narratives circulating on the event, and feel connected with different groups and identities involved in the event. Consequently, a multiplicity of shared ex-

periences is created in this hybrid media event. Thus, the question of power embedded in social integration as underlined in the classic theory of media events by Dayan and Katz (1992) needs to be addressed on several levels, including a variety of hybrid constellations of sociality (cf. Sumiala & Korpiola, 2016).

In the following sections, we suggest a multi-method approach for the empirical study of hybrid media events, using the *Charlie Hebdo* attacks as a case study. To meet this goal, we first provide a brief outline for our theoretical framework—the theory of media event—which is necessary to contextualise the methodological model. Secondly, we introduce a multi-method approach developed for the analysis of hybrid media events. In this approach, computational social science—or more specifically, a combination of automated content analysis (ACA) (Boumans & Trilling, 2016) and computational social network analytics (SNA) (Huhtamäki, Russell, Rubens, & Still, 2015)—is used in concert with a qualitative approach—specifically, digital ethnography. The article outlines three key phases for research in which the interplay between quantitative and qualitative approaches is played out. In the first phase a preliminary digital ethnography is applied to provide an initial sketch of the event. In the second phase, quantitative social network analytics is applied to construct the digital field for research. In this phase, it is necessary to map a) what is circulating on the websites and b) where this circulation takes place. In the third and final phase, a combination of the qualitative approach and digital ethnography is applied to provide a more nuanced, in-depth interpretation of what (substance/content) is circulating and how this material connects with the 'where' in the digital landscape, hence constituting links and connections in the hybrid media landscape. In conclusion, the article reflects on how this multi-method approach contributes to the understanding of the workings of today's hybrid media events—how they create and maintain symbolic battles over certain social imaginaries of solidarity, belonging, contestation, and exclusion. This is a topic of core value for the theory of media events.

2 Theoretical Framework: Re-Thinking Media Events

Since the birth of the modern mass media, many sociologists, cultural theorists, and communication scholars have examined the interplay between modern society and mass-media saturated gatherings (Bennet & Segerberg, 2012; Boorstin, 1973; Debord, 1967; Kellner, 2003; Rojek, 2013; Shils & Young, 1956). A key focal point in creating this tradition of thought in media studies is *Media events: The live broadcasting of history*, published by Daniel Dayan and Elihu Katz (1992). According to Dayan and Katz, a media event is a special genre that is powerful enough to interrupt everyday media flow, bring the viewer into touch with society's central values, and invite the audience to participate in the event (Dayan & Katz, 1992, pp. 5-9). In their lexicon, media events have their own grammar, their own meaning structure (story form or script), and their own practices, which are characterised by live broadcasting: the interruption of daily media rhythms and routines, the scripting and advance preparation of the event, a huge audience (the "whole world" is watching), social and normative expectations attached to viewing ("must see"), the ceremonial tone of media narration, and the intention to connect people.

As the story forms, media events can be divided into "conquests", "contests", and "coronations". According to Dayan and Katz (1992), these scripts constitute (i) the main narrative possibilities within the genre, (ii) the distribution of roles, (iii) and the ways in which these roles are enacted. In many cases, the three story forms are closely intertwined, and historical events correspond to and resonate with each other at different levels. One event may have certain features of each form; the form of an event may also change, transforming into another story form as the event develops. It is also important to acknowledge that all these scripts are embedded in deeper meaning structures in any given culture (Dayan & Katz, 1992, pp. 28-29). The common denominator for Dayan and Katz's (1992) original work is the ceremoniality associated with media performance. The authors indicate that the significance of media events is in their ability to reach a larger audience than any event that requires physical presence. The audience itself is well aware of this, as they follow the unfolding media event in different locations, which may be private, semi-public, or public.

Since its publication in the 1990s, the media events theory has stimulated vigorous scholarly debate, with its value believed to be in its theoretical and methodological innovation (Cottle, 2006; Couldry, 2003; Dayan, 2010; Fiske, 1994; Hepp & Couldry, 2010; Hepp & Krotz, 2008; Katz & Liebes, 2007; Kyriakidou, 2008; Liebes, 1998; Nossek, 2008; Roel, 2009; Rothenbuhler, 1998; Scannell, 1995, 2001; Sumiala, 2013). The main criticisms of Dayan and Katz's approach have addressed (i) the assumed ceremonial and integrative functions of media events, (ii) the attempt to exclude any disruptive or traumatic events from the focus of their theory, and (iii) the strong focus on television and broadcasting, which may result in inadequate study of global web-based media events.

In other words, many argue that Dayan and Katz's initial account of media events assumes too straightforward a relationship between media coverage and audience endorsement, thereby obscuring the ideological construction of social order, as well as the challenges and disruptive potential that are implicit in many media events (Cottle, 2006; Couldry, 2003; Fiske, 1994; Kellner, 2003; Kyriakdou, 2008; Rothenbuhler,

2010; Scanell, 1995, 2001). In addition, given the globalisation of communication through the Internet and social networking websites, critics have called for a recontextualisation of the explicit focus on TV and broadcasting in the media events theory.

Hepp and Couldry (2010, p. 9) argue that in theorising media events today we should not perceive them as placed at a defined locality, but rather as disembedded, or even ubiquitous, communicative practices. Drawing from the work of Hepp and Couldry, we postulate that today's media events should be understood as multi-sited, multi-temporal, multi-actor and multi-voiced phenomena articulated by a simultaneous connectivity of a variety of communication processes. These media events may be simultaneously structured around relatively centralised power structures, such as national and global mainstream media—for example the BBC or CNN—and multi-centred power structures, such as social networking sites (Hepp & Couldry, 2010, p. 9). Hepp and Couldry (2010, p. 12) offer a new working definition for contemporary media events to better grasp their fluid nature:

> "Media events are certain situated, thickened, centring performances of mediated communication that are focused on a specific thematic core, cross different media products and reach a wide and diverse multiplicity of audiences and participants."

Dayan and Katz have responded to the criticism of their original theory of media events and have readjusted their ideas in different public forums. Katz and Liebes (2007, 2010) suggest that the focus of analysis should be shifted from conquests, contests, and coronations to disaster, terror, and war. According to Katz and Liebes (2007, p. 157):

> "We believe that cynicism, disenchantment, and segregation are undermining attention to ceremonial events, while the mobility and ubiquity of television technology, together with the downgrading of scheduled programming, provide ready access to disruption. If ceremonial events may be characterized as 'co-productions' of broadcasters and establishments, then disruptive events may be characterized as 'co-productions' of broadcasters and anti-establishment agencies, i.e. the perpetrators of disruption."

Furthermore, Katz and Liebes suggest that marathons of terror, natural disaster, and war—media disasters—should be distinguished from media events as a separate genre. These mediatized disasters of different kinds have become far removed from the ceremonial roots of the original media events (Cottle, 2006; Liebes, 1997; Liebes & Blonheim, 2005). Daniel Dayan (2010) has written extensively about the changing nature of media events. For him, the "macabre accoutrements to televised ordeals, punishments, and tortures" and the emphasis on "stigmatization and shaming" in today's mediatized public events have caused media events to lose their potential to reduce conflict; instead, they 'foster divides, and install and perpetuate schisms' (Dayan, 2010, pp. 26-27). As a result, media events tend to lose their distinct character and instead migrate towards other genres: new media events are no longer clearly differentiated entities, but exist on a continuum. Dayan (2010, p. 27) suggests this 'banalization of the format' produces what he calls "almost" media events. Dayan reminds us that the pragmatics of media events have changed as messages have become multiple, audiences selective, and social networks ubiquitous. Dayan (2010, p. 27) summarises the difference between original and current media events in the following manner:

> "Interpersonal networks and diffusion processes are active before and after the event, mobilizing attention to the event and fostering intensive hermeneutic attempts to identify its meaning. But during the liminal moments we described in 1992, totality and simultaneity were unbound; organizers and broadcasters resonated together; competing channels merged into one; viewers gathered at the same time and in every place. All eyes were fixed on the ceremonial centre, through which each nuclear cell was connected to all the rest."

Dayan leaves the reader in a state of scepticism. For him, in today's "contested territory of media events", disenchantment and the loss of the "we"—the most critical functions of media events—are the most likely consequences. Although it is reasonable to ask whether this "we" ever existed, it is nevertheless inevitable that the dimensions of media events have changed with the changing media environment and the contemporary multiplicity of the media.

The concept of hybrid media events is one attempt to respond to the criticism offered by Hepp and Couldry (2010) and the response offered by Dayan (2010) and Katz and Liebes (2010). The idea of hybrid media events acknowledges the situated nature of transnationally or even globally mediated communication of a certain thematic core (here, the killings and related public reactions), while underscoring the fluidity of the movement in the circulation of the related posts, memes, images, news, and reports. The concept of the hybrid media event highlights the complex intermedia dynamics between the different media platforms (namely, mainstream news media and social media) in communicating those solidarities, belongings, and controversies associated with the event.

One of the key challenges for the study of contemporary media events is a methodological one. As the

media landscape changes and media events become more transnational and global, the right methodological tools need to be developed to better grasp these changing conditions. During the 1980s and 1990s in the empirical study of media events, the methodological focus was mainly on qualitative research. The empirical analysis focused on the study of national broadcast media, such as the BBC, or the national press, and the focus was on observation, textual analysis, and interviews related to the production, representation, and reception of media events (cf. Couldry, 2003; Couldry, Hepp, & Krotz, 2010; Eide, Kunelius, & Phillips, 2008).

While dividing the empirical focus between the production, representation, and reception of media events has proven a useful strategy for understanding national media events, this approach lacks the tools to analyse those communicative processes that go beyond the national frame and take place between and *betwixt* production, representation, and reception of media events. In these new conditions, messages, tweets, posts, memes, images, and symbols circulate and travel from one context to another. The categories between production, representation, and reception become blurred. It takes only one click to transform the person receiving a message into the one who produces it. As a result, new methodological approaches and tools need to be developed to capture these processes of communication that are crucial for today's hybrid media events. This suggests a new type of methodological dialogue between qualitative and quantitative approaches.

Here, the quantitative methods that make it possible to deal with a large amount of data circulating on a variety of media platforms are combined with more in-depth qualitative methods, such as digital ethnography, that enable researchers to go deeper into the data and trace pieces of meaning associated with symbolic battles carried out in the process of communicating about the events. In the following section, we introduce our methodological model for the study of hybrid media events with global appeal. This model, developed in the project on the 2015 *Charlie Hebdo* attacks in Paris, was created primarily for researching digital media—specifically Twitter.

3. Studying Hybrid Media Events on Twitter

Twitter is a micro-blogging website created in 2006 that enables users to send up to 140 character messages, commonly called "tweets". According to Twitter's own statistics from December 2015, the service has 320 million monthly active users (Twitter, 2016). Although the user growth has stalled in 2016—as Twitter is having its tenth birthday—it is still among the most popular social networking sites, along with Facebook and Instagram (Statista, 2016). On Twitter, messages are public by default, although the service also offers a feature called direct message (DM), which is private. On Twitter, the model of social relationships is directed and non-reciprocal, meaning that users can subscribe to other users' tweets in order to follow them. However, those they follow don't have to follow them back. When a user follows other users, the tweets of those followed will be visible on the user's main Twitter homepage, constituting a "tweet timeline" that appears in reverse chronological order. The characteristic practices for Twitter communication allow individual tweets to be liked and retweeted, which can increase the visibility and popularity of a single tweet. The retweet practice can also push a single tweet into a circulation that crosses the borders of different media platforms. Users can make a reference to other user with the @ symbol. With the prefix @ followed by a username, users can mention or reply to other users. An important feature is the hashtag—a word or phrase prefixed with the # symbol. Hashtags provide means for labelling tweets under certain topics, which gives structure to the communication on Twitter and enables users to find the information that interests them. Additionally, Twitter allows users to post images, videos, and hyperlinks.

As Twitter communication is limited to short messages that can be enriched with other communicative elements, such as images, videos, and hyperlinks, it is suitable for fast information sharing. Due to its public nature, it is popular among journalists, authorities, and organisations, as well as ordinary people. It is a prominent platform in the context and construction of different types of media events, varying from sports and politics to crises and disasters. A recent report on Twitter states that typical content on Twitter is twofold: either conversational, with thousands of people engaging with a particular topic for an extended period of time, or breaking news stories that drive large spikes in traffic over shorter periods of time (Parse.ly, 2016). Studies focusing on Twitter during political elections and sport events such as the Olympics give emphasis to idea of the audiences as co-producers of a media event, in addition to the traditional mass media (cf. Girginova, 2015; Kreiss, Meadows, & Remensperger, 2014). In the field of crisis communications, Twitter has been at the centre of many discussions. From the Arab Spring to the 2011 London riots, Twitter has been identified as a prominent platform for citizen communication in several revolutions, protests, and movements, as it connects people and bypasses the gatekeepers, whether they be the authorities or journalists (cf. Bennet & Segerberg, 2012; Procter, Vis, & Voss, 2013). From the journalistic viewpoint of crisis reporting, the 2010 Haiti earthquake has been called the first "Twitter disaster". This title underlines the fact that during the first 24 hours of the Haiti earthquake, news organisations were depending on social media, and especially the rapid and easily accessible flow of information provided by Twitter (Bruno, 2011). In times of crisis, ordi-

nary people can actively produce information, and they can also link and share published news stories from mainstream news media (Utz, Schultz, & Glocka, 2013; Parse.ly, 2016). In this context, Twitter has been perceived as a symbol of change in the media landscape:

> "If we allow ourselves to paraphrase the CNN effect of the 1990s, this changeover in the media landscape could be called the Twitter effect. As was true for the CNN effect, which was caused by more than just the CNN organization, the Twitter effect must also be considered as a symbol of a much broader phenomenon, concerning several online tools oriented to the publication of user-generated, real-time content (Twitter, Facebook, YouTube, etc.)." (Bruno, 2011, p. 8)

For our research, Twitter offers a fruitful context for the study of a hybrid media event. First of all, in contrast to Facebook, Twitter provides an Application Programming Interface (API) that allows access to the majority of the data published through the service (cf. Vis, 2013). Secondly, although we fully acknowledge that Twitter is only one platform in the hybrid media system, we state that it played a key role as a prominent platform during the unfolding of the *Charlie Hebdo* attacks, through which information, images, videos, and links about the incident were circulated. Thirdly, Twitter has become a key platform for breaking news, and therefore events that draw attention tend to surface on the platform. Finally, Twitter offers rich data that also sheds light on other forms of media. Several media organisations, politicians, and authorities use Twitter, and the content and actors from other media platforms are also present through a hypermedia chain (cf. Kraidy & Mourad, 2010). To give an example, a tweet that contains an image taken of the TV screen showing the news is a common convention that constructs a chain of different media.

4. Towards a Multi-Method Model

4.1. Automated Content Analysis and Computational Social Network Analytics

In our multi-method model, we combine computational social science—more specifically, automated content analysis (ACA) (Boumans & Trilling, 2016) in concert with computational social network analytics (SNA)—with a qualitative approach—particularly, digital ethnography. The computational approach allows for analysing both what is being said and by whom. Moreover, the individual actors can be connected to each other through their interactions for richer context to content, and this allows, for example, the identification of densifications in interaction between actors. More specifically, methods of automated content analysis allow us to identify the content that is circulating in the context of the hybrid media event under investigation. Social network analytics give us the means to investigate the overall structure between the actors that discuss and share content related to the event.

Here, the computational approach is used primarily to support digital ethnographic investigations. In terms of content analysis, the computational approach allows us to identify the key topics that are discussed in the data collected on the event. Four main approaches exist for automated content analysis: counting and dictionary, unsupervised learning, semi-supervised learning, and supervised learning (Boumans & Trilling, 2016; Laaksonen, Nelimarkka, Tuokko, Marttila, & Kekkonen, 2015). In its simplest form, automated content analysis is implemented by counting the number of times individual keywords or, in Twitter's case, hashtags and usernames, are included in the data. Unsupervised learning allows, for example, the creation of content-based clusters from the data to identify topics and their combinations or, in other words, to "identify potentially significant fragments" (Procter et al., 2013). In supervised learning, part of the data is categorised manually, and this learning data is used to teach an algorithm to categorise the rest of the material according to its category. Examples of approaches for automated content analysis include keyword extraction, topic modelling, natural language processing (NLP), and entity recognition (Boumans & Trilling, 2016; Finkel, Grenager, & Manning, 2005).

Compared to the situation that Procter et al. (2013) faced when they started mining tweets and found that there was very little existing infrastructure to support them, the availability of tools supporting analysis has improved over the last few years. Online services and social media analysis platforms, including Pulsar and others, provide investigators with dashboards that are able to manage millions of tweets. Using such environments for research is, however, far from trivial. Transparency of data and analysis routines remains a key issue. For ethnographic research, this limitation is not as major, as the investigation is done first and foremost on a qualitative basis, and therefore representative sampling is not a major issue. It is, however, important for the ethnographic research to understand what, in fact, is "the field" where the research takes place. This can be a problem when using commercial analytics services, as, due to technical and business restrictions, it is not always possible to gain the necessary information on how the data has been obtained.

The key approach into the analysis of structure that emerges from the interaction between individual actors in the data is social network analysis (SNA). Here we follow the insight of Yang and Leskovec (2014, p. 1892) as they maintain that, "networks provide a powerful way to study complex systems of interacting objects". SNA supports investigators in observing latent

structures and patterns in source data and in sharing their findings with others (Freeman, 2000). When analysing communication networks, actors are represented as network nodes and connected to each other through interactions. Network analysis allows us to quantify both structural properties of networks, as well as the structural positions of individual actors. Moreover, cluster identification can be used to identify groups of nodes that are interconnected to each other.

Network-level metrics come into play when individual network representations are compared to each other. Moreover, network metrics support the temporal analysis of network structure. Size, connection count, density, diameter, and average path length are examples of metrics that can be used as indicators as to which way a network under investigation is evolving. In investigating hybrid media events, one can, for example, create network representations of interactions that are related to a particular topic (identified using automated content analysis) and use network-level metrics to compare the properties of these topic networks.

Cluster identification is a particularly useful method for supporting early exploration of communication data. Clusters emerge from the topology of the network and challenge the investigators to make sense of why a particular cluster emerges. To support the sense-making process, the investigators can use the cluster membership to volumes of hashtags and other topic identifiers and therefore name or label the clusters according to their content signature.

Node-level metrics can be used for a number of purposes. Nodes with a high "betweenness" value, for example, are likely to act as bridges or boundary spanners (Hansen, Shneiderman, & Smith, 2011, "bridge scores for boundary spanners") connecting the different clusters of the overall network. Nodes with a high "in-degree", receive attention from the other actors. Nodes with a high "out-degree" are active in producing new content. Closeness centrality allows us to make a distinction between nodes with peripheral position and those close to the core of the network.

4.2. Digital Ethnography

Concepts such as digital ethnography, virtual ethnography, web ethnography, netnography, mobile ethnography, ICT ethnography, and virtual ethnography have emerged to describe fieldwork conducted in digital environments and landscapes (Boellstorff, 2008; Hine, 2015; Kozinets, 2015; Wittel, 2000). Online access to vast amounts of archived social interactions, along with live access to the human beings posting, changes the practice of ethnography. Researchers of the media are not dealing merely with words, but with images, drawings, photography, sound files, edited audiovisual presentations, website creations, and other digital artifacts (Kozinets, 2015, p. 4).

A characteristic of this qualitative methodological approach is that the researcher conducts fieldwork in the digital environment and applies participatory observation as a means to analyse human–technology interactions in the media and the social and cultural implications this interaction has for the present day digitalized life. In more practical terms, a digital ethnographer constructs his/her field by following ortracing the event, phenomenon, oractivity in question. The fieldworker makes notes, keeps field diary, takes screen shots, downloads material, and he or she may also interview informants by meeting them face-to-face or via digital communication media. It is not untypical that digital or online ethnography is combined with offline ethnography (cf. Postill & Pink, 2012).

Ethnographic understanding of the digital environment and its related interactions aims at in-depth, holistic, and situational understanding and knowledge of the studied event, phenomenon, activity and people (Hine, 2015, pp. 2-3). Considering the global, fluid, and continuously changing nature of the digital landscape, the issue of proximity and situational knowledge also needs re-framing. As Hine (2015, pp. 3-4) argues:

> "When we watch a fight break out on Twitter we cannot be sure whether any of the followers of those involved are seeing the same fight, at the same time, and understanding it in the same way that we do…The very notion of singular 'situation' as a pre-existing object breaks down when we look closely…An ethnographer in such circumstances must get used to a perpetual feeling of uncertainty, of wondering what has been missed, and attempting to build interpretations of events based on sketchy evidence."

In digital ethnography, the researcher has to deal with his/her limited human capacity to encompass the whole of the situation. For this challenge, computational social science offers valuable tools to map the digital landscape and provide a broader frame for the communicative and social processes taking place in that landscape. The value of ethnographic thick description and situational understanding lies in the depth, detail and the ability to grasp more profound layers of meaning in those actions and activities taking place in Twitter and elsewhere in digital media. To follow Hine (2015, p. 5):

> "Ethnography is highly necessary for understanding the Internet in all its depth and detail, and yet it can be challenging to develop way of conducting ethnographic studies which both embrace all that mediated communication offers and still provide us with robust, reliable insights into something in particular."

5. Three Phases

In the following, we will explain in more detail how the computational social science methods, automated content analysis (ACA) and social network analytics (SNA), can be combined with digital ethnography, and how this methodological interplay contributes to developing a new multi-method model for the study of such media events. This method has three phases:

1) Digital ethnography provides the first outline of the event;
2) Automated content analysis and social network analytics construct the digital field for research;
3) Digital ethnography provides an in-depth interpretation of what (substance/content) is circulating and how this material connects with the 'where' in the digital landscape, hence constituting links and connections in the hybrid media landscape necessary for the social meaning making of the event.

5.1. Digital Ethnography Sketches the Event

Like traditional media events, hybrid global media events interrupt the daily routines of the media and of the everyday. In the case of disruptive events, not only the mainstream news media, but also the social media environment turns to a disaster mode and begins to broadcast and circulate news, comments, tweets, posts, and images on the events as they unfold. This moment of massive media saturation and circulation of information produces the first methodological challenge for the study of hybrid media events. This first phase of chaotic information flow demands a digital ethnographic scope—a perspective in which the events are followed and structured into a timeline. In the case of *Charlie Hebdo* attacks, we started our pilot study immediately as the events unfolded. As digital ethnographers, we traced the news in the mainstream media, such as the *BBC*, the *New York Times*, the *Guardian* and *Le Monde*, as well as on Twitter, YouTube, and Facebook. Our personal media streams also included national news outlets, as well as friends and family members located in our native Finland and in different parts of the world, reporting and commenting on the events from different local perspectives. We identified certain prominent messages, hashtags, posts, memes, and images circulating in those media environments. To give one example, the hashtag #JeSuisCharlie was soon announced as the most-tweeted message in the history of Twitter, offering a simple and interesting lead to be followed in the course of the events. This first ethnographic phase of the analysis is best described as suggestive, and its findings may well be challenged in the later process of quantitative and qualitative analysis. Yet, it is a necessary stage for the process to follow, as it is this first stage of the project in which the chaotic information flow around the events is given its first suggestive sketch. This phase provided insight to what might be interesting, relevant, and peculiar in the events as they evolve and, thus, direct the analysis in the next phase. As a concrete way of gathering data, this phase results in many field notes, screenshots, memes, images, videos, and links, as well as a timeline of the events.

5.2. Using Automated Content Analysis and Social Network Analytics to Map the Field

In the next phase, social network analytics are applied to draw a more general overview of communication around the events with more data. In the case of the Charlie Hebdo attacks, the media platform analysed was Twitter. In this so-called "helicopter stage" of the analysis, social network analytics are used to construct the research field and give an overview of the data as well as map certain elements considered relevant based on the first phase of the pilot study. Prior to the analysis, the data needed to be collected. In this case, it was acquired through the social media analytics platform Pulsar using several search words[1]. The number of hits for #JeSuisCharlie totalled 2.3 million.

At the second stage, it is important to make a distinction between *what* is circulating in Twitter and *where* this circulation is taking place. In the *Charlie Hebdo* attacks project, we began with the hashtag #JeSuisCharlie and identified certain key groups: actors including ordinary media users, professional media houses; sites such as connected media platforms, countries and connections associated with it - both communication and non-communication between the different virtual communities created around this particular hashtag. As a result, this mapping can be further expanded to identify hashtags and actors that are related to #JeSuisCharlie. This mapping helps us to empirically illustrate communicative networks created around the events—where and when they take place and how they exist in relation to each other.

Human-in-the-loop analysis is particularly important when ethnographic and computational methods and approaches are used together. Therefore, we point to the Ostinato Model (Huhtamäki et al., 2015) for a structured process for data-driven visual network analytics that allows for balancing between exploration and automation (i.e. reproducibility) of analysis. This way, a multidisciplinary group of investigators can develop the rich description of a hybrid media event in an iterative and incremental fashion through a process that resembles peeling an onion and, thus, to begin to quan-

[1] The list of search words applied is the following: je suis charlie, #jesuischarlie, je ne suis pas charlie, #jenesuispas charlie, je suis ahmed, #jesuisahmed

titatively identify what is circulating (which hashtags) and in which digital media landscapes this circulation takes place. Starting with Twitter helps us to follow the circulation of certain hashtags and actors tweeting and re-tweeting onto new networks of communication (for example, Facebook, Instagram, or online media sites such as Huffington Post have been identified).

However, it must be noted that while collecting data from Twitter is relatively straightforward, given that the investigative team has the required technological capabilities, hybrid media introduce a major issue into data collection. The two public APIs that Twitter offers for developers, response-request based REST API and real-time streaming API, only allow data collection at the time it is published on Twitter. REST API allows the collection of limited amounts of data dating back to a number of days, and the streaming API operates in real-time by definition. The only way to collect extensive data on *Charlie Hebdo*, for example, is to acquire (buy) the data either directly from Twitter or through a social media listening service such as Pulsar. The data ecosystem has transformed since Procter et al. (2013) conducted their research regarding the 2011 London riots. Importantly, Twitter acquired Gnip[2] in 2014 that is currently the only company through which Twitter data can be purchased.

5.3. Applying Digital Ethnography in Tracing the Social

In the final stage of the empirical analysis, networks mapped by using quantitative analysis and social network analytics and its visual illustrations are taken into an ethnographic reconsideration. The quantitative analysis draws a map of the field and helps to orientate the ethnographic immersion. After choosing an interesting incident within the larger event, this case is followed in and through different media platforms. This phase aims to develop a holistic understanding of the chosen research object. Thus, the fieldwork in a digital landscape integrally involves a dense description of the observations in the form of field notes as well as documentation and recording of data by any means available, such as screenshots and prints (cf. Sumiala & Tikka, 2013). In order to capture the research object in a highly complex and dynamic landscape, it is useful to go back to the timeline of the events and re-evaluate the first sketch of the events against the quantitative framework and, consequently, make necessary re-orientations. In this phase, the researcher needs to re-evaluate the incident's relationship with the larger event and the key nodal points in this process. This can be carried out by searching for facts connected to the events and identifying certain key elements such as time, place, and people. In the digital landscape, this can be challenging as hybrid media events host and entice myriads of interpretations, misunderstandings, rumours as well as intentional misinformation. After re-locating the basic elements in the event, the researcher can begin to add layers of meanings to the event. This can be done in two overlapping ways; it is possible to conduct ethnographic fieldwork by following paths and trails of links, streams, and algorithmic suggestions offered by Twitter and other social media platforms, but it may also be useful to conduct digital ethnography by approaching the event simultaneously from different directions for example by making searches in search engines. In these overlapping processes digital ethnographer develops a more nuanced and in-depth understanding of the event and can begin to make interpretations of those more or less visible and hidden representations, discourses, actors and symbols and related communicative practices that contribute to creating and maintaining different types of social imaginaries of solidarities, belongings, and exclusions embedded with the events.

6. Conclusions

In this article, we suggest a new multi-model methodological approach to the study of hybrid media events developed for the study of the *Charlie Hebdo* attacks. In this new condition of hybrid media, events, messages, tweets, posts, memes, images, and symbols spread simultaneously and are constantly on the move. Old hierarchies between the centre and peripheries in the media event need to be reconsidered as hybrid media events appear more horizontal and multi-sited, multi-temporal, multi-actor, and multi-voiced social phenomena.

As a result, new methodological approaches and tools need to be developed to capture these processes of communication and better understand the workings of today's hybrid media events. This multi-method model proposed in the article consists of combining quantitative automated content analysis (ACA) and social network analytics (SNA) with qualitative digital ethnography. The key for the model is a close interplay between the different approaches and their careful adaption in the different phases of the research. This offers a unique possibility to bridge the gap between situational, in-depth knowledge achieved by qualitative methods in the study of media events and their understanding in the more global communication context.

The theory of media events was first established to explain the social dynamics activated as people gathered together around their TV sets to watch national rituals as live history to be performed on the screen. As discussed earlier, later developments in this theory have challenged the assumed social cohesion created by these events and emphasised instead the disruptive nature of media events. This has implied a certain contested view on the issue of sociality. The hybrid charac-

[2] https://blog.twitter.com/2014/twitter-welcomes-gnip-to-the-flock

ter of contemporary global, disruptive media events such as the *Charlie Hebdo* attacks makes the issue of social integration even more complicated. The question of establishment and anti-establishment is also changed in the hybrid media system. A terrorist attack is as carefully crafted and designed as a coronation, despite aiming at disruption rather than integration. The hybridity of the media environment causes a situation where no individual actor is able to control the flows of information, attention and effect. Despite this, hybrid media events also represent and reproduce existing social solidarities and antagonisms. Continuity and change take place through the circulation of meanings. Twitter is a particular environment for the circulation to take place. Its specific properties contribute to a culture of circulation (Lee & LiPuma, 2002) that seems complex and dispersed. This complexity calls for the multi-method approach.

The question of social integration in media events is not only a theoretical one. It is important to ask how we should empirically study the social dynamics activated in hybrid media events. In this article we suggest a methodological model that has potential to move from one research scale to another. The wider scale observations of the *Charlie Hebdo* attacks as a hybrid media event suggest a multiplicity of social dynamics were activated during the events. Hence, it suggests an interpretation that emphasises the heterogeneity as well as the ephemerality of those social dynamics. To understand more profoundly what kind of meanings and interpretations are associated with those messages and actors circulating in the digital landscape, an ethnographic perspective is necessary. In the future, more empirical research is needed to grasp these complex dynamics of social imaginaries of solidarity, belonging, but also exclusion. The multi-method approach is one attempt to point to this direction.

Acknowledgements

This article is based on the research project *Je Suis Charlie*—The symbolic battle and struggle over attention, conducted in the School of Communication, Media and Theatre in the University of Tampere. This two-year-project (2015–2017) is funded by the Helsingin Sanomat Foundation. The writers wish to thank the anonymous reviewers for their insightful comments.

Conflicts of Interests

The authors declare no conflicts of interests.

References

Allan, S., & Thorsen, E. (Eds.) (2009). *Citizen journalism: Global perspectives*. New York: Peter Lang.

Bennett, L. W., & Segerberg, A. (2012). The logic of connective action. *Information, Communication & Society, 15*(5), 739-768.

Boellstorff, T. (2008). *Coming of age in second life. An anthropologist explores the virtually human*. Princeton: Princeton University Press.

Boorstin, D. J. (1973). *The image—A guide to pseudo-events in America*. New York: Atheneum.

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit. *Digital Journalism*, *4*(1), 8-23.

Bruno, N. (2011). *Tweet first, verify later? How real-time information is changing the coverage of worldwide crisis events*. Reuters Institute for the Study of Journalism. Oxford: University of Oxford.

Castells, M. (2009). *Communication power*. New York: Oxford University Press.

Chadwick, A. (2013). *The hybrid media system: Politics and power*. New York: Oxford University Press.

Cottle, S. (2006). Mediatized rituals: beyond manufacturing consent. *Media, Culture & Society*, *28*(3), 411-432.

Couldry, N. (2003). *Media rituals: A critical approach*. London: Routlegde.

Couldry, N., Hepp, A., & Krotz, F. (Eds.). (2010). *Media events in a global age.* Routledge: Abingdon.

Dayan, D. (2010). Beyond media events. Disenchantment, derailment, disruption. In N. Couldry, A. Hepp, & F. Krotz (Eds.), *Media events in a global age* (pp. 23-42). London: Routledge.

Dayan, D., & Katz, E. (1992). *Media events. The live broadcasting of history.* Cambridge, MA: Harvard University Press.

Debord, G. (1967). *Society of the spectacle.* Buchet-Chastel.

Eide, E., Kunelius, R., & Phillips, A. (Eds.), (2008). *Transnational media events: The Mohammed cartoons and the imagined clash of civilizations.* Göteborg, Sweden: Nordicom.

Finkel, J. R., Grenager, T., & Manning, C. (2005). *Incorporating non-local information into information extraction systems by Gibbs sampling* (pp. 363-370). Stroudsburg, PA: Association for Computational Linguistics.

Fiske, J. (1994). *Media matters: Everyday culture and political change*. Minneapolis: University of Minnesota Press.

Freeman, L. C. (2000). Visualizing social networks. *Journal of Social Structure*, *1*(1). Retrieved from http://www.cmu.edu/joss/content/articles/volume1/Freeman.html

Girginova, K. (2015). New media, creativity, and the Olympics: A case study into the use of #NBCFail during the Sochi winter games. *Communication & Sport*. doi:10.1177/2167479515576101

Hansen, D. L., Shneiderman, B., & Smith, M. A. (2011). *Analyzing social media networks with NodeXL: Insights from a connected world*. Burlington, MA: Morgan Kaufmann.

Hepp, A., & Couldry, N. (2010). Introduction: Media events in globalized media cultures. In N. Couldry, A. Hepp, & F. Krotz (Eds.), *Media events in a global age* (pp. 1-20). Abingdon: Routledge.

Hepp, A., & Krotz, F. (2008). Media events, globalization and cultural change: An introduction to the special issue. *Communications. The Journal of European Communication Research*, *33*(3), 265-272.

Hine, C. (2015). *Ethnography for the internet: Embedded, embodied and everyday*. London: Bloomsbury.

Huhtamäki, J., Russell, M. G., Rubens, N., & Still, K. (2015). Ostinato: The exploration-automation cycle of user-centric, process-automated data-driven visual network analytics. In S. A. Matei, M. G. Russell, & E. Bertino (Eds.), *Transparency in social media: Tools, methods and algorithms for mediating online interactions* (pp. 197-222). Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-18552-1_11

Katz, E., & Liebes, T. (2007). 'No more peace!' How disaster, terror and war have upstaged media events. *International Journal of Communication, 1,* 157-166.

Katz, E., & Liebes, T. (2010). 'No more peace!' How disaster, terror and war have upstaged media events. In N. Couldry, A. Hepp, & F. Krotz (Eds.), *Media events in a global age* (pp. 32-42). Abingdon: Routledge.

Kellner, D. (2003). *Media spectacle.* London: Routledge.

Kozinets, R. (2015). *Netnography: Redefined.* London: Sage.

Kraidy, M. M. (2002). Hybridity in cultural globalization. *Communication Theory, 12*(3), 316-339.

Kraidy, M., & Mourad, S. (2010). *Hypermedia space and global communication studies lessons from the Middle East*. Retrieved from http://repository.upenn.edu/cgi/viewcontent.cgi?article=1256&context=asc_papers

Kreiss, D., Meadows, L., & Remensperger, J. (2014). Political performance, boundary spaces, and active spectatorship: Media production at the 2012 Democratic National Convention. *Journalism, 16*(5), 677-595.

Kyriakidou, M. (2008). Rethinking media events in the context of a global public sphere: Exploring the audience of global disasters in Greece. *Communications. The Journal of European Communication Research*, *33*(3), 273-291.

Laaksonen, S.-M., Nelimarkka, M., Tuokko, M., Marttila, M., & Kekkonen, A. (2015). Data augmented ethnography: using big data and ethnography to explore candidates' digital interactions. In *Proceedings of the Empiricist's Challenge: Asking Meaningful Questions in the Age of Big Data Conference.* Mannheim: University of Mannheim.

Lee, B., & LiPuma, E. (2002). Cultures of circulation: The imaginations of modernity. *Public Culture, 14*(1), 191-213.

Liebes, T. (1997). *Reporting the Arab Israeli conflict: How hegemony works*. London: Routledge.

Liebes, T. (1998). Television's disaster marathons. A danger for democratic processes? In T. Liebes & J. Curran (Eds.), *Media, ritual and identity*, (pp. 71-84). London: Routledge.

Liebes, T., & Blondheim, M. (2005). Myths to the rescue: How live television intervenes in history. In E.W. Rothenbuhler & M. Coman (Eds.), *Media anthropology*, (pp. 188-198). California: Sage.

Nossek, H. (2008). 'News media'—Media events: Terrorist acts as media events. *Communications. The Journal of European Communication Research*, *33*(3), 313-330.

Parse.ly (2016). Does Twitter matter to news. The authority report. *Parsely.com* Retrieved from http://learn.parsely.com/rs/314-EBB-255/images/authority-report-10.pdf

Postill, J., & Pink, S. (2012). Social media ethnography. The digital researcher in a messy web. *Media International Australia, 145*, 123-134.

Procter, R., Vis, F., & Voss, A. (2013). Reading the riots on Twitter: Methodological innovation for the analysis of big data. *International Journal of Social Research Methodology*, *16*(3), 197-214.

Roel, P. (2009). Intense media coverage. *Communications. The Journal of European Communication Research*, *34*(1), 1-20.

Rojek, C. (2013*). Event power: How global events manage and manipulate.* London & Thousand Oaks: Sage.

Rothenbuhler, E. W. (1998). *Ritual communication: From everyday conversation to mediated ceremony*. Thousand Oaks: Sage.

Rothenbuhler, E. W. (2010). Media events in the age of terrorism and Internet. *The Romanian Review of Journalism and Communication, IV*(2), 34-41.

Scannell, P. (1995). Media events (review). *Media, Culture & Society, 17*(1), 151-157.

Scannell, P. (2001). Editorial. *Media, Culture & Society, 23*(6), 699-705.

Shils, E., & Young, M. (1956). The meaning of the coronation. *Sociological Review*, *1*(2), 63-82.

Statista. (2016). Leading social networks worldwide as of April 2016, ranked by number of active users (in millions). *Statista.com.* Retrieved from http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users

Sumiala, J. (2013). *Media and ritual. Death, community and everyday life*. London: Routledge.

Sumiala, J., & Korpiola, L. (2016). Tahrir 2011: Contested dynamics of a media event. In B. Mitu & S. Poulakidakos (Eds.), *Media events: A critical contemporary approach*. London: Palgrave McMillan.

Sumiala, J., & Tikka, M. (2013). Broadcast yourself—Global news! A netnography of the 'flotilla' news on YouTube. *Communication, Culture & Critique, 6*(2), 201-352.

Twitter. (2016). Twitter usage/company facts. *Twitter*.

Retrieved from https://about.twitter.com/company
Utz, S., Schultz, F., & Glocka, S. (2013). Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review*, *39*, 40-46.
Vaccari, C., Chadwick, A., & O'Loughlin, B. (2015). Dual screening the political: Media events, social media, and citizen engagement. *Journal of Communication, 65*(6), 1041-1061.
Vis, F. (2013). A critical reflection on big data: Considering APIs, researchers and tools as data makers. *First Monday, 18*(10). doi:10.5210/fm.v18i10.4878
Whitehead, T. (2015, January 1) Paris Charlie Hebdo attack: Je Suis Charlie hashtag one of most popular in Twitter history. *Telegraph.* Retrieved from http://www.telegraph.co.uk/news/worldnews/europe/france/11336879/Paris-Charlie-Hebdo-attack-Je-Suis-Charlie-hashtag-one-of-most-popular-in-Twitter-history.html
Wittel, A. (2000). Ethnography on the move: from field to net to Internet. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 1*(1).
Yang, J., & Leskovec, J. (2014). Overlapping communities explain core-periphery organization of networks. *Proceedings of the IEEE*, *102*(12), 1892-1902.

About the Authors

Johanna Sumiala (University of Tampere) is Associate Professor (Docent) at the Department of Social Research/Media and Communication Studies, University of Helsinki. She is an expert in the fields of media sociology and media anthropology, virtual ethnography and visual culture. She has recently published articles on mediatized violence and ritualized (online) communication and she is the author of the book *Media and Ritual: Death, Community and Everyday Life* (Routledge, 2012).

Minttu Tikka (University of Tampere) is a PhD candidate at the Department of Social Research/Media and Communication Studies, University of Helsinki. She has published issues of crisis and social media, YouTube and news and methods of studying internet.

Jukka Huhtamäki (Intelligent Information Systems Laboratory at Tampere University of Technology, TUT) is a co-founder of Innovation Ecosystems Network. He is specialized in developing methods for extracting, accessing and analyzing heterogeneous data for system-level insights on various kinds of networked phenomena.

Katja Valaskivi (University of Tampere) is the Research Director of Tampere Research Centre for Journalism, Media and Communication (COMET). She is a media scholar, who has published widely in the issues of media and the nation, nation branding, media and social theory, and circulation in media society. The title of her most recent book is *Cool Nations: Media and the Social Imaginary of the Branded Country* (Routledge 2016).

Media and Communication (ISSN: 2183-2439)
2016, Volume 4, Issue 4, Pages 109-117
doi: 10.17645/mac.v4i4.578

Article

Participation in Social Media: Studying Explicit and Implicit Forms of Participation in Communicative Social Networks

Mikko Villi [1,*] and Janne Matikainen [2]

[1] Department of Communication, University of Jyväskylä, 40014 Jyväskylä, Finland; E-Mail: mikko.villi@jyu.fi
[2] Department of Social Research, University of Helsinki, 00014 Helsinki, Finland; E-Mail: janne.matikainen@helsinki.fi

* Corresponding author

Submitted: 31 January 2016 | Accepted: 2 May 2016 | Published: 10 October 2016

Abstract
The diverse forms of participation in social media raise many methodological and ethical issues that should be acknowledged in research. In this paper, participation in social media is studied by utilising the framework of explicit and implicit participation. The focus is on the communicative and communal aspects of social media. The aim of the paper is to promote the reconsideration of what constitutes participation when online users create connections rather than content. The underlying argument is that research on social media and the development of methods should concentrate more on implicit forms of participation.

Keywords
connectivity; ethics; explicit participation; implicit participation; methods; social media

Issue
This article is part of the issue "Successes and Failures in Studying Social Media: Issues of Methods and Ethics", edited by Epp Lauk and Niina Sormanen (University of Jyväskylä, Finland).

© 2016 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

Participation in social media is multifaceted and, thus, methodologically challenging. In this paper, we shed light on the complexities of social media research by means of a literature review, building on the theoretical framework of explicit and implicit participation. The review of explicit and implicit participation can provide insights for research and methodology in the context of social media. In particular, we suggest that the review can aid in developing methods for studying implicit participation in social media.

First, it is necessary to examine the general concept of participation in social media. The word 'social' in the term 'social media' implies that these platforms facilitate communal activities. In the same sense, the term 'participatory' emphasises human collaboration (van Dijck, 2013, p. 11). However, all participation is not the same. Participation is an ambiguous concept that can take on many different forms, and it can even be seen as a problematic concept that is often overused and overstretched (Carpentier, 2011, pp. 351-353). One approach is to stress participation as a political-ideological concept that is intrinsically linked to power. Carpentier (2011) claims that participation is deeply embedded within our political realities and that it is therefore the object of long-lasting, intense ideological struggles. Carpentier also considers participation as structurally different from interaction. According to him, interaction remains an important condition of participation, but it cannot be equated with participation. Unlike participation, interaction has no such political meanings because participation entails power dynamics and interaction does not (Matikainen, 2015, p. 43).

Many other views on participation are less political and less occupied with notions of power; therefore, they may be closer to interaction. According to Jenkins (2006), participatory culture contrasts with older notions of passive media spectatorship and consumption. Rather than talking about media producers and con-

sumers as occupying separate roles, we might see them interacting and collaborating. Jenkins refers to this as convergence culture. From this perspective, participation is associated more with acting together and communicating with each other and less with power, politics and struggle, although these elements still play a role.

The action within inter*action* is of essence. Former passive audiences turn into active participants and agents of cultural production on the Internet (Schäfer, 2011, p. 10). Thereby, audiences become participants or users. The term 'users' is often used to refer to people who are more or less active online. In this sense, 'users' is a better term than 'audience' to indicate interactive participants in social media. Being an audience only captures one segment of the contemporary media experience because an increasing part of online activity involves 'doing' things, such as messaging, sharing, tweeting, using Facebook, chatting, commenting, editing, posting and uploading (Merrin, 2009, p. 24).

Livingstone (2005) states that 'users' as a term does not work very well either. It is too broad, and it has less to do with communication. Users use computers when engaging online, but they also use, for example, washing powder, which does not have much to do with human interaction, communication or communality. Thus, Livingstone proposes that 'people' would be a good term to describe participants in the realm of social media.

It is important to acknowledge that being a user and being an audience member are not mutually exclusive roles. People can be online users and audience members simultaneously. The roles are not always defined by the level of activity, but the role can also be a question of attitude, i.e. how people themselves consider their positions in social media. In addition, the roles can depend on the platform, i.e. what kind of affordances different online platforms offer for interactivity. However, it is reasonable to question the significance of novel communication technology and platforms for participation in general. The presumption that new networked technologies lead to participation is rather generalising (Terranova, 2004; van Dijck, 2009). Nevertheless, users of 'old media' are often stereotyped as a passive audience, and users of digital, participatory media are equally stereotyped as hyperactive, co-creative people (Olsson, 2010, p. 101).

2. Explicit and Implicit Participation

To further examine participation, we will focus on explicit and implicit forms of participation (Schäfer, 2011). From the users' standpoint, explicit participation involves producing media texts and artefacts. Explicit participation has a 'pro' aspect, in the sense that it is about production, produsage and prosumerism and is closely linked with user-generated content (UGC). Explicit participation is also connected to co-creation when consumers are no longer satisfied with their traditional end-user roles, and they want to be involved in creating and developing digital products and services (Ramaswamy & Gouillart, 2010, pp. 3-6).

In contrast to explicit participation, implicit participation does not involve conscious production. Rather, it often comprises unacknowledged labour or even unconsciously performed labour. Implicit participation shows a trend towards automated user participation because it is channelled through easy-to-use interfaces and the automation of user activity processes (Schäfer, 2011, pp. 51, 78.). In implicit participation, the behaviour patterns, interests and consumption habits of users are turned into valuable data. For example, implicit participation occurs when a user connects his/her Spotify account with Facebook and information about music consumption is thereafter automatically communicated to his/her Facebook friends (Karppi, 2014, pp. 37, 72). Connections between users are essential to implicit participation. If UGC represents explicit participation, implicit participation is represented by user-generated behaviour (UGB) (Netzer, Tenenboim-Weinblatt, & Shifman, 2014, p. 625).

Most social media platforms are created to generate data about users for improved information management and targeted advertising (Schäfer, 2011, p. 78). In implicit participation, the actions of users and producers (as in produsage, see Bruns, 2012) do not necessarily blur; instead, the actions of users and information technology blur because the labour is performed by both the information system and the user (Schäfer, 2011, p. 78). An obvious example of this is Facebook. The social media platform connects users and technology (e.g. interface and platform infrastructure) to gather data from users. In this sense, implicit participation is closely related to the concept of connectivity (van Dijck, 2013), which we will return to shortly.

Implicit forms of user activity involve sustaining connections and togetherness rather than producing content. In this sense, implicit participation is connected to the ritual view of communication as defined by Carey (1989). Ritual communication concentrates on relationships between people and serves to maintain communality and community. The ritual view of communication exploits the mutual roots of commonness, communion, community and communication (Carey, 1989). However, many people might object to calling implicit online activities participation, at least in the sense that Carpentier has approached participation.

We adapt Schäfer's classification by replacing the dual construction of participation with a continuum or dimension of participation. Activities such as writing a blog post or contributing to a Wikipedia article lie at the explicit end of the continuum. In contrast, at the implicit end of the continuum, participation is non-productive from the users' point of view. Popular social media activities, such as sharing and liking content, lie between explicit and implicit participation on the con-

tinuum; however, in our view, they are closer to implicit participation.

2.1. Users of Users

Participatory social media users can be regarded, among others, as consumers, producers, distributors or products. We will now focus on the latter two roles, which represent more-implicit forms of participation on the continuum. First, we will discuss how participation can be regarded as a product or commodity. One well-known view associated with social media is that 'if you are not paying for it, you are not the customer, you are the product' (Noguera Vivo, Villi, Nyirő, de Blasio, & Bourdaa, 2013). From this view follows the argument that social media companies are *users of users* because the users and the data they provide about themselves are the products social media companies use in their business.

Van Dijck (2009, 2013) writes about this in a convincing manner. Many social media companies are less interested in selling paid memberships than they are in customers who do not pay anything for services. In exchange for free services, social media companies require the use of users' data. According to van Dijck (2009, p. 49), a user's role as a data provider is infinitely more important to social media companies than the user's role as a content provider. Personal data are turned into public, observable connections, such as liking something on Facebook. Privacy is the currency in social media.

Users profile themselves, e.g. by liking and using social plugins, while social media companies, which can also be called corporate social media (CSM), sell the resulting data to marketers and advertisers (Gehl, 2015). Likes and social plugins create new forms of connectivity between websites, generating a 'like economy' in which the acts integrate more objects into the social graph and enable social media companies to expand their activities throughout the Internet to monetise connections and data flows outside their own platforms (Gerlitz & Helmond, 2013; Langlois & Elmer, 2013, p. 9). Users willingly and often unknowingly provide important information about their profiles and behaviours to CSM companies (van Dijck, 2009, p. 47). This view is very much in line with Schäfer's conception of implicit participation. It can be argued that social media services are built on a double exploitation of explicit and implicit participation. First, social media companies thrive on the free labour of content creators (Terranova, 2004). Second, they exploit the labour of all site visitors, who generate valuable data about their interests and activities as they like and share content. This more-implicit user participation produces data that are valuable to marketers, who want to know what people are interested in and how those interests link to other interests (Gauntlett, 2011, p. 191).

In other words, in the context of participatory culture, users are a commodity sold by the media as eyeballs to advertisers, in the classic view of Smythe (1977), and also the communal activities of the audience, such as liking or distributing news stories, are a commodity (Fuchs, 2010). Social media users represent a labour power that produces attention to ads as well as data, which then feeds into the production of ads (Fisher, 2015, p. 65). Much of the user work or user labour in social media is associated less with production and more with communication and interaction. This also means that user work is described well by implicit forms of participation.

In relation to implicit participation, the role of connectivity has been emphasised as part of social media logic (van Dijck, 2013). Connectivity refers to how social media companies gather data from users and their connections and then exploit this data for their purposes, e.g. by selling that information to advertisers (van Dijck & Poell, 2013, p. 8). When engaged in social media, users are positioned within an algorithmic connectivity in which machine processes mine their data, which are then used to affect and engage other users (Karppi, 2014). Users produce a precious resource: connectivity (van Dijck, 2013, p. 16).

The term 'connectedness' is closely related to connectivity and implicit participation. In a sense, connectedness involves social media users connecting with each other in a multitude of ways. Commoditising relationships and connections to turn connectedness into connectivity is the essence of the business undertaken by social media companies (van Dijck, 2013). CSMs are less interested in co-creation or UGC than in users making connections, which yield valuable information about who the users are and what they are interested in (van Dijck & Nieborg, 2009, pp. 865-866). Accordingly, social media platforms record not only what is being said but also the more broad act of communication itself, including information about the profile of a user sending out a message, the users receiving that message and how users interact with a message by reading, liking and sharing it or not (Langlois & Elmer, 2013, p. 2).

Importantly, social dynamics on social media platforms depend on the platforms themselves (Gillespie, 2015; Helmond, 2015). Social media platforms have simplified the communication process and expanded potential communicative opportunities, but they have also been built to harness communication in an effort to monetise it (Langlois & Elmer, 2013, p. 2). Social media platforms are designed to enhance human connections and constant connectivity because smaller friendship networks and less communication would lead to less user data to market (Langlois, 2014, p. 7; Langlois & Elmer, 2013, p. 10). As Gillespie (2015) notes, 'Platforms matter…The platforms, in their technical design, economic imperatives, regulatory frame-

works, and public character, have distinct consequences for what users are able to do, and in fact do'.

2.2. Sharing in User Communities

The context for connectedness can be labelled a 'user community'. When users in social media regularly communicate among themselves, they can be said to form an actual community; otherwise, impermanent, transient and atomised user structures simply represent a crowd (Bruns, 2012, p. 819). The traditional definition of a community, i.e. a spatially compact set of people with a high frequency of interaction, interconnections and a sense of solidarity, does not always fit well with user communities. Social media services in particular do not necessarily make people feel as if they belong to a community. Instead, social media can host sets of interlinked personal communities built on interpersonal commitments (Gruzd, Wellman, & Takhteyev, 2011, pp. 1296-1299). One way to illustrate a user community is to view it as connections between personal communities formed around friends, followers and other contacts in social media. Social networking services do not form mutually shared communities; instead, interaction takes place in many interlinked personal communities, such as one's Facebook friends. Overall, connectedness is significantly related to communality.

On the continuum of participation, sharing content can be placed at a position closer to implicit participation than to explicit participation. Sharing in social media involves connectedness and connectivity. Those who share content connect with their peers while also connecting with Facebook, Twitter and Google (van Dijck, 2013, p. 45). In general, the consumption of media content has gone from being an individual activity to one in which consumers have the opportunity to interact with others. In a sense, the consumption of content is a social experience (Villi, 2012, p. 627) and a networked practice (Jenkins, 2006, pp. 244–255). Thus, sharing is becoming central to the way in which people experience media content (Hermida, Fletcher, Korrell, & Logan, 2012, p. 7), such as when they read news stories recommended by others on Facebook. The word 'sharing' fittingly describes participation in social media at large: sharing can be an act of distribution, communication and consumption (Belk, 2010, p. 730; John, 2013). Sharing involves social exchange on the one hand and distribution and dissemination on the other (Wittel, 2011, pp. 3, 8).

We will now focus on sharing in the context of social media specifically from the viewpoint of user-distributed content (UDC). UDC describes the role of horizontal connections in disseminating media content; users take part in distribution processes by enclosing content with social relations (Villi, 2012; Villi & Matikainen, 2015). Users make personal referrals and guide their peers to consume content that they consider interesting and relevant by recommending the content on Facebook, tweeting links or sending messages through email or instant messaging applications. UDC involves sharing as a method of distribution. UDC refers to audience activities that serve to amplify the extent, visibility and impact of existing online media content. In this sense, it is possible to draw clear boundaries between UDC and user-generated content (UGC), which clearly involves producing new content.

Terms and concepts that closely relate to UDC include superdistribution, i.e. forwarding media through social networks (Anderson, Bell, & Shirky, 2013, p. 14), and social curation, which illustrates the networked distribution of media content by adding qualitative judgements and imbuing the content with personal and social significance (Villi, 2012, p. 615; Villi, Moisander, & Joy, 2012). Singer (2014, p. 68) has coined the term 'user-generated visibility', which is very close to UDC. The idea of friendcasting (Lee & Cho, 2011) is implicit in these approaches to social media use. User-copied content (UCC) can also be linked to UDC (van Dijck, 2013, p. 119).

UDC has the advantage of not being a very time-consuming activity for users. In many ways, UDC represents 'easy participation' (Newman & Levy, 2013, p. 70), consisting of simple, everyday actions (Jenkins, Ford, & Green, 2013, p. 199) that represent a light version of participation, i.e. 'participation lite'. Using a social plugin, such as pressing the Facebook recommend button placed next to a news item or another story, is less demanding than writing a post. In UDC, personal recommendations by online contacts, especially by friends, are of utmost importance; having media content recommended by a friend is very effective when deciding what content to consume (Enda & Mitchell, 2013; Matikainen & Villi, 2013). Influencers—people with large networks of connected followers and friends—are also important (van Dijck, 2013, p. 40). The term 'alpha distributors' can be used to refer to users who act as key nodes in social media and whose recommendations have an extensive reach and influence.

UDC involves connectedness in the sense that it entails communicative interactions and relationships between two or more individuals. At the same time, it represents connectivity because users lure their peers to consume media content, which then leads to data generating audience traffic. Users who distribute content in social media also engage in the work of advertisers and marketers by endorsing brands and their content (Napoli, 2010, p. 512) and mobilising themselves in the promotion of brands (Fisher, 2015, p. 50).

2.3. Participation and the Media

Explicit and implicit forms of participation can also be examined in the context of mass media and journalism.

Until now, UGC has been the most-recognised form of audience participation in journalism (Noguera et al., 2013), and legacy media organisations' strategies concerning audience participation have concentrated largely on UGC (Napoli, 2010; Netzer et al., 2014; Singer et al., 2011). However, an increasing number of media scholars have argued that for legacy media organisations, engaging, encouraging and assisting the audience in the circulation of media content is more important than having them participate in content production (Hermida et al., 2012; Singer et al., 2011). Naturally, these two ideas are interlinked in the sense that content in social media often acts as a social glue that connects people, such as when users view photographs on Instagram.

Marshall (2009) argued that the successful operation of the media industry in social media is as much related to content production as it is to facilitating the maintenance of social connections among its audience. In other words, the work of editorial teams is transformed from content production through creating platform concepts to coordinating, managing and nurturing audience communities (Malmelin & Villi, 2015).

Jönsson and Örnebring (2011) offer a simple distinction between a) features that require a low level of participation in the media, mainly addressing audience members as consumers; b) features that require a medium level of participation, addressing audience members as prosumers; and c) features that require a high level of participation, addressing audience members as producers. An example of the low level of participation is the use of the social media audience as a 'radar' (Villi, 2012, p. 616). Within the radar, journalists can observe which stories and content circulate the most in social media and make content decisions accordingly. Another term used to describe journalists watching what users discuss online is 'public sensor' (Heinonen, 2011, pp. 37–38). When users are monitored, the simple act of liking or sharing content can influence decision-making processes in mass media; thereby, users can take part in news selection and gatekeeping (Netzer et al., 2014, p. 628). Moving from a low level to a high level of participation correlates with the transition from implicit to explicit participation.

2.4. Methodological Challenges of Studying Implicit Participation

Now, we move on to discussing the methodological challenges of social media research. In connection to the continuum of explicit and implicit participation, a fundamental methodological challenge is the tension between human action and the data-based environment. As van Dijck and Poell (2013) demonstrated, programmability is an essential element of social media logic. Sociality is produced on social media platforms, and their code, data, algorithms, protocols and interfaces impose preconditions for social interactions and social being. The corpus is not just the product of people's participation; it is also crafted by social media platforms according to the logic of their algorithms, the imperatives of their business models and the enforcement of community guidelines (Gillespie, 2015).

Here, we can again refer to Schäfer's (2011) claim that in implicit participation, a user's actions and information technology are strongly interlinked. Therefore, there is both a human–social aspect and a data-based aspect of interaction and participation. When studying implicit participation, both of these aspects should be considered. However, the aspects are methodologically and pragmatically different. In most cases, user data are quantitative, but participation is human and social in nature; the human–social aspect of participation emerges in meanings and discourses.

To understand this context, Lessig (2001) has developed a useful specification. He presents three distinct layers of the Internet. The first layer is the physical layer, which is the basis of communication, i.e. computers and the wires that link computers. The physical layer, i.e. the hardware, enables operations in information networks. The second layer is the code layer, which makes the hardware run. This layer includes protocols and code. The third layer is the content layer, which is the visible part of the Internet, including images, texts and videos. Lessig's specification of the three layers helps us understand that implicit participation and participation in general operate on three levels, and the content layer is the only one visible to users.

In social media, the power of algorithms, i.e. the code layer, is stronger now than it has ever been. As Bucher (2012) notes, algorithmic architectures dynamically constitute certain forms of social practices. For example, Facebook's algorithms are based on the assumption that users are not equally connected to their friends. This assumption has two methodologically important consequences. First, explicit and implicit participation are partly constituted by the algorithm. Second, the data produced by algorithms is relevant when studying participation in social media.

A further methodological challenge involves determining how to gather research data, which in social media could be rather extensive. The data can be either user-based or platform-based, following the continuum from explicit to implicit participation. Traditionally, research data in social sciences have been gathered through surveys, interviews, observations and documentation, which are generally manual methods. Social media and the Internet in general generate a need for new forms of data analysis and software-supported data capture (Rieder, Abdulla, Poell, Woltering, & Zack, 2015), such as network analysis. For example, Facebook produces diverse, broad data, which can be gathered using computational techniques (Sormanen et al., 2016), such as an application programming interface (API). API-

based data can highlight the role of administrators as connective leaders, measure the size and composition of the participating audience, engage in various types of periodisation and investigate issues that have been raised in comment sections (Rieder et al., 2015).

Even with new forms of data, interaction and participation in social media cannot be analysed fully using either qualitative or quantitative methods. The relationship between quantitative and qualitative methods is vague because they are based on different assumptions of reality and human action. The challenge is then to combine qualitative and quantitative methods in a sensible way. Quantitative and qualitative data also require various analytical skills, and multidisciplinary research is often necessary. In social media research, the lack of technical skills may lead to limited data and biased analysis (Weller, 2015). In addition, we should not become blinded by new data and methods. Big data and network visualisation can be impressive and imposing, but analytically, they are rather conventional. We as researchers must not only preserve analytical and critical thinking but also remain curious about new data and methods.

The challenge of data access and ownership of user data is of essence to research. As mentioned earlier, social media companies utilise user data as a product in their businesses. Thus, the data are commonly considered trade secrets. This causes many problems from a methodological perspective. At worst, research could be limited only to data that are freely available. This phenomenon is already evident, as shown in the amount of social media research focusing on Twitter due to the platform's comparatively easy and stable data access (Weller, 2015). Another point is that a majority of online participation is increasingly unreachable and hidden from researchers. The popularity of messaging apps, such as WhatsApp (Newman & Levy, 2014, p. 70), LINE and iMessage, in online communication is growing; however, data from these apps cannot be analysed in the same manner as traffic on Twitter.

2.5. Ethical Considerations

When conducting research on social media, ethical aspects should be considered. Implicit and explicit participation especially relate to the question of privacy, which depends on individual and cultural definitions and expectations (Markham & Buchanan, 2012, p. 6). These definitions are extremely complicated in the context of participation because users do not often perceive on their own whether their participation is implicit or explicit, or do not realise that their activity is participation at all.

As Sveningsson Elm (2009) notes, public and private are not univocal states; instead, they are based on users' perceptions. In this situation, researchers should be sensitive to users' rights and privacy. One essential notion is that like implicit and explicit participation, the concepts of public and private should be considered as a continuum and not a dichotomy (Sveningsson Elm, 2009). Many social media environments are semi-public, where some parts of the environment are public and some parts are not; in these cases, the definition of privacy is ambiguous. As Henderson, Johnson and Auld (2013, p. 550) stated, 'In the context of social media, it becomes increasingly difficult, if not impossible, to ethically claim a dichotomy of private or public'.

An important ethical aspect is the question of consent (Henderson et al., 2013, p. 549). If the research subject is identifiable, consent is needed. However, in the case of social media, from whom should consent be obtained and how should it be obtained? Consent is closely related to the above-mentioned dimension of private and public. Basically, consent is needed when content or communication is private. Because drawing the line between private and public in social media is challenging and sometimes even impossible, researchers should be extremely careful when gathering research data without consent. Often, the difficulty of obtaining consent adds to the challenge.

Research participants or subjects are regularly promised anonymity (Dawson, 2014). In social media, data are diverse because they also include information regarding people's preferences and social relationships, which is often a result of implicit participation. In relation to implicit participation, it should be noted that personal data in social media are also produced by platforms. Platform-based data can often reveal a user's identity easier than when the user makes an explicit decision to reveal or hide his or her identity.

One important question revolves around how researchers can penetrate social media, especially implicit forms of participation. Corporate social media companies are not shy about collecting information on implicit participation, and many users are not aware that they are leaving traces and collectable information. Can researchers be as straightforward and ethically blunt when digging into the forms and practices of implicit participation? How active as participants should the researchers themselves be? In relation to these questions, Markham (2013, p. 440) questions whether participation is always necessary for ethnographers working in online environments.

The ethical challenges are increased by the fact that children and young adults are remarkably active on social media. Scientific ethical rules emphasise that all humans, especially children, should be protected. It is interesting to note that despite many rules and guidelines related to research ethics, there are few empirical studies on ethics in online research (Dawson, 2014). Therefore, we suggest that in addition to concrete, detailed ethical rules concerning social media research, more meta-analysis of the fulfilment of research ethics in social media research is needed.

3. Conclusions

In this paper, we emphasised that implicit and explicit participation are the two ends of the continuum of participation. At the implicit end of the continuum lies fully automatic participation, as exemplified by the programmed publication of music listening habits on Facebook after listening to music on Spotify. At the explicit end lies writing a blog or publishing a picture on Instagram. Sharing and liking lie between the two ends of the spectrum, although they more closely represent implicit participation. Both explicit and implicit participation in social media are motivated at least partly by connectedness, and they benefit and support connectivity. Connectivity represents the relationship between the user and the information system or the platform (such as Facebook and Google), whereas connectedness represents the relationship between users.

We argue that the more-implicit forms of user participation in the communicative social networks of social media are of essence, especially for social media companies that offer services and platforms for interaction and that rely on connectivity in their businesses. We suggest that this is also increasingly the case for legacy media organisations. Therefore, we claim that when attempting to describe the instrumental participatory activities of users in a social media environment, discussion of UGC and other forms of more-explicit participation should be supplemented with discussion of more-implicit forms of participation. Accordingly, the development of methods should focus more on the study of implicit participation, which has not been the case in previous research.

The continuum of explicit and implicit participation should be considered as a theoretical, methodological, ethical and pragmatic phenomenon. Therefore, the elaboration of theoretical concepts is necessary but insufficient. We should remember that many methodological and pragmatic aspects can restrict the operationalisation of participation in social media and the nature of the data. For example, Langlois and Elmer (2013, pp. 10-11) argue that the promise that social media data offer a transparent trace of human behaviour is false because the behaviour is affected by the corporate logic of social media platforms. This leads to several challenges when studying modes of participatory culture on social media platforms, including access to data, the ethics of data research, the data, and what the data claim to stand for.

Conflict of Interests

The authors declare no conflict of interests.

References

Anderson, C. W., Bell, E., & Shirky, C. (2013). *Post-industrial journalism: Adapting to the present*. Retrieved from http://towcenter.org/wp-content/uploads/2012/11/TOWCenter-Post_Industrial_Journalism.pdf

Belk, R. W. (2010). Sharing. *Journal of Consumer Research*, *36*(5), 715-734. doi:10.1086/612649

Bruns, A. (2012). Reconciling community and commerce? Collaboration between produsage communities and commercial operators. *Information, Communication & Society*, *15*(6), 815-835. doi:10.1080/1369118X.2012.680482

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, *14*(7), 1164-1180. doi:10.1177/1461444812440159

Carey, J. W. (1989). *Communication as culture. Essays on media and society*. Boston: Unwin Hyman.

Carpentier, N. (2011). *Media and participation: A Site of ideological-democratic struggle*. Bristol, UK: Intellect.

Dawson, P. (2014). Our anonymous online research participants are not always anonymous: Is this a problem? *British Journal of Educational Technology*, *45*(3), 428-437.

Enda, J., & Mitchell, A. (2013). Friends and family—Important drivers of news. *The Pew Research Center's Project for Excellence in News Journalism*. Retrieved from http://www.stateofthemedia.org/2013/special-reports-landing-page/friends-and-family-important-drivers-of-news

Fisher, E. (2015). "You media": Audiencing as marketing in social media. *Media, Culture & Society*, *37*(1), 50-67. doi:10.1177/0163443714549088

Fuchs, C. (2010). Labor in informational capitalism and on the internet. *The Information Society*, *26*(3), 179-196. doi:10.1080/01972241003712215

Gauntlett, D. (2011). *Making is connecting: The social meaning of creativity, from DIY and knitting to YouTube and Web 2.0*. Polity Press: Cambridge, UK.

Gehl, R. W. (2015). The case for alternative social media. *Social Media + Society*, *1*(2). doi:10.1177/2056305115604338

Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*, *15*(8), 1348-1365. doi:10.1177/1461444812472322

Gillespie, T. (2015). Platforms Intervene. *Social Media + Society*, *1*(1-2). doi:10.1177/2056305115580479

Gruzd, A., Wellman, B., & Takhteyev, Y. (2011). Imagining Twitter as an Imagined community. *American Behavioral Scientist*, *55*(10), 1294-1318. doi:10.1177/0002764211409378

Heinonen, A. (2011). The journalist's relationship with users: New dimensions to conventional roles. In J. B. Singer, D. Domingo, A. Hermida, A. Heinonen, S. Paulussen, T. Quandt, . . . M. Vujnovic (Eds.), *Participatory journalism: Guarding open gates at online newspapers* (pp. 34-56). Chichester, UK: Wiley-Blackwell.

Helmond, A. (2015). The Platformization of the web: Making web data platform ready. *Social Media + Society*, *1*(2). doi:10.1177/2056305115603080

Henderson, M., Johnson, N. F., & Auld, G. (2013). Silences of ethical practise: Dilemmas for researchers using socila media. *Educational Research and Evaluation*, *19*(6), 546-560.

Hermida, A., Fletcher, F., Korrell, D., & Logan, D. (2012). Share, like, recommend: Decoding the social media news consumer. *Journalism Studies*, *13*(5-6), 815-824.

Jenkins, H. (2006). *Convergence culture*. New York: NYU Press.

Jenkins, H., Ford, S., & Green, J. (2013). *Spreadable media: Creating value and meaning in a networked culture*. New York: New York University Press.

John, N. A. (2013). The Social logics of sharing. *The Communication Review*, *16*(3), 113-131. doi:10.1080/10714421.2013.807119

Jönsson, A. M., & Örnebring, H. (2011). User-generated content and the news: Empowerment of citizens or interactive illusion? *Journalism Practice*, *5*(2), 127-144. doi:10.1080/17512786.2010.501155

Karppi, T. (2014). *Disconnect.Me: User engagement and Facebook*. Turku: University of Turku.

Langlois, G. (2014). *Meaning in the age of social media*. New York: Palgrave Macmillan.

Langlois, G., & Elmer, G. (2013). The research politics of social media platforms. *Culture Machine*, *14*, 1-17.

Lessig, L. (2001). *The future of ideas: The fate of the commons in a connected world*. New York: Random House.

Livingstone, S. (2005). People living in the new media age: Rethinking "audiences" and "users". In *New approaches to research on the social implications of emerging technologies*. UK: Oxford Internet Institute.

Markham, A. N. (2013). Fieldwork in Social media what would malinowski do? *Qualitative Communication Research*, *2*(4), 434-446. doi:10.1525/qcr.2013.2.4.434.434

Markham, A. N., & Buchanan, E. (2012). *Ethical decision–making and internet research: Recommendations from the AoIR ethics working committee.* Retrieved from http://aoir.org/reports/ethics2.pdf

Marshall, D. P. (2009). New media as transformed media industry. In J. Holt & A. Perren (Eds.), *Media industries: History, Theory, and method* (pp. 81-89). Chichester, UK: Wiley-Blackwell.

Matikainen, J. (2015). Motivations for content generation in social media. *Participations: Journal of Audience & Reception Studies*, *12*(1), 41-58.

Merrin, W. (2009). Media Studies 2.0: Upgrading and open–sourcing the discipline. *Interactions: Studies in Communication and Culture*, *1*(1), 17-34. doi:10.1386/iscc.1.1.17/1

Napoli, P. M. (2010). Revisiting "mass communication" and the "work" of the audience in the new media environment. *Media, Culture & Society*, *32*(3), 505-516. doi:10.1177/0163443710361658

Netzer, Y., Tenenboim-Weinblatt, K., & Shifman, L. (2014). The construction of participation in news websites: A five–dimensional model. *Journalism Studies*, *15*(5), 619-631. doi:10.1080/1461670X.2014.895527

Newman, N., & Levy, D. A. L. (2013). Reuters institute digital news report 2013: Tracking the future of news. *Reuters Institute for the Study of Journalism*. UK: University of Oxford.

Newman, N., & Levy, D. A. L. (2014). Reuters institute digital news report 2014: Tracking the future of news. *Reuters Institute for the Study of Journalism*. UK: University of Oxford.

Noguera Vivo, J. M., Villi, M., Nyirő, N., de Blasio, E., & Bourdaa, M. (2013). The role of the media industry when participation is a product. In N. Carpentier, K. C. Schrøder, & L. Hallet (Eds.), *Audience Transformations: Shifting audience positions in late modernity* (pp. 172-190). New York: Routledge.

Olsson, T. (2010). From the ecology of broadcasting to the ecology of participation critical reflections. *Nordicom Review*, (2-3), 95-104.

Ramaswamy, V., & Gouillart, F. (2010). *The power of co-creation*. New York: Free Press.

Rieder, B., Abdulla, R., Poell, T., Woltering, R., & Zack, L. (2015). Data critique and analytical opportunities for very large Facebook Pages: Lessons learned from exploring "we are all Khaled Said." *Big Data & Society*, *2*(2). doi:10.1177/2053951715614980

Schäfer, M. T. (2011). *Bastard culture! How user participation transforms cultural production*. *culture*. Amsterdam: Amsterdam University Press.

Singer, J. B. (2014). User–generated visibility: Secondary gatekeeping in a shared media space. *New Media & Society*, *16*(1), 55-73. doi:10.1177/1461444813477833

Singer, J. B., Domingo, D., Hermida, A., Heinonen, A., Paulussen, S., Quandt, T., . . . Vujnovic, M. (2011). *Participatory journalism: Guarding open gates at online newspapers*. Chichester, UK: Wiley-Blackwell.

Smythe, D. W. (1977). Communications: Blindspot of Western marxism. *Canadian Journal of Political and Social Theory*, *1*(3), 1-27.

Sormanen, N., Rohila, J., Lauk, E., Uskali, T., Jouhki, J., & Penttinen, M. (2016). Chances and challenges of computational data gathering and analysis. *Digital Journalism*, *4*(1), 55-74. doi:10.1080/21670811.2015.1096614

Sveningsson Elm, M. (2009). How do various notions of privacy influence decisions in qualitative Internet research? In A. N. Markham & N. K. Baym (Eds.), *Internet inquiry: Conversations about method*. Los Angeles: Sage.

Terranova, T. (2004). *Network culture*. London: Pluto Press.

van Dijck, J. (2009). Users like you? Theorizing agency in user-generated content. *Media, Culture & Society*,

31(1), 41-58. doi:10.1177/0163443708098245

van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford: Oxford University Press.

van Dijck, J., & Nieborg, D. (2009). Wikinomics and its discontents: A critical analysis of Web 2.0 business manifestos. *New Media & Society*, *11*(5), 855-874. doi:10.1177/1461444809105356

van Dijck, J., & Poell, T. (2013). Understanding social media logic. *Media and Communication*, *1*(1), 2-14. doi:10.12924/mac2013.01010002

Villi, M. (2012). Social curation in audience communities: UDC (User-distributed content) in the networked media ecosystem. *Participations: The International Journal of Audience and Reception Studies*, *9*(2), 614-632.

Villi, M., & Matikainen, J. (2015). Mobile UDC: Online media content distribution among Finnish mobile Internet users. *Mobile Media & Communication*, *3*(2), 214-229. doi:10.1177/2050157914552156

Villi, M., Moisander, J., & Joy, A. (2012). Social curation in consumer communities: Consumers as curators of online media content. *Advances in Consumer Research*, *40*, 490-495.

Weller, K. (2015). Accepting the challenges of social media research. *Online Information Review*, *39*(3), 281-289. doi:10.1108/OIR-03-2015-0069

Wittel, A. (2011). Qualities of sharing and their transformations in the digital age. *International Review of Information Ethics*, *15*, 3-8. Retrieved from http://www.i-r-i-e.net/inhalt/015/015-Wittel.pdf

About the Authors

Mikko Villi, Ph.D., is a Professor of Journalism in the Department of Communication at the University of Jyväskylä, Finland. His research and teaching is focused on journalism, media management and media work, concentrating on mobile, visual and social media. His international experience includes working as an international research fellow at the University of Tokyo.

Janne Matikainen, D.Soc.Sc., works as a researcher in Media and Communication Studies in the Department of Social Research at the University of Helsinki. His expertise is concentrated on the social aspects of the Internet. His research focuses on interaction in online communication, online participation and content-generation motives in social media. In recent years, he has studied the relationship between traditional and social media, especially from the audience's perspective.

Media and Communication (ISSN: 2183-2439)

Media and Communication is an international open access journal dedicated to a wide variety of basic and applied research in communication and its related fields. It aims at providing a research forum on the social and cultural relevance of media and communication processes.

www.cogitatiopress.com/mediaandcommunication

COGITATIO