This is the published version:


Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30041098

Reproduced with the kind permission of the copyright owner.

Copyright: 2011, The Authors
Can Qualitative Content Analysis be Adapted for use by Social Informaticians to Study Social Media Discourse? A Position Paper

Craig Parker
Dilal Saundage
Chia Yao Lee
School of Information Systems
Deakin University
Melbourne, Australia
Email: cparker@deakin.edu.au
Email: dilal@deakin.edu.au
Email: chia.lee@deakin.edu.au

Abstract

IS research on social media (e.g. Facebook, Twitter, blogs) has so far used user surveys or quantitative content analysis (QuantCA) research methods almost exclusively. There is considerable potential for social informatics research to use qualitative content analysis (QualCA) to explore social media discourse and its appropriation by people “in situ”. This paper presents the position that QualCA offers researchers the flexibility to identify emergent research questions and units of analysis which they may not have preconceived. This is likely to be important for IS research because of the infancy and evolving nature of social media discourse. The paper puts forward suggestions on how the QualCA research method can be adapted for this type of research.

Keywords

Social media, content analysis, qualitative, research method.

INTRODUCTION

Social informatics—a topic of research for decades—is defined by Kling as “… the interdisciplinary study of the design, uses, and consequences of information technologies that takes into account their interaction with institutional and cultural contexts” (2000, p 218). The context could include the individual, group, organisation or community level (Yoo 2010). Such inquiry argues against the technological deterministic view characterised by research questions of such forms as “Will IS solution X result in A or B?” which assumes discrete answers (Kling 2000; 2007). Instead, social informatics recognises that the impact of an IS depends on its context (Kling 2000; 2007) and that IS is shaped by social structures and by relationships among people and between people and the IS (Lamb and Sawyer 2005; Orlikowski and Scott 2008). Such research has explored various IS such as intra-organisational systems and those on personal devices such as mobile phones (Lamb and Sawyer 2005). In the case of the latter there have been calls in MIS Quarterly for IS researchers to explore the use by individuals of computerised devices (and IS) in everyday life rather than focusing only on organisations (Yoo 2010).

An emerging area of enquiry for social informaticians is social media (Wang et al. 2007). Constantinides and Fountain (2008) classify social media as blogs (e.g. Twitter), social networks (e.g. Facebook, LinkedIn), content communities (e.g. YouTube), discussion boards (e.g. Whirlpool) and content aggregators (e.g. iGoogle). All types of social media share the trait that users can generate, share and consume content (asynchronously and synchronously), while with traditional online media visitors only consume content. The interactive nature of social media gives social informaticians the opportunity to explore how people appropriate these applications (especially their online discourse or electronic textual conversations) and how social media shapes relationships and social structures (Wang et al. 2007; Yoo 2010) in the context of the individual, group, organisation and community (Yoo 2010).

The dominant research methods for studying social media have been quantitative user surveys (e.g. Baumgartner and Morris 2010; Ellison et al. 2007; Norris 2002; Saldanha and Krishnany 2010) and quantitative content analyses (e.g. Hinduja and Patchin 2007; Jones et al. 2008). The long-established quantitative content analysis (QuantCA) research method in particular typically involves: hypothesis testing; identifying content samples using probabilistic techniques; developing predefined coding schemes; counting phrases, features or topics of the content against the coding scheme; and performing statistical analysis to arrive at conclusions about the content (Hsieh and Shannon 2005; McMillan 2000; Zhang and Wildemuth 2009).

While this is useful research, the discourse by people using social media presents many opportunities for social informaticians who wish to engage in qualitative analysis. Romand et al. (2003) argue that such discourse can
provide richer insights into people’s viewpoints, feelings, attitudes and intentions than directive questions in surveys and focus groups (and associated researcher bias) especially for consumer researchers.

There are a range of established qualitative research methods which could be used to study social media such as grounded theory, discourse analysis and the qualitative content analysis (QualCA) research method. These methods are all similar in that they involve a semiotic mode of analysis but they also differ. For example, discourse analysis focuses on the language (phrases) (Myers 2011) and grounded theory focuses on data analysis simultaneously with data collection (Corbin and Strauss 1990). In this paper we focus on the QualCA research method which typically involves: purposive sampling; an iterative, inductive approach to content coding (Krippendorff 1980; Zhang and Wildemuth 2009); descriptive conclusions; and can result in theory building such as typologies (Zhang and Wildemuth 2009). It should be noted that for the purposes of this paper we describe QualCA in its extreme form to differentiate it clearly from QuantCA.

Despite the potential for QualCA we have not found any studies which use this (or any other qualitative) research method for examining online discourse in social media. More surprisingly, we were only able to find a few studies which used QualCA to examine discourse on websites more generally (Groschl 2011; Hall and Irvine 2009; Hashim et al. 2007; Parker et al. 2009; Parker et al. 2010c; Parker et al. 2010a; 2010b) and some studies which have applied QualCA to study university student discourse in online discussion forms (e.g. Hara et al. 2000). We believe that QualCA has a lot of potential for the study of social media discourse because the inductive nature of this method gives researchers the flexibility to identify new, emerging research questions and units of analyses which may not be apparent until “immersed” in the discourse. A key challenge, however, is adapting this research method for the peculiarities and complexities associated with social media.

In this paper we propose an adaptation of the QualCA method which will give social informaticians the flexibility to identify emergent research questions and units of analyses in situ and which minimises the complexities of applying QualCA to social media research. The paper starts by providing an overview of the QualCA method and then summarises the challenges associated with applying it to social media research. The paper then proposes an adaptation of the QualCA method which researchers can use to identify emergent research questions and units of analyses based on the inductive processes inherent in this method.

OVERVIEW OF THE QUALCA RESEARCH METHOD

QualCA is a research method which uses subjective interpretations by researchers of content-based phenomena within its context in an inductive, iterative process to infer explicit, implicit and/or multiple meanings (Hsieh and Shannon 2005; Krippendorff 1980; Mayring 2000; Zhang and Wildemuth 2009). This process typically leads to the identification and discussion of categories, themes or patterns which characterise the content being examined, often using quotations as evidence for any conclusions drawn (Hsieh and Shannon 2005; Romand et al. 2003).

Various authors have suggested different issues which must be considered when conducting QualCA generally (Hsieh and Shannon 2005; Mayring 2000; Zhang and Wildemuth 2009) or more specifically when studying online content (Herring 2004b; 2004a; 2010; Romand et al. 2003). We do not subscribe to the view that there are sequential steps involved in QualCA. Indeed, Krippendorff (1980) states that qualitative researchers often resist prescribed sequential steps. For example, he suggests that a research question determined at the outset may be redefined or reconceptualised by QualCA researchers based on the inductive interpretation and analysis of the content. Similarly, researchers may gain new ideas for sampling content during content analysis, which implies the potential for an iterative process of sample selection and content analysis. The four major types of considerations for QualCA researchers are now summarised based on a review of the literature.

Defining the Research Objective and Unit of Analysis

An important aspect of a QualCA study is defining a unit of analysis which can include (in increasing levels of communication sophistication) structure, meaning, interaction, participation and social behaviour (Herring 2004b; see also Herring 2010; Mayring 2000). Krippendorff (1980) suggests other generic units of analysis such as the collection of content which represents a coherent context in which to study a phenomenon (a context unit). Social informaticians may prefer to explore the social end of the continuum such as the inferred relationships between participants or the observable social behaviours and interactions based on the researchers’ repeated subjective interpretations of the dialogue in the content. The objective of research using QualCA is typically to produce categories or themes which characterise the meanings inferred from the content by the researchers (Zhang and Wildemuth 2009). The inductive nature of QualCA means that a potential outcome is theory building in the form of typologies (Zhang and Wildemuth 2009) which characterise the content, and that it can also be used for exploratory research when little is known about the content-based phenomenon being studied.
Selecting the Content

QualCA researchers can select the content by identifying instances of the content to be analysed. QualCA typically involves purposive selection of content. This content could be an exemplar, extreme or atypical in a similar vein to selecting cases in case study research (Yin 2009). The purposive selection of content will usually be driven by the research objective, or additional content may be sought after some initial analysis. QualCA researchers also need to preserve context to increase the authenticity of interpretation and understanding of meanings during iterative readings or analysis of the content (Hsieh and Shannon 2005; Mayring 2000; Zhang and Wildemuth 2009) – see Analysing the Content below. Herring (2004b) suggests that QualCA researchers can use content selection methods which preserve context such as selecting based on theme (e.g. messages in a thread of online communication) and time (e.g. all content over a specified period) or by convenience (e.g. the content is “at hand”). It is therefore essential to collect, capture or describe the context of the content or communication (not just the content itself) in a form which can facilitate repeated readings. This is important because QualCA researchers often recontextualise the content being studied based on their analysis (Krippendorff 1980).

Analysing the Content

QualCA involves an inductive, iterative process to identify themes, patterns and categories which emerge from the content (Hsieh and Shannon 2005; Mayring 2000; Zhang and Wildemuth 2009). Mayring (2000) suggests that QualCA researchers can commence the inductive process by deciding on a preliminary level of abstraction which will be guided by the research objective or question. Mayring does not provide guidance here. An example might be that looking at social behaviours versus interaction styles (based on unit of analysis) in dialogue-based content would imply a higher and lower (respectively) level of abstraction when determining coding categories. Further, the way in which categories, patterns and themes emerge from the data can vary. Herring (2010) argues that various forms of analysis (e.g. theme, feature, language, network) drawing from various disciplines (e.g. sociology, communication, linguistics) can be used depending on the unit of analysis and research objective. Regardless of the approach used, the ultimate objective of QualCA is typically to arrive at categories, patterns or themes which characterise the meaning and interpretation of the content, whereby new ones are found and old ones subsumed during each iteration of analysis (Mayring 2000; Romand et al. 2003).

Interpretation of the Content

It is important for QualCA researchers to identify their personal values, assumptions and biases as early as possible (Creswell 2003) because of the key role they play in this method. There is debate about the extent to which QualCA researchers need to be concerned with reliability and validity of their interpretations of content. Krippendorff (1980) states that QualCA researchers often do not see this as a requirement because interpretation is subjective and the goal of such researcher is often to uncover the multiple meanings in the content. QualCA researchers often prefer terms such as trustworthiness and credibility, rather than reliability and validity. This is often achieved by having teams of QualCA researchers on the one project engaging in independent coding of content and then coming together to discuss any discrepancies found (e.g. see Romand et al. 2003).

Drawing Conclusions

Findings from QualCA research focuses on descriptions and quotes which provide evidence for subjective, interpretative conclusions drawn by the researchers (Zhang and Wildemuth 2009). Inferences and reconstruction of meanings derived from data are conducted to explore and describe themes and categories which have emerged. Furthermore, the reconstruction of meanings enables QualCA researchers to describe the relationships between categories and across themes. As the QualCA approach is used to uncover hidden meanings, patterns, themes and categories, the reporting of findings will have to include sufficient details of the methods used and examples of content in the form of quotes to establish the trustworthiness and credibility of the conclusions.

Next we summarise the major characteristics of the different types of social media and the challenges these pose for using the QualCA method.

CHALLENGES OF RESEARCHING SOCIAL MEDIA

Table 1 summarises the characteristics of the major types of social media which underpin many of the challenges associated with applying QualCA.

The challenges of website content analyses have been widely acknowledged (Herring 2010; Kim and Kuljis 2010; McMillan 2000; Romand et al. 2003; Weare and Lin 2000). The major challenges, especially for QualCA, relate for instance to: handling the ephemeral nature of website content and the websites themselves (Herring 2010; Kim and Kuljis 2010; McMillan 2000); finding websites which meet research criteria using search
engines and website directories (McMillan 2000; Weare and Lin 2000); defining the context unit or unit of analysis due to the variety of possibilities such as website home pages, hyperlinked content, other content types and the entire website (McMillan 2000; Weare and Lin 2000); coding multimedia and interactive content which is not in textual form (Kim and Kuljis 2010; Weare and Lin 2000); and handling the sheer volume of data which can be obtained for analysis (Herring 2010; McMillan 2000; Romand et al. 2003; Weare and Lin 2000).

Social Media

<table>
<thead>
<tr>
<th>Social Media</th>
<th>Example</th>
<th>Format of the content</th>
<th>Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blogs</td>
<td>The blogger posts entries to an online diary that is shared with contacts. The blog entries may include multimedia content such as photographic images, animations, audio and video content. Readers of blogs may interact with the bloggers as well as with other readers through the blogs.</td>
<td>Text and Multimedia (Images, Animations, Audio, Video)</td>
<td>Interactive, Dynamic, Ephemeral</td>
</tr>
<tr>
<td>Social Networks</td>
<td>Social Network users exchange private messaging; make public broadcast; share photo, video, calendar bookmarks and applications.</td>
<td>Text and Multimedia (Images, Animations, Audio, Video)</td>
<td>Interactive, Dynamic, Ephemeral</td>
</tr>
<tr>
<td>Content communities</td>
<td>Users share content with contacts as well as the wider community. They also enable viewers and consumers of the content to provide direct feedback to the content contributor, or tag the content with textual commentaries and geographical information. Content community entries include multimedia content such as photographic images, animations, audio and video.</td>
<td>Multimedia (Images, Animations, Audio, Video)</td>
<td>Interactive, Dynamic, Ephemeral</td>
</tr>
<tr>
<td>Discussion Boards</td>
<td>Enable users to participate in forums that are organised according to themes and threads.</td>
<td>Mostly text</td>
<td>Interactive, Dynamic</td>
</tr>
<tr>
<td>Content Aggregators</td>
<td>Enable users to subscribe to content sourced from various streams, such as mainstream news station, publishers and RSS feeds.</td>
<td>Text and Multimedia (Images, Animations, Audio, Video)</td>
<td>Ephemeral</td>
</tr>
</tbody>
</table>

Table 1: Characteristics of the major types of social media

Social media more specifically shares these same issues but can also exacerbate website complexities. First, social media exacerbates the dynamic nature of website content because voluminous “postings” by people can accumulate quite rapidly in a short space of time. Second, social media is more ephemeral that websites because people have more control over how social media content is stored, organised, distributed and disposed. For example, users who contribute photographic or video content may withdraw the content later, whereas original content may be combined by subsequent users to produce “mash-ups”. Third, social media “conversations” can occur over long periods of time with no clear end-point. Fourth, contributions to conversations can take new forms such as reader feedback comments and tags (e.g. on photos) and ratings (e.g. “like” in Facebook). These issues present challenges for QualCA researchers such as considering: time-frames for data collection; whether to collect further data in multiple stages as content changes occur; what constitutes the context of and unit of analysis for online discourse; how many content coders to use given the volume of content and when they should code given the dynamism of content, as well as the implications this might have for the trustworthiness and credibility of interpretations and conclusions; and how to present evidence of conclusions for multimedia content.

Social media is also blurring the boundaries of where online conversations are situated in technological space. First, people can contribute to a conversation within a single social media application (e.g. Twitter, Facebook) using multiple tools such as smartphones, web browsers and emails. Second, the increasing integration of social media applications means that a single conversation might occur via multiple social media. For example, Facebook and Twitter are integrated and traditional websites can include Facebook and Twitter “widgets” so that people can contribute to conversations on social media via traditional websites. Indeed, this reinforces the social informatics notion of the blurring of social and technological contexts in the sense that, over time, the technological space in which online discourse occurs will become less relevant than the social context.
Nonetheless, this increasing blurring of technological boundaries presents a challenge for QualCA researchers in determining how to observe and capture this richness. In addition, social informaticians and QualCA researchers in particular could explore whether online discourse via various social media and devices are different owing, for instance, to the disparate ways in which content is presented to people. For example, would the limited content visible on smartphones result in different discourse patterns compared to using a web browser?

We now explore how QualCA can be adapted to enable social informatician researchers who are interested in studying online discourse in social media to address or reduce these challenges.

**ADAPTING QUALCA FOR QUALITATIVE SOCIAL MEDIA RESEARCH**

Herring (2010) argues that it is not always possible to determine the phenomenon of interest prior to data collection when studying new media. Although Herring wrote about blogs rather than social media broadly, we argue that her observation would be just, if not more, applicable to social media. This implies that researchers may find that their research question may change as they start to analyse social media discourse from what they originally set out to explore. Similarly, our earlier discussion suggests that framing the unit of analysis could be a challenge because of the unique nature of the social media content. The interactive, dynamic, transient and abstract nature of the content makes it difficult to articulate the boundary of the unit of analysis. In some instances, it might warrant more than one unit of analysis or having to articulate an aggregate unit of analysis. For this reason, social informaticians require a qualitative research method which gives them the flexibility to allow their research questions and units of analysis to emerge (or change) inductively throughout the process.

We believe that adapting traditional QualCA can help serve this purpose.

We propose that social informaticians could commenced QualCA by identifying one or more purposively selected social media applications (and/or one or more devices such as smartphones and web browsers) and observing the online discourse. Richter et al. (2010) has noted that the current IS research tradition has been to study only a single social media channel at a time. However, given the increasingly integrated nature of social media and associated discourse (as highlighted above), selecting and observing multiple applications as part of the one study may be more appropriate. The goal here might be for researchers to become highly immersed in the discourse across a wide range of social media (and possibly devices). This may be guided by initial views on an appropriate research question, or the research team may wish to engage in this process with any preconceived question. While doing this, each researcher would write down (in free-form) any observations of interest about the discourse. They could capture illustrative examples of the discourse using software such as Camtasia (which can capture the researcher’s “browsing”), offline website browsers or software tools such as FireShot (which captures an image of the webpages). The research team can then view or discuss the observations to arrive at one or more potential research questions and the associated unit(s) of analysis.

We believe that this approach has the potential to lead to innovative research questions and units of analyses which emerge inductively from observations of social media discourse in situ. This is because social media discourse and appropriation is likely to evolve and therefore requires an approach which encourages social informaticians to “think outside the box” when formulating research questions and identifying units of analysis.

A further advantage of this adapted form of QualCA is that it does not necessitate that researchers capture and store the content (to preserve the context of the online discourse) prior to doing the analysis. Instead, the context is preserved in the researchers’ free-form descriptions of their observations. Indeed, there would be no need to capture or store content at all, except maybe some “exemplars”. This is important because, as we have stated earlier, there could be considerable quantities of content associated with social media discourse, especially when examining such discourse across multiple social media applications. It is therefore essential, at least in the early stages of identifying (and narrowing) research questions and units of analysis, to keep the data capture/storage to a minimum. Once these potential questions and units of analysis are defined using this adapted use of QualCA, it will then be possible later to be more targeted in the content to be captured/stored.

The potential issue of the trustworthiness/credibility of researcher interpretations leading to the identification of emergent (or changes to initially proposed) research questions and units of analysis (if this needs to be defended in research publications) could be addressed if each member of the research term makes independent observations and then discuss any differences in views. In addition, researchers can report on their exemplars of the online discourse as part of the evidence for their conclusions.

**CONCLUSION**

In this paper we state that there is considerable opportunity for social informaticians to engage in qualitative research on social media discourse in order to understand how people appropriate this evolving media. We argue that, given the infancy of social media and associated research, we need a method which gives social
informaticians the flexibility to allow research questions and units of analysis to emerge inductively. In this paper we suggest a way in which the QualCA research method can be adapted to achieve this objective.

REFERENCES


Parker, C.M., Fraunholz, B. and Zutshi, A. 2009. "How do Australian small and medium enterprises communicate their corporate social responsibility activities online?", 23rd Annual Australian and New Zealand Academy of Management Conference, Melbourne, Australia, 1-4 December.

Parker, C.M., Zutshi, A. and Fraunholz, B. 2010a. "How Australian small businesses are using their websites to communicate CSR initiatives", 2010 Academy of Management Annual Meeting, Montreal, Canada, 6-10 August.


COPYRIGHT

Parker, Saundage & Lee © 2011. The authors assign to ACIS and educational and non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The authors also grant a non-exclusive licence to ACIS to publish this document in full in the Conference Papers and Proceedings. Those documents may be published on the World Wide Web, CD-ROM, in printed form, and on mirror sites on the World Wide Web. Any other usage is prohibited without the express permission of the authors.