Web 2.0 technology has supported the arrival of various social and collaborative applications. Those applications, which is known as social media, provide an opportunity for users to express their opinions and share their knowledge and thoughts with others in online space. This provides opportunities for researchers to extend their research setting and generate data within a broader environment. As a result, researchers have increasingly turned to social media sites to practice their research to expand the scope and improve the analytic power of their studies. Yet there is still relatively little direction on and much confusion about how to gather data using qualitative and quantitative approaches on social media sites. Through an extensive literature review from Information Technology and Information System journals, this study proposes new data gathering strategies within a broader environment of social media sites. This study contributes to understanding of using social media for new research settings. The limitation and future research of this study is addressed.

Key words: social media, data gathering strategies, quantitative and qualitative research

1. Introduction

Web 2.0 technology has supported the arrival of various social and collaborative applications (e.g. Facebook, Twitter, and LinkedIn). Those applications, which is known as social media, provide an opportunity for users to express their opinions and share their knowledge and thoughts with others in online space. Currently, social media has been widely used across the globe. For this study purposes, social media is defined as internet-based application that allow users to create and exchange content in online space [1]. The social media is used in various aspect of live, for example in religious activities [2, 3], education [4, 5], Social [6], Economic and business [7], and politic [8, 9], government [10].

When users use social media applications, they produce online content, such as by posting texts, photos, and videos, and at the same time the users also consume the content by viewing others’ post [11]. This social media sphere can be functioned as an intersection of activities for the producers and users and these activities create settings for researchers to investigate a phenomena [7, 12]. This implies researchers will have opportunity to extend their research setting and generate data wit-
hin a broader environment where they are not limited by space and time.

Researchers who utilize social media for their new research settings flourished. They use different methods and data gathering strategies in their research practices. Quantitative researchers, for example, they gather data through web surveys [13] and emails [14]. On the other hands, qualitative researchers gather data through online content analysis coding [15] and online interviews [16].

Instead research on social media-based has grown exponentially in recent years, very little is known how data on social media sites should be gathered using quantitative or qualitative approach to increase research robustness. This research, therefore, proposes some solutions regarding data gathering strategies using both quantitative and qualitative approaches based on previous seminal studies on various social media sites. The objective is to provide new data gathering strategies and paradigm in social media-based research.

In this study, the author will address the following research question: How do data gathering strategies should be practiced on social media with both quantitative and qualitative approaches and what are the benefits?

The structure of this article is as follows; the next section presents theoretical issues on social media, research in online space, and qualitative and quantitative research paradigm. In the third section, the approach of this article is discussed which is followed by finding and discussion in section four. Section five present conclusion and future research as well as limitation of this study presented in final section.

2. Methods

This section presents literature reviews on social media definition, reseach in online space, and understanding of qualitative and quantitative paradigm.

Understanding Social Media

The term social media has been used interchangeable with the term Web 2.0 and social networking [17]. The definition of social media proposed by Kaplan and Heinlein reflects the interchangeable nature of this term [1] who define social media as ‘a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content’. This implies that social media is part of Web 2.0 technology that allows users to communicate, participate, contribute, share, make friends through activities such as discussion, feedback, and reply to enquiries [18].

Kaplan and Haenlein [1] categorize social media into six types: collaborative activities (for example, Wikipedia), blogs and microblogs (for example, Twitter), content collection in the form of video sharing (for example, Youtube), online social networks (OSN) (for example, Facebook and Twitter), game worlds or games (for example, World of Warcraft), and social virtual worlds (for example. Second Life). However, O’Rielly (2007) categorizes Facebook, Twitter, and You Tube as social networks, while Mangold & Faulds [19] categorise Blog, Facebook, and MySpace as social media. This article takes a broad view of social media that includes the six types identified by Kaplan & Haenlein.

Research in Online Space

Instead online technology is rampant used within community, limited scholars practice their research in this technology. The technology for online survey research is young and evolving. Until recently, creating and conducting an online survey was considered a time-consuming task requiring familiarity with web authoring programs, HTML code, and scripting programs. Thousands of groups and organizations have moved online, many of them aggressively promoting their presence through the use of search engines, email lists, and banner advertisements. These organizations not only offer information to consumers, they also present opportunities for researchers to access a variety of populations who are affiliated with these groups [20]. According to Wright (2005) benefits of online re-
TABLE 3.
DATA GATHERING STRATEGIES ON SOCIAL MEDIA-BASED QUALITATIVE RESEARCH

<table>
<thead>
<tr>
<th>Qualitative research approach</th>
<th>Example authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect textual and media contents generated from users and producers. These textual and qualitative contents provide rich insight for a topic being studied.</td>
<td>[7, 11]</td>
</tr>
<tr>
<td>Access qualitative data from users’ profile and personal pages. For example, track users’ comment and opinion overtime at multiple time or base time segmentation (weekly, monthly, daily).</td>
<td>[39, 41]</td>
</tr>
<tr>
<td>Generate qualitative data (such as communication, discussion, and opinion) from social media sites</td>
<td></td>
</tr>
<tr>
<td>Use observation strategy to observe wide range of informants on social media sites. For example, Susarla et al., (2012) observed 913 YouTube users, and not limited by geography &amp; demography.</td>
<td>[10, 46]</td>
</tr>
<tr>
<td>Compare similarities and differences of qualitative data from different social media sites. This enhance understanding and allow triangulation</td>
<td>[41, 50]</td>
</tr>
<tr>
<td>Practice online content analysis coding to allow researchers early coding development. This early coding help analysis develop early insight in the following data gathering stages.</td>
<td>[37, 48, 53]</td>
</tr>
<tr>
<td>A researcher can contact more informants for an interview or observation because a social media user may connect with number of friends, ties, and followers as practiced by Hong (2013). The researcher also can conduct semi-structured in-depth interviews.</td>
<td>[15, 56]</td>
</tr>
<tr>
<td>Collect rich qualitative data from wide range of participants’ demographic information regarding motivational, behavioral, and psychological on a certain issue.</td>
<td>[57-59]</td>
</tr>
<tr>
<td>Practice long term participants behaviour monitoring or observation for example using certain social network site such as BeHive. This long-term monitoring can be used to observe people behavior and attitude toward topics being studied.</td>
<td>[63, 64]</td>
</tr>
<tr>
<td>Crawling qualitative data which can be generated from users’ profiles or postings such as positive or negative comments. The data can be users’ comments, discussions, and other postings, and it can be positive or negative review.</td>
<td>[67, 68]</td>
</tr>
</tbody>
</table>

search include; access to uniqueness population, time saving, cost reduction.

Some scholars argue that practicing online research could limit researcher to understand verbal and oral phenomena such as gestures and auditory data. However, there is ample evidence that online users compensate textually for missing auditory and gestural cues, and that online phenomena can be richly expressive [21]. Herring [21] adds that as people practice “doing things” online, and that “doing things” become online discourse and behaviour requires researchers to study with certain methods or it can be suplemented by surveys, interviews, ethnography observation, or other me-thods, and it may involves quantitative or qualitative analysis.

Research in online space can be carried out through qualitative, such as observation text discourse phenomena, or quantitative, such as phenomena can be coded and counted, and summaries of the frequencies relative produced [21]. Online space allow researchers to understand discourse behaviour which include as depicted in Table 1.

Gathering data in online space can be carried out with strategies such as conducting ‘traditional’ methods online and using ‘naturally occurring’ online data [22]. Benninger et al., [22] add that researchers can employ traditional methods such as interviews and group discussions in online space to gather data, or they can use social media as a rich source of ‘naturally occurring’ data. In traditional methods online, researchers might gather several people in an online discussion or a chat room across regions to obtain data, while in social media contexts researchers might obtain ready-available data such as textual, photographic, and videos.

Quantitative Versus Qualitative Research

Qualitative research is often associated with interpretive study, while quantitative research is considered as positivist. The majority of research in information technology has been carried out in two broad paradigms; positivist and interpretive [23].

Positivist studies assume “an objective physical and social world that exists independent of humans, and whose nature can be relatively easily apprehended, characterized, and measured” [23]. Lee [24] argue that positivist research involves the manipulation of theoretical propositions using the rules of formal logic and the rules of hypothetico-deductive logic, so that the theoretical propositions satisfy the four requirements of falsifiability logical consistency, relative explanatory power, and survival. Positivist research posits that humans are not considered as active actors that create and recreate the reality.

For example, government organizations may have structure and reality beyond the actions of their employees. A researcher needs to discover the government organizations’ physical and social reality through creating precise instruments for measurement. In other words, understanding phenomena by positivist, a researcher is about modelling, measurement, building constructs, and developing a set of accurate measurement instruments [23]. As a result, positivist researchers are able “to remain sufficiently detached so as to be objective” [25].
The data gathered is in the form the form of numbers that can be quantified and summarized, (3) the mathematical process is the norm for analysing the numeric data and (4) the final result is expressed in statistical terminologies [26].

Interpretive studies, in contrast, assert that the "social world is produced and reinforced by humans through their action and interaction" [23]. This paradigm suggests that reality and humans cannot be separated as independent identities in understanding the phenomenon because meanings are cooperatively produced by humans and their world. The implementation and use of information technology within government organizations, for example, involves the complex relationship between government employees, the technology, and other social institutions [27]. This complex reality cannot be discovered, but should be constructed and reconstructed through the interaction of humans and each person within the government organization, who should also construct his or her own reality [28]. As a result, qualitative research is understood as "any kind of research that produces findings not arrived at by means of statistical procedures or other means of quantification" [29]. Instead, qualitative research produces findings arrived from real-world settings where the "phenomenon of interest unfold naturally" [30].

Approach to this Research

This article was developed from extensive review of top journals in information technology, Information System, and research method literature. Some of those journals are listed in the following table 2.

Meanwhile, research method literature were generated from seminal books wrote by well-known authors such as John Mingers and George Walsham, M.D Meyer and M. Newman, M.Q Patton, J.W Creswell, W. Olsen, and R.B Johnson.

The author also used other online data such as web-based pages analysis [31] to identify the current phenomenon of social media use in various research setting. The analysis of Information Technology and Information System journals focus on how researchers utilized social media sites in their studies. The analysis focus on both qualitative and quantitative method. For example, how the content analysis (such as texts, photo, discussions, postings) and observation (users) were gathered and analyzed in social media based research. Some advantages of the use of social media as research settings were also extracted from the literature.

Research method literature reviews was aimed to broaden understanding of various approaches used in data gathering and analysis method. This analysis was also used to show how researchers may benefit from the use of social media in their studies. The outcome of the analysis was presented to show how social media can be applied in quantitative and qualitative research contexts. To identify the major topic areas by which to organize the review, the author undertook a systematic series of theme-extraction procedures commonly used in qualitative research [32].

<table>
<thead>
<tr>
<th>Quantitative Research approach</th>
<th>Example authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use social media sites to collect panel data and survey data from wide range of respondents</td>
<td>[40]</td>
</tr>
<tr>
<td>Deliver online survey instrument through users, social media pages by using a survey hyperlink. It also can be done through direct attachment to a user social media page (e.g: Survey Monkey site)</td>
<td>[42-45]</td>
</tr>
<tr>
<td>Collect quantitative data (such as symbols, numbers, graph, etc) from social media sites using a tool such as Graph API, Technorati.com, etc. The data can be used for statistical analyses.</td>
<td>[47-49]</td>
</tr>
<tr>
<td>Involve wide range of respondents to generate more data. For example Morris (2010) surveyed 624 and Maier (2015) surveyed people on social network sites to understand why people often turn to their friends, families, and colleagues when they have questions.</td>
<td>[51, 52]</td>
</tr>
<tr>
<td>Practice data mining and statistical content analysis on various social media sites. This data can be analysed with quantitative approach such as certain statistical procedures. As an individual on a social media may connect number of friends, ties, and followers a researcher may have opportunity to recruit more sample easily. A member may leave a public note which attract more respondents to involve</td>
<td>[54, 55]</td>
</tr>
<tr>
<td>Ablle to study a “hard-to-reach” populations. the population can be reached across region without limited by geography through virtual survey such as virtual snowball sampling. This can be done through a viral survey distribution.</td>
<td>[57, 58]</td>
</tr>
<tr>
<td>Practice 24/7 online survey. For example, Lin &amp; Lu (2011) advertised the survey on social media sites for certain period as well as on Taiwan’s largest bulletin board system (BBS), PTT (telnet://ptt.cc).</td>
<td>[60-62]</td>
</tr>
<tr>
<td>Use data crawling in which the data is generated from user profiles. The data can be users’ comments, discussions, and other postings. This data is analysed with quantitative approach</td>
<td>[65, 66]</td>
</tr>
<tr>
<td>Researchers may practice online experiment. For example researchers may test certain tools for human benefits such as remote and data analysis tools or to experiment effect of consumers relationship with certain brands.</td>
<td>[69, 70]</td>
</tr>
</tbody>
</table>
3. Result and Analysis

Social media adoption and use phenomena

Currently about 3.36 billion or about 46 percent of world population use internet and 1.6 billion or 48 percent of them live in Asia continent [33]. This number is increasing significantly every year due to improvement in technology infrastructure, regulation, and politics. Western population has adopted and used the Internet in every aspect of life earlier compared to Asian and African population. But Asian countries are considered experience higher Internet adoption and use in last few years due to economic development.

In Indonesia, the number of citizens who have access to the Internet and telephone lines has also increased. According to Asia World Stats, Internet subscribers in Indonesia is about 78 million or about 30.5% of Indonesia population. However, overall internet penetration in Indonesia is lower (23%) compared to other South East Asia countries such as Malaysia (67.5%), Brunei (72%) Singapore (82%) Philippines (43%), and Thailand (56%) [34]. According to Asia World Stats, 503 million of Facebook users live in Asia and 78 million of them are Indonesian [35]. This impacts the increase of social media users across countries around the globe. For example, daily active Facebook users until 31 March 2015 is 936 million which around 29 percent (270 million) of them are Asian [35]. Meanwhile 78 million of Indonesia population use Facebook in 2015 with 29 million of them also use Twitter, 1.3 million use LinkedIn [36]. Indonesia is also known with higher mobile connection which is 308 million mobile users and 62 million of them use mobile devices to access social media sites [37]. Indonesia is the third largest Face book users after USA (192 million), India (136 millions) and Indonesia (78 million). The trend of Face book users in Indonesia is sharply increased from year to year which most of users are in the age 18 to 25 age group [38].

Data gathering Strategies on Social Media sites

The use of social media for research setting provides some benefits for researchers. Those benefits, for example, include continue access to previous and ongoing online data [39], data is publicly available [7], and researchers are to obtain wide range of textual and media contents [11]. Based on extensive literature review, the author would like to show various data collection strategies on social media-based studies to help better understanding of using social media for research setting. The data collection strategies are summarized based on a number research conducted in various fields and used both quantitative and qualitative methods. Each of the strategy was presented together with examples depicted in Table 3 and 4.

The purpose of this article was to provide data gathering strategies for both quantitative and qualitative approaches. This purpose was achieved through extensive review of seminal studies in social media based research. This section discusses those two aspects of the findings; benefits data gathering for qualitative and quantitative studies on social media-based research.

Quantitative Research Social Media

Researchers with positive paradigm (quantitative approach) are able to reach more respondents, the validity of a research could also be increased. More respondents might help researchers to obtain more responses which potentially increase generalizability of the research findings. While small population would reduce researchers' confident in understanding the findings [74]. Relationship between number of population and generalizability has been addressed by Radhakrishna & Doamakor [75].

Generalizability is understood as “to form general notions by abstraction from particular instances” [76]. This implies that when a researcher is able to obtain more samples from a large population [52], the researcher can produce more valid conclusion compared to small sample size. Even though there is no precise number of sample should be taken from a population, but fair samples would provide valid estimation of a population characteristics being studied. This quite reasonable concept evolved in common usage so that “population” became synonymous with “all persons or all social media cases being studied or taken for sample in a study [77].

By employing social media as research settings, researchers have more opportunities to study those fair sample because samples or respondents on social media sites are within the “reach” area of researchers. More samples or respondent from social media sites, a research result would be more robust and the result can also be generalized to the larger population [78]. This increase researcher validity. This also supports a paradigm that a quantitative researcher attempts to fragment and delimit phenomena into measurable or common categories that can be applied to all of the subjects or wider and similar situations [26].

Social media platform is also can be used online experiment. For example researchers may test certain tools for human benefits such as remote control devices and data analysis tools or to experiment effect of consumers relationship with certain brands [33, 34, 73]. Other than increase researcher convinience in collecting data, this strategy might also produce new insights of doing experiment on social media sites.
Qualitative Research on Social Media

Interpretive researchers (qualitative approach) can obtain significant benefits from social media-based research. For example, social media help interpretive researchers to start early coding in online space [e.g: 15, 56]. Most of social media users post comments, textual, photos, videos, etc which allow researchers to develop their early understanding before they pursue to the next steps of their research. If the researcher plan to conduct face to face interview, the researcher might start built knowledge on a phenomenon being studied on social media sites used by the respondents.

A researcher might combine traditional and online data gathering. This strategy might be used as supplementary or additional data gathering. For example, Maier et al., [52] surveyed social network site users which then followed with in-depth interviews to understand social overload phenomena. This means understanding of a topic being studied can be enhanced when a researcher employs traditional approach together with social media data gathering. Combination between offline and online data gathering is common practiced as found by Debatin [79] and Benninger et al., [80]. For example, xxxxxx, Stockdale & Scheepers [81] in their study of e-procurement perspective, they conducted offline interview which then followed up with social media conversation observation to understand how local companies perceived the e-procurement system.

A researcher may found similarities or differences in their findings because people behave differently offline and online [22], for example a researcher might found results from face to face interviews differ from online interviews or observation. The differences could occur due to different environment of data gathering. In this case, research participants may say different thing about the enquired phenomenon but they might be under time pressures in the offline interview sessions which reflect that the “interviews [take place in] a very artificial situation” [82] compared to online social media interviews. This may lead to a lack of descriptive validity or factual accuracy of a topic being explored [83] which means what participants say may not take place in the actual context. This supports the view that that conventional interview can be problematic and participants might “create their own phenomena to satisfy or to frustrate researchers or control researchers” [84], such as by saying what they want to say to meet researchers expectation [85].

These challenges could be reduced when a study took place in social media sphere. The author argues that participants in social media space express their opinion freely through the production and consumption of content in online space. As the users are not confronted by face to face interaction with researchers and time limitation, (for example, in offline context a researcher usually exerts significant power in their engagement with a participant [86] which left participants little room to express idea, while the production and consumption of the digital texts are expressed freely. This allows researchers to observe the phenomenon in a more natural setting which become a foundation to the integrity of the observation [87] in a scientific study.

Social Media-based Research for Triangulation

Triangulation is a strategy to “validate data and results by combining a range of data sources, methods, or observers” [88]. Many studies e.g. Cress-wel and Angen [e.g. 87, 89, 90, 91] argue that triangulation can be used to achieve validity through building consistency across data source and approaches. Lack consistency across the data and appro-ached is considered weaken the evidences. However, Patton [30] suggests to view this inconsistency as an opportunity to deepen understanding of a phenomenon.

Data on social media sites can be used to validate the findings from offline study such as how far the phenomenon found on the social media observation correspond with findings from the offline interviews and field visits. This type of validation can be considered as environment triangulation which “involves the use of different locations, settings, and other key factors related to the environment in which the study took place” [92]. For example, Sarker & Sahay [93] observed participants (communicate face-to-face) physically and virtually to understand collaboration in information system development. Sarker & Sahay [93] argue that even though the research process become complex due to data overlapping between both environments but this overlapping “provided opportunities for triangulation of the ‘virtual’ and the ‘real’ understanding” (p.7).

Social media observations also enable triangulation based on data sources. As researchers obtained data from different social media sites (Facebook and Blog), they are able to triangulate the data similarities and differences across Facebook and Blog sites [e.g: 48]. The data from Facebook, for example, can be used to validate data from the Blog and otherwise. This provides opportunity for researchers to build consistency among the findings.

In case of the social media findings contradict with offline findings, this can be functioned as medium to deepen understanding as argued by Patton [94] and Olsen [95] that inconsistency can deepen researchers’ understanding on the topic being stu-
died. This could lead to new insight that require further investigation to understand how such different could occur and what is the implication. Therefore, the author argues that social media-based research is not only enable triangulation but also deepening understandings of a topic being studied.

4. Conclusion

The findings show that social media has been used in every aspect of human life. This creates new space for researchers to extend their data gathering methods. A number of data collection strategies can be practiced to increase research robustness and generalizability.

For qualitative research approach, the data can be gathered through observation of online discussion, behavior, comments, and opininion. Meanwhile, quantitative research approach can gathered data through online survey direct to social media users and positivist researcher may also practice online experiment to gather data.

When the social media is used for new research setting, some benefits are obtained by researchers. For example, researchers can use their time in online survey because they may survey directly to a social media user. Then, data from a “hard-to-reach” population can also be gathered because the social media users are usually connected in an online network ties.

Those data gathering strategies and benefits reflect social media has emerged as new potential research platform that requires researchers to redesign their research paradigm and views in doing research. As social media use has become so intense in human life, researchers also become so reliance with social media in their research activities. A study of more than 2000 researchers shows that social media impact on all points of the research lifecycle, from identifying research opportunities to disseminating findings at the end [96]. This is required re-searchers to incorporate social media in research practices to improve research benefits and out put to human kind as well as to academic community.

The author acknowledges the limitation of this study which was carried out based on literature review. However, this study has been carried in a scientific methodology to contribute to new knowledge and perspective in social media-based research to promote new paradigm and data gathering strategy. Dubin [97] argues that a research can be carried out through predicting and understanding a phenomenon, this study can be categorized into understanding the use of social media for research settings. Future research might need to address ethical issues and challenges in social media-based research data gathering. Ethical issue can be addressed by exploring ethic discourse in social media study as argued by Minger & Walsham [98].

References


Nurdin Nurdin, Research in Online Space: The Use of Social Media for Research Setting


