

Author(s):

Dr. Marcus Goncalves (corresponding)
Associate Professor of International Business,
Management
Nichols College
marcus.goncalves@nichols.edu

Dr. Erika Cornelius Smith
Assistant Professor of Political Science,
International Business
Nichols College
erika.smith@nichols.edu

Mailing Address:

International Business Program
Nichols College
124 Center Road
Dudley, MA 01571

Office phone: 508-213-2429

Fax: 508-213-2490

Title: Social Media as a Data Gathering Tool for International Business Qualitative Research: Opportunities and Challenges

Abstract:

Lusophone African (LA) multinational enterprises (LAMNEs) are becoming a significant pan-African and global economic force regarding their international presence and influence. However, given the extreme poverty and lack of development in their home markets, many LA enterprises seeking to internationalize lack resources and legitimacy in international markets. Compared to higher income emerging markets, Lusophone enterprises in Africa face more significant challenges in their internationalization efforts. Concomitantly, conducting significant international business research in these markets to understand these MNEs internationalization strategies better can be a very daunting task. The fast-growing rise of social media on the Internet, however, provides an opportunity for international business (IB) researchers to examine new phenomena in these markets in innovative ways. Unfortunately, for various reasons, qualitative researchers in IB have not fully embraced this opportunity.

This article studies the use of social media in qualitative research in the field of IB. It offers an illustrative case based on qualitative research on internationalization modes of Lusophone African MNEs conducted by the authors in Angola and Mozambique (2017) using social media to identify and qualify the population sample, as well as interact with subjects and collect data. It discusses some of the challenges of using social media in those regions of Africa and suggests how scholars can design their studies to capitalize on social media and corresponding data as a tool for qualitative research. This paper underscores the potential opportunities and challenges inherent in the use of social media in IB-oriented qualitative research, providing recommendations on how qualitative IB researchers can design their studies to capitalize on data generated by social media.

Keywords: Lusophone Africa; Internationalization, Frontier Markets, Angola, Mozambique, Qualitative Research, Quantitative Research, Social Media, Twitter, Facebook, LinkedIn

I. Introduction

A growing online community, composed of individuals and enterprises posting messages, pictures, and videos online, has embraced the emergence of social media technologies (Duggan, 2013; Duggan, Ellison, Lampe, Lenhart, & Madden, 2015). Reports of worldwide social networking activity suggest that there were 1.96 billion users in 2015 with predictions of 2.44 billion users by 2018 (Statista, 2015). Kaplan and Haenlein (2010), as well as Jukic and Merlak (2016, 98), 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” (p. 61). Of all the social networking sites, Facebook, LinkedIn, Twitter, and YouTube is among the most popular ranking within the “Top 10” of a list of most heavily visited sites on the Internet (Alexa, 2015). There is an abundance of qualitative data pouring into the Internet every day on Twitter, Facebook, LinkedIn, blogs, wikis and so forth, all of which are available for download, interpretation, and analyze by qualitative researchers. A few researchers are even attempting to categorize trends to specific social media platforms, including Facebook (Błachnio, Przepio, & Rudnicka, 2013; Caers et al., 2013; Hew, 2011; Manca & Ranieri, 2013; Nadkarni & Hofmann, 2012; Wilson, Gosling, & Graham, 2012), Twitter (Dhir, Buragga, & Boreqqah, 2013; Williams, Terras, & Warwick, 2013), and YouTube (Snelson, 2011). The combination of abounding user activity and production of user-generated content has captured the attention of scholars and researchers across numerous fields as they seek to understand social media and its role in contemporary society.

In reviewing the relevant literature and providing examples from our own qualitative research using social media, this paper will expand understanding of the application of social media to qualitative and mixed methods research within the IB discipline. The use of social

media data, or as a data gathering tool in IB research, especially when investigating frontier markets, is a relatively new field of study that has emerged in conjunction with the development of social media technologies and the global upsurge in their use (Duggan et al., 2015). There has been very little attention to IB qualitative and mixed methods using social media in the academic literature, and research trends using social media in the selection of research design, data collection techniques, and analytic approaches are not well documented.

II. Literature Review

Unlike broader studies of social media use, such as those noted above, there is not enough data about trends in qualitative and mixed methods approaches of the use of social media for international business research, especially in frontier markets. The literature includes discussions of trends in research approaches but provides a more global classification of the general patterns (e.g., Best et al., 2014; Hamm et al., 2013; Jones et al., 2013; Williams et al., 2013). At the moment, it seems, international business quantitative researchers are the ones taking advantage of the flood of data available, by using, for instance, big data analytics to analyze things such as statistical relationships between users of Twitter and their information sharing behavior (Shi, Rui, & Whinston, 2014). By contrast, qualitative studies in the international business (IB) field using social media data are few and far between (Müller, Junglas, vom Brocke, & Debortoli, 2016). These findings suggest to us that we, as qualitative researchers in IB, have a tremendous opportunity to use social media to provide additional insights to those supplied by quantitative researchers. This is especially so given that 90 percent of all digital content on the Internet today is estimated to be unstructured data (Vijayan, 2015), with most of it of qualitative nature.

The use of social media in mixed methods IB research also seems to be of significant value, mainly when exploring far-reaching markets and attempting to gather both quantitative and qualitative data, integrating them and then drawing interpretations based on the combined strengths of both sets of data to understand research problems. Mixed methods research “in which the researcher gathers both quantitative (closed-ended) and qualitative (open-ended) data, integrates the two and then draws interpretations based on the combined strengths of both sets of data to understand research problems” (Creswell, 2014, p. 2). A recent study by Chareen Snelson (2016) offers a comprehensive review of these trends in qualitative and mixed methods social media research literature published from 2007 through 2013. She cited, for example, Morgan, Snelson, and Elison-Bowers (2010) use of qualitative analysis of social media content along with a survey to find models of behavior and attitudes concerning depictions of alcohol and marijuana usage by young adults on social media websites. Snelson (2016) found that the most often used research methods included gathering data from individuals through interviews, focus groups, and other survey methodologies. The content interpretation was the second most generally used procedure whereby researchers use Facebook posts, Twitter posts (Tweets), YouTube videos, or other social media content as a data source.

The emergence of social media provides IB qualitative researchers, particularly those investigating economies challenged by extreme poverty and underdevelopment, as in the case of Lusophone African countries, a new avenue for connecting with individuals, organizations, and enterprises in those regions. In our recent qualitative study on internationalization modes of MNEs in Angola and Mozambique (Goncalves and Cornelius Smith, 2017), we faced significant challenges in being able to identify and connect with local in-country MNEs. As scholars with experience utilizing a variety of social media platforms, we contemplated the possibility that

quantitative researchers in international business, especially those researching frontier economies where extreme poverty and lack of infrastructure development is pervasive, could make good use of social media as both data. A literature review of primary and notable journals with articles on qualitative IB studies utilizing social media yielded limited results.¹

Research on the use of social media in qualitative, quantitative, and mixed methods IB research was further complicated by the many alternative lexicons for social media used by researchers in the field. Terms ranged from *new media of the Internet* (Vaast, Davidson, and Mattson, 2013), to *Internet social networks* (Ameripour, Nicholson, and Newman, 2010). Hence, the authors decided to expand the search using various terms. Additional terms included social media, microblog, wiki, enterprise 2.0, online social network, LinkedIn, online community, Twitter, social media, web 2.0, and blog (Wang, Min, & Liu, 2014; Aviles & Bauman, 2017). Taking into account differing terminologies and levels of use of social media tools, we find that the vast majority of studies in the IB field used quantitative research methods, with only a minor number of them using a qualitative methodology of some kind (often as part of a mixed methods study).

To improve our comprehension of this phenomenon, we decided to expand our revision of the literature and performed multidisciplinary search queries on the use of social media in journals from other business and management disciplines, and found the same pattern. For instance, for management, marketing, human resources, and information system disciplines, most investigations using social media data were quantitative, not qualitative. Also, we realized there was a distinction in the type of data used by quantitative and qualitative international

¹ Academic journal survey included: *Corporate Governance: An International Review*, *Cross-Cultural Research*, *European Journal of International Management*, *Global Strategy Journal*, *Journal of International Business Studies*, *Journal of International Economics*, *Journal of International Management*, *Management International Review*, and *Journal of Management & Public Policy*.

business researchers. Most quantitative researchers used data directly extracted from social media platforms. The types of quantitative data, while it tended to vary, typically included data items such as message counts, messages downloaded, number of connections, number of likes (or dislikes), number of hits, friend counts, number of posts, or level of participation. Only a few quantitative articles and papers used data from online surveys. Conversely, most qualitative articles and papers using social media as a data gathering tool did not use qualitative data extracted directly from social media platforms. The most common data collection method was interviews of social media users, which is in line with our research approach for Angola and Mozambique MNEs. Only a small number of articles and papers used qualitative data directly gathered from social media platforms (e.g., Ameripour et al., 2010; Germonprez & Hovorka, 2013; Payton & Kvasny, 2012; Vaast et al., 2013; Vaast & Levina, 2015).

Although the findings of this study are not generalizable, the review of literature above suggests a significant loss of opportunity for the field of IB. There are a variety of ways to use social media to improve qualitative research in IB, especially in areas where not much information is available, such as frontier markets. The potential usefulness of examining trends in the use of social media for IB's qualitative research approaches (e.g., semi-structured interviews, focus groups, and qualitative content analysis) lies in the discovery of how researchers are designing their studies. There is value in gaining insights into the various processes of such research, from the identification and selection of the population target to data collection and analysis. Studies outside the field of IB, including those by Snelson (2011, 2016), prompted us (see Goncalves and Cornelius Smith, 2017) to construct similar research design. We used LinkedIn's emerging and frontier market groups to conduct a qualitative analysis aimed at uncovering patterns of internationalization process regarding MNEs from Lusophone African

countries, specifically in Angola and Mozambique. We utilized LinkedIn, Twitter, WhatsApp, and WeChat to search for potential MNEs in Angola and Mozambique that were willing to participate in our research and collect data to help us better understand how these enterprises were internationalizing. A second study by Vyas, Landry, Schnider, Rojas, and Wood (2012), combined a survey with follow-up interviews to examine short message services and social media use among Latinos. This study prompted us (see Goncalves and Cornelius Smith, 2017) to utilize Twitter and Constant Contact to conduct a short quasi-quantitative survey and another more extensive qualitative one, to qualify respondents and set up in-country interviews. Greene, Choudhry, Kilabuk, and Shrank (2011) conducted qualitative research evaluating Facebook communities posts dedicated to diabetes to reveal how patients, family members, and friends share information and receive emotional support. Such strategies prompted these researchers to evaluate how the MNEs interviewed in Angola and Mozambique were using LinkedIn and Facebook in their internationalization process as well as the promotion of their products and services abroad, which turned out to be a significant trend in the research findings. These examples help illustrate the potential of social media as a tool in qualitative and mixed methods for IB research to uncover new insights through the corresponding combination of these methods.

One probable reason for this state of affairs is that there are few qualitative research methods articles and papers about the use of social media and big data in international business, in particular when used to gather data from frontier markets. Hence our motivation in writing this paper, as we would like to encourage qualitative researchers in international business to start using this valuable and potentially intriguing source of data. The following sections of this research will suggest how qualitative researchers in IB can use social media as a tool, as well as

its data. Although the value of qualitative social media data addresses specific purposes, such as supplementing quantitative social networking studies (e.g., Whelan, Teigland, Vaast, & Butler, 2016), this paper is only one of the few qualitative research methods contributions to the IB research literature about the use of social media. These findings should be of interest to Ph.D. students, to international business professionals and researchers who are unsure about the conduct of research on social media and its implications, and to researchers in general who are interested in combining traditional qualitative techniques with social media studies.

III. Research Method

For this paper, we define social media as computer-based tools (such as websites and apps) that enable people to create and share content with other people and participate in a community. Bradley (2010), from the Gartner group, says that at their foundation all types of social media are a set of technologies that can construct and enable a potentially extensive community of participants to collaborate. Whereas IT tools to support collaboration have existed for decades, new social-media technologies allow collaboration on a much grander scale (Bradley, 2010).

The term “social media” typically identifies web 2.0 applications such as blogs, social networking sites, or video/image/file sharing platforms, and wikis (Fuchs, 2013). Kaplan and Haenlein (2010), as well as Green (2014), 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 definition introduces us to two key concepts common to social media: technology and content. The authors of this paper see social media as dependent on mobile and web-based technologies, which enables the creation of highly interactive media through which individuals and communities share, co-create, discuss,

and modify user-generated content. Shirky (2011, p. 20) notes that social media are tools that “increase our ability to share, to co-operate, with one another.” Boyd (2009) claims that social media is a collection of software allowing people to collaborate, play, share, and communicate, characterized by user-generated content. Lovink (2011) observes that social media facilitates social interactions, while Meikle and Young (2012) add that social media includes the creation of a profile, contacts, and the communication between those contacts.

The following two sections of the article explore these issues, and others noted in the Literature Review. The part immediately following, "Social Media in Qualitative IB Research," discusses challenges related to data validity, barriers to access, and ethics in qualitative social media research. Section V utilizes an illustrative case study from the authors’ recent experience studying internationalization strategies of MNEs from Angola and Mozambique to illustrate these challenges in practice, and Section VI provides recommendations from the integration of our experience with multi-disciplinary bodies of literature referenced previously. Generally speaking, in our research of MNEs in Angola and Mozambique, social media did allow for the search and introduction of the research prospectus in a very ubiquitous and pervasive way. It enhanced the ability to communicate with MNE’s executives in the country, as well as other professional community of practices.

IV. Social Media in Qualitative IB Research

While qualitative research of any kind can be challenging, regardless if of a traditional nature or a relatively new method, there are added challenges when using social media as a tool, as well as its data (Hanna, 2012; Hunt & McHale, 2007; Jowett, Peel, & Shaw, 2011). Some of these challenges may include:

- **The validity of data** – Unlike traditional research methods, on social media participants typically tend to be anonymous or use pseudonyms (Christopher, 2009), making it difficult to confirm the authenticity of the data. If the identity of the interviewee cannot be verified, data may not be trusted. Moreover, the literature suggests that anonymous use, while offering users a high degree of privacy, tends, at the same time, to give users the license to “misbehave.” Posting inappropriate, offensive or illegal content without fear of punishment (Tsang, Au, Kapadia, & Smith, 2010), are some examples.
- **Data sources** - A fundamental difference between data obtained from traditional face-to-face interviews and those gathered via social media is that the researcher has to generate interview data. In the case of social media data, although it can be produced (e.g., online surveys, chats, virtual conferences, and focus groups, and so forth.), it can also be the result of self-generated users content. In a traditional qualitative interview, the researcher has more control since they tend to direct the conversation with focused questions (Myers & Newman, 2007). With user-generated data, however, there is less control and less knowledge about the origin of the data, meaning there is potentially much more “noise” in the data, which will require filtering. Also, the social media data collected may not contain the specific points the researcher is interested. Additionally, there may be questions about the reliability and validity of the data, as noted above.
- **Digital Divide** - The use of social media might exclude some members of a society who do not have access to or are uncomfortable with their use. In Iran, for example, blogging is viral but only among a minority of the population, typically younger individuals or the educated elite (Ameripour et al., 2010). The same is true for frontier economies such as Angola and Mozambique with an aggravating factor that such technologies are costly there, and the public services (e.g., electric power, the Internet, and bandwidth) are not as reliable or in some cases, unavailable. The existence of a digital divide, therefore, might exclude some individuals from a study if the researchers use social media only, hence making the use of social media alone nonviable for particular research projects.
- **Digital Texts** - The type of data available on social media platforms, which frequently contain new forms of digital texts (Urquhart & Vaast, 2012), can be a challenge. For instance, it is quite common to find emails, chat threads, images, wikis, avatars, YouTube clips, microblog posts, and emoticons/emojis on these platforms. Diaz Andrade, Urquhart, and Arthanari (2015) advocate the use of images in information systems research, not just as contextual information to other data sources, but also as a source of information in their own right. Of course, most social media platforms contain images of one sort or another. Urquhart and Vaast (2012) suggest that there is a need to theorize these social media-related environments given that they provide a wealth of digital text data.
- **Ethics** – according to Ackland (2013), there are three leading ethical concerns relating to research using social media as a research tool:
 - **Informed consent** - Given some of the concerns mentioned earlier, it is just not realistic to obtain the permission of everyone on a social networking site. For instance, when web-crawling for data from social media sites, the volume of data is often substantial. Trying to obtain consent from every contributor or author of the data, let alone verify their identity, is impractical. The consensus of most scholars, however, seems to be that researchers are free to use data accessible in the public domain, such as websites, newsgroups, and blogs, but one needs to be careful not to infringe copyrights or invade people’s privacy. Therefore, it is advisable that

researchers attempt to obtain consent whenever possible when researching such sites where some expectation of privacy exists. The challenge then becomes, from who should one receive permission? In the first author's research in Angola and Mozambique, this became a significant issue, primarily due to local laws, culture, and privacy expectations, not to mention the foreign language, with Portuguese being their local language and frequently used in their blogs and postings. The issue of informed consent became a lot more complex for us, prompting the creation of a multi-lingual consent form and research prospectus. The works of Thelwall and Stuart (2006) were very insightful in this case, as they explored these issues and a few other ethical concerns in great depth.

- **Making the distinction between public and private** - The difference between these two spheres can become a lot fuzzier in online environments. For instance, bloggers may reveal personal information about themselves in a public manner, but with the belief, they are only interacting with a small group of people. In such case, people tend to assume that they are having a conversation in a private place, with the expectation that their peers will not use, or disseminate, the information. So, is the data public domain, and if so, can it be used? The distinction tends to be more explicit in social networking sites, such as Facebook or LinkedIn, because of the use of privacy controls.
- **Participant anonymity** – When using social media as a research tool, it is frequently not clear when to grant anonymity to participants. For instance, in the first author's study of internationalization of MNEs from Angola and Mozambique, many of the data gathered from the executives were from LinkedIn and Facebook. The first author collected this data by asking questions and reading of their blogs, requesting them to provide additional evidence about themselves, to confirm the authenticity of the subjects and validity of their data, and also to assist in our research population selection criteria (Ameripour et al., 2010). Many of these bloggers and interviewees, however, were not willing to reveal their full identities (beyond their user/screen name). They feared for their safety (when posting or commenting on government-related themes, financial and strategic information of local MNEs), and breach of confidentiality (Ameripour et al., 2010). Accordingly, the first author decided to change the research design to include face-to-face semi-structured interviews in the country, after earning their trust and offering a commitment to maintaining interviewee anonymity. Despite these extra steps, and international travel expenses, not all of them felt comfortable with the pledge to protect their identity. In these cases, the use of social media, where the subjects did not have the chance interact with face-to-face and in a more personal way before the interview, was not helpful. In fact, it was detrimental, as in some cases, it negatively influenced the interviewee's ability to trust the researcher. Some of them did not believe the first author due to his screen name in Google (e.g., MGRider, as the word "rider" did not inspire trust in them), generating a bad and mistrusting connotation (due to the word "rider"). Two of the interviewees were so suspicious that they asked if the first author worked for the U.S. government and if the first author was with the CIA. Although this is conceivably an extreme illustration, it is real, and it demonstrates that considering anonymity along with issues such as authenticity and validity can be a challenge when resorting to social media and

networking tools for research, especially qualitative ones. Light and McGrath (2010) and Zimmer (2010) examine this and other social media ethical concerns at length and in more detail.

- **Levels of Access** - In traditional face-to-face interviews, the level at which a researcher enters an organization is critical (Wasko, Teigland, Leidner, & Jarvenpaa, 2011) and will affect the researcher's ability to move around the place. Often a researcher may only talk to senior managers and other key personnel, but not employees at the front line. The same issue can apply to social media. For instance, when collecting data on LinkedIn, basic access to the platform, which limited the number of companies and professionals eligible for contact, does not provide much of the relevant information being sought in the previous study (Goncalves and Cornelius Smith 2017). It was necessary to upgrade to a fee-based membership (premium) to be able to access additional information and large numbers of users. Without the required permissions and payment of a monthly fee for the duration of the research, access to these areas of the site and the much more significant number of contacts made in a month would have been denied or limited. In fact, in some social media environments it might not be sufficient to obtain permission just to enter the site; instead, the researcher may need to create one or more avatars (Schultze, 2010) or create some other type of online presence, such is the case with Second Life (www.secondlife.com). While there are many MNEs in Second Life, the first author decided not to use it, as he would need to attain to a certain level of skill with their avatar to access and move around that (virtual) world. However, the fact remains that, a researcher might be able to access more subjects using social media than would otherwise be possible in a traditional organizational setting.
- **New Types of Behavior** - Researchers may experience unique types of demeanor in social media that are not usually found in face-to-face interviews. For instance, behavior such as “flaming” (Papacharissi, 2002), “lurking” or “whispering” (Garcia, Standlee, Bechkoff, & Cui, 2009) can occur only in social media. For instance, lurking is where participants in social media adopt passive behaviors: they listen to, observe, and perhaps record the "conversations," but do not engage with the contributors to the social media to any great extent, if at all. One cannot account for such passive behavior if it is not visible or noticeable.
- **Sheer Volume of Data** - The most obvious challenge in using qualitative data from social media platforms is the large volume of data typically available. Although qualitative research already tends to produce large amounts of data (Myers, 2013), the size of the social media data set can be very daunting. For instance, Twitter generates about 6,000 tweets per second, which corresponds to over 350,000 tweets sent per minute, 500 million tweets per day and around 200 billion tweets per year (Internet Live Stats, 2016). In one information system study, the researchers collected a total of 1,915,429 tweets from 50,778 Twitter users on their chosen topic in just two months (Oh, Eom, & Rao, 2015). Electronic data management tools such as Nvivo are undoubtedly necessary to store, categorize and manage the data, but these tools in themselves are not sufficient. Since qualitative researchers tend to study a particular topic in depth with a focus on the context, an efficient way to filter, or "clean," the data is a must, so that irrelevant data can be ignored, while the richness of the story is revealed.
- **Visual Cues** - In traditional face-to-face interviews, there may be a variety of cues generated by both parties in the social encounter. Such prompts may vary from facial

expressions and jokes to encouraging sounds and mannerisms. Body language during semi-structured interviews often provides essential cues that help validate data, especially during the transcription of the conversations, which can be useful in supplementing the words and statements embodied in the transcripts. It may even provide pointers about the demeanor of the interviewees and how open they appear to be at the time of the interview (Myers & Newman, 2007). Depending upon the nature of the social media platform, however, many or all of these hints may be absent or replaced by electronic ones.

As IB researchers, these various characteristics of social media research challenge us to expand our thinking when designing methods for IB research studies, both qualitative and quantitative. For instance, as depicted in the illustrative case analysis in later sections, one possibility relates to the depth of involvement it enables. Walsham (1995) considers two extremes about the roles of the researcher in qualitative research: that of the outside observer and that of the involved researcher. In our qualitative IB study of Lusophone African MNE's internationalization, the outside observer role could take the form of a passive observer on a social media platform. In our case, it was LinkedIn, where the platform's paid (Premium) advanced features were used to extract specific data related to its members and Facebook. Likewise, as an outside observer, as depicted by Vaast et al. (2013), the first author analyzed the discursive practices of a group of senior executives from Angolan and Mozambican MNEs, obtaining data through an aggregator website, in this case, Constant Contact, as well as LinkedIn groups again. The first author's involvement and role-playing, as a researcher, was through voluntary membership in these social networks and groups (community of practice), hence enabling closer access to data, many confidential or sensitive, as he became an active contributor and participant in the social network. Unlike traditional qualitative methods tools, by using social media, as already observed by Coleman (2010) and Walsham (2006), our role as researchers, was more of a spectrum, where our position stood somewhere in between the two extremes of complete outside observer and fully involved researcher. This role, at times, did change.

Undoubtedly, there are several ways to design qualitative studies integrating social media as research methods, but the researchers must consider their role alongside with the topic and the adequate type of social media to use. Scholars should be aware of the various potential challenges, such as those noted above, when using such tools. Notwithstanding these inherent difficulties, we believe that the use of social media for qualitative researchers in international business holds much promise. In the following section provides a summary of some of these difficulties, with a comparative qualitative study conducted by the first author, adding our own experience and perspective whenever appropriate.

V. An Illustrative Qualitative Study: Internationalization of MNEs from Angola and Mozambique

In this section, we describe our recent qualitative research conducted using social media as a data gathering instrument, and illustrate portions of the aggregated data. This discussion is intended to provide an illustrative example and context for the recommendations offered in the following sections. Given the various studies on the use of social media as part of a research method, the example of this qualitative study is not meant to be representative of social media studies more broadly and instead focuses on the challenges and opportunities inherent in this method of IB qualitative research.

Internationalization theories suggest that enterprises from emerging and frontier markets will adopt different entry modes than those in advanced economies. There are very few studies to date, however, examining the process of how multinational enterprises (MNEs) from frontier markets internationalize or evaluating which factors influence their mode of entry into global markets. Our previous research (Goncalves and Smith, 2017) investigated the internationalization strategies of Lusophone Africa MNEs from Angola and Mozambique, more

specifically their entry mode, to expand the framework for entry mode strategies to include the motivations and issues of MNEs from emerging and frontier economies.

To conduct this study, we used a variety of data gathering techniques, including netnography, also known as virtual ethnography or online ethnography (Hine, 2000; Kozinets, 2010; Ruhleder, 2000). Netnography is a form of ethnography (Harvey & Myers, 1995; Myers, 1999) that involves participation and interaction with community members over the Internet, which in our study, we relied on LinkedIn and Facebook. The authors obtained textual data from other sources as well including blog posts, discussion forums and a community of practice website from Angola. Note that, in the subsequent discussion, we are focusing solely on how we conducted the research, to demonstrate how we dealt with some of the challenges in using social media as a data gathering instrument, and not on the findings of such research. The conclusions of the study have been and are being published elsewhere (Goncalves & Cornelius Smith, 2017).

The role of the researchers was that of *involved researchers* (Walsham, 2006). After obtaining permission from prospect participants in LinkedIn to conduct the research, the first author immersed himself in discussions and forum postings. He joined a few emerging and frontier economies groups and participated in many online activities such as discussions, debates, writing of opinions (op-eds) for publication at LinkedIn, with the intent to attract the targeted group of professionals for the research. These groups and their facilitators/moderators were aware of the presence of the researcher playing the involved researcher role. During the fieldwork, field notes such as digital texts and images were taken, as suggested by Urquhart and Vaast (2012), and stored in NVivo, as well as Excel spreadsheets and memos in Word format. We found that the discussion forum data in LinkedIn's groups were a valuable data source. The discussion forum data was downloaded from the website with a script to automatically extract

the content of each forum post, the thread it belonged to, the entire thread, the dates stamps of each post, and the name of the poster. The researcher saved this dataset into NVivo as well, but later on moved it to Excel, for more straightforward manipulation.

In total, the first author spent over 2200 hours engaging with LinkedIn groups. Table 5.1 lists the multiple sources of data that were collected from or through the use of social media throughout this research project. The challenges of using this data are discussed next, although we will not address the data obtained from LinkedIn, Facebook, or any other websites, as the way we approached the analysis of this textual data was not particularly unique to social media and no different from traditional text analysis approaches.

Table 5.1 – Data sources originated from or through the use of social media

Data Source	Nature of data	The quantity of data collected	Type
LinkedIn group's forums	Discussion posts from the emerging and frontier economies groups	837 discussion posts, dating back to 2013.	Text
Chat logs	Chat logs from WhatsApp and WeChat with selected executives from targeted MNEs	Approximately 14 months' worth of chat logs.	Text
Facebook	Textual information relating to background information about the MNEs interviewed and their internationalization process, through exporting, using Facebook.	Approximately eight pages	Text

Given that we were studying frontier economies MNEs at online community of practices, the field notes based on participant observations took a different form than in traditional qualitative studies. For instance, the researcher could not actively take field notes during interactions with LinkedIn, WhatsApp, WeChat, or Twitter members, as he was using the mouse and keyboard for other tasks, such as moving through applications, positioning text, taking screenshots, posting emojis, and so forth. The researcher, therefore, had to record these interactions by saving the discussion at the end of a session, taking screenshots, and at times cutting and pasting blog posts.

This approach to collecting data during fieldwork meant that text and images of graphs and diagrams were obtained from the participant observation, as well as textual data. The first author took screenshots, as recommended by Kozinets (2010). As Urquhart and Vaast (2012) point out, digital texts can include images or graphs (and diagrams) of the cyberspace. By analyzing these images, graphs, and charts, we were able to understand better how data were cross-linked with different businesses and industries. Hence, the presentation of pictures and diagrams might not only increase the contextual understanding of the reader but can also provide empirical evidence to support a theoretical point (in this case concerning the interaction between internationalization modes of MNEs across two distinct countries at the same time).

Adopting a qualitative approach method with an exploratory and reflexive nature (Alvesson and Sköldbberg, 2009), to allow for more in-depth cross-cultural understanding, (Marschan-Piekkari and Welch 2004; Arber, 2006), this research included a series of qualitative online surveys. It also included semi-structured interviews and a descriptive case study method to investigate the central research questions (Yin, 2003). Invitations were sent out via email through LinkedIn, Twitter, and corporate email venues to 716 prospect corporate officials, senior management and decision-makers at Angolan MNEs, and 566 to Mozambicans MNEs' inviting them to participate in the online survey and semi-structured interviews. Of 1,282 invitations issued (in at least three attempts on different occasions via alternate modes, using social media (i.e., via LinkedIn, Twitter whenever possible, corporate email and introductions), only 29 prospects answered the online survey and 26 agreed to meet in-person, at Luanda-Angola and in Maputo-Mozambique.

The online survey, using Constant Contact as a data aggregator, and LinkedIn and Twitter to invite surveyees to the study, contained questions to measure the manner in which

MNEs rated internationalization motives and factors as well as control variables concerning the characteristics of the MNEs themselves. The control variables (selection criteria) provided information about the type of MNE that could respond to the survey. The most important part of the online survey, however, was the grading of the motives for internationalization. The MNEs were asked to grade these motives on a five-point Likert-scale. The scale started with “not important” (ranking=1) to “very important” (ranking=5). The responses were analyzed for both single motives and groups of motives (details of ranking are provided in such instances).

The first part of the online survey consisted of control variables. Another consideration was with the cultural, administrative, geographic, and economic (CAGE) distance framework, which helps managers identify and assess the impact of distance when internationalizing. Dunning's OLI framework questions, as well as Uppsala and the Network theories, guided analysis of how these LAMNEs are internationalizing. When asked about their personal information and *position* they held at their enterprises, the online survey found that all the respondents were senior executives of managers. They occupied various decision-making positions varying from Country Managers, COO, CEO, CIO, President, General Manager, Founder, CTO, Country Ops Manager, Economist, Central Auditing Director, VP of International Marketing, Commercial Team Coordinator, in addition to many undeclared, for confidentiality reasons.

When asked if they had interest in being personally interviewed, 21 (19+2) of the respondents agreed to be interviewed. To be personally interviewed, a respondent had to belong to an enterprise that had already internationalized or was in the process of internationalization. After commencement of interviews in these countries, other enterprises agreed to be interviewed taking the total number of respondents to 26. Others declined for various reasons, including but

not limited to “being too busy,” “only available via email,” “not being the right person to be interviewed.”

Semi-structured interviews were digitally recorded and conducted face-to-face to the extent possible (12 face-to-face meetings in Luanda-Angola, and 10 in Maputo-Mozambique), or over video conferencing using Skype, WhatsApp, or WeChat (two with Angola and two with Mozambique). We avoided using the telephone to prevent international calling charges. All in all, we completed a total of 14 interviews in Angola and 12 in Mozambique. These semi-structured interviews explored the views and experience of respondents in greater depth and allowed them to go beyond the constraints of the multiple-choice survey. In such open-ended conversations, they discussed their views on the internationalization mode and process of their MNEs. This method of in-depth, phenomenological interviewing (Seidman, 2006), provided an opportunity to gain greater insight into the language, data, and stories of the survey population targeted, including senior management officials and decision-makers from MNEs in Angola and Mozambique, as ways of knowing and understanding the issue and context involved in the internationalization process of their MNEs.

The online survey and semi-structured interviews conducted produced data that allow some conclusions to be drawn regarding the internationalization modes of LAMNEs, specifically from Angola and Mozambique. Consistent with the literature (Anderson, and Gatignon, 1986; Okoroafo, 1990) the findings show the main reasons these LAMNEs internationalized were for profit, expansion, market opportunity and to achieve growth, taking advantage of OLI framework, mainly location and internalization advantages.

As discussed earlier, one of the most significant challenges in using social media as an instrument for gathering data is analyzing the large volume of such data, which was also the case

in our study. Given that we collected 837 discussion posts, we needed to find a way to filter the data so that we could focus our attention only on those posts considered significant for answering our research questions.

In addition to traditional qualitative data analysis software like Leximancer, there are also many other application tools available for more advanced text mining, such as QDA Miner and KNIME. We found NVivo to be useful for analyzing large amounts of text as the first author already had it at his disposal. Although Leximancer's use of machine learning for content analysis, to analyze massive amounts of qualitative data sets and to display the results in a visual format, would have been of a substantial help with we were more familiarized with it.

Data Analysis

Data analysis involved the use of NVivo and Excel spreadsheets. The first step was to load the entire dataset from ConstantContact, LinkedIn, Facebook, Twitter, WeChat, and WhatsApp into NVivo. Since our research was exploratory, we adopted an unsupervised approach. This created a set of NVivo themes. However, the dataset contained far too many posts which were irrelevant to our original research questions. Consequently, we had to determine a way to reduce the dataset.

Therefore, our second step was to reduce the dataset by reading the blogs and chats notes in an attempt to discover which had the most impact on the internationalization mode's overarching research question. After analysis, we identified the central themes with stronger influence. Focusing on these themes, we were able to filter the data from the discussion forums in LinkedIn by looking at the headers and extracting only those posts of a recurring theme. We also performed keyword searches using keywords from our theoretical approaches along with

keywords based on our knowledge of the topic. Therefore, we were able to disregard a large number of posts that were creating "noise" in the data, focusing on those we considered to be useful for answering our primary research question.

The third step was to reload this reduced dataset into NVivo again for manual coding with theoretical sensitivity to actor-network theory (ANT)²Please note that a discussion of theory choice is beyond the scope of this paper, but is discussed elsewhere (McKenna et al., 2012).

This created a new set of themes and knowledge pathways which were then more manageable.

The fifth step was to further codify, manually, with theoretical sensitivity to ANT using MS Excel to generate some visual graphs for analysis. The codes from both rounds of NVivo were compared with each other, along with the charts created in Excel.

VI. Recommendations

From the above challenges noted in qualitative social media IB research in Angola and Mozambique, and insights from the survey of the multi-disciplinary literature, we offer several recommendations for using social media as a tool for qualitative research in IB. Table 6.1 lists a summary of these recommendations.

Table 6.1 – List of recommendations

CHALLENGE	DESCRIPTION	RECOMMENDATIONS
Data source	Social media data may not be reliable or readily available or useful	A. Consider the use of data qualification techniques B. Consider data triangulation C. Consider qualitative data analysis software
Digital divide	Social media use may exclude some people or groups	D. Consider supplementing social media with traditional data gathering techniques

² Actor-Network Theory (ANT) is based in science and technology studies. As a process for conducting in-depth research, it has now been used in other areas of science as well. ANT does focus on the connections that are being formed and transformed between personal and non-personal realities that are part of the issue at stake. It goes past the borders that are usually set: ANT does not stop the study when it enters contexts or so-called underlying structures. Tracing back links can be done by observation, document analysis or in-depth interviews.

The multiplicity of digital text	Social media platforms contain a variety of types of digital texts, images, and videos	E. Data, such as screenshots, images, videos, and texts may need to be gathered and analyzed, often requiring real-time analysis, while actively engaged in online fieldwork.
Informed consent	It may not be realistic to obtain the consent of a whole group in a social networking site	F. Consider posting a disclaimer informing participants about the data collection intent and purpose G. Attempt to obtain consent whenever possible on sites where expectation of privacy exists
Level of access	Researcher may need to access certain (restricted) areas of the social media platform	H. Get permission to access social media area I. If needed, become a temporary member of restricted area J. Consider creating and using avatar(s)
New types of behavior	People may exhibit different types of behavior when using social media than in a face-to-face environment	K. Get acquainted with and socialized into the social media platform (e.g., memberships in LinkedIn, Facebook, Twitter, and so forth)
Participant anonymity	Participants may be anonymous or use pseudonyms, potentially raising questions about the authenticity of the data	L. Use mixed methods to triangulate various types of data M. Subject's identity should not be important when developing research questions
Public and private distinction	Distinction between these two spheres can be fuzzy in online environments	N. Inquiry the group on whether they consider available information public or private
Validity of data	Unlike traditional research methods, on social media participants typically tend to be anonymous or use pseudonyms (Christopher, 2009), making it difficult to confirm the authenticity of the data.	O. Take steps to ensure the reliability and validity of research findings
Visual Cues	Visual cues may be in digital form (through use of avatars, pictures, and so forth.) rather than face-to-face	P. Become familiarized with the various social media platforms and features (e.g., the use of avatars, filters, augmented reality, and so forth)
Volume of data	The researcher gathers large amounts of data, often containing significant noise, which requires filtering before it can be analyzed.	Q. Apply a filtering or data mining technique R. Use qualitative data analysis software

A. Consider the use of data qualification techniques

For qualitative researchers to develop proper internal representations for pattern mapping, noise in the dataset must be managed. Due to the many challenges associated with manually validating data, we recommend using decision table techniques as a possible, domain-independent way of optimizing dataset formulation. Decision tables accommodate a variety of devices whereby datasets can be prepared to eliminate vagueness, contradictions, and other types of noise.

B. Consider data triangulation

Data triangulation enables researchers to validate data and research by cross verifying the same information. This triangulation of data strengthens research as data credibility and validity increases. Also, triangulation of data reinforces investigation by adding additional sources of information, which often give more insight into the topic at hand. It also enables inadequacies found in one-source data to be minimized when multiple sources confirm the same data (data saturation). Triangulation, among many other features, allows for various

sources of data verification and validity while complementing similar data. It allows for more comprehensive data to be obtained.

C. Consider qualitative data analysis software

Considering the large volume of data, it is just not feasible to analyze social media-generated data manually. It is essential to use a qualitative data analysis (QDA) software package to help in the management and interpretation of data. Not all QDA software packages, however, have the same features. Hence, it is essential to choose the most appropriate one. In our illustrative case of Angola and Mozambique MNEs, it would be incredibly time-consuming to code 58,653 discussion forum posts on LinkedIn, in addition to 26 90 minutes digitally recorded face-to-face semi-structured interviews using a spreadsheet. NVivo, which enabled for auto-coding, was utilized to a limited extent. A more appropriate software would have been a text mining software such as Leximancer, which provides automated qualitative data analysis. Of course, researchers need to be careful when using automated text analysis tools, as one needs to be aware of the “garbage in garbage out” problem (Kim, Huang, and Emery 2016), which is mainly problematic when using live data such as discussion forums.

D. Consider supplementing social media with traditional data gathering techniques

To address the potential problem of the digital divide, especially in frontier economies where environmental and infrastructure challenges, among others, tend to be present, IB qualitative researchers should consider supplementing social media as a data gathering tool with traditional data gathering techniques. These conventional techniques may include interviews, participant observation and the use of documents. This recommendation is similar to recommendation B, except that in that case the triangulated data is obtained from traditional data sources. Also, the purpose of data triangulation is to address the challenge of ensuring authenticity, whereas in this case, the goal is an attempt to ensure that some members of a population sample are not excluded (assuming their participation would be of relevance to the research).

E. Data, such as screenshots, images, videos, and texts may need to be gathered and analyzed, often requiring real-time analysis, while actively engaged in online fieldwork.

Social media platforms contain many different kinds of digital texts (Urquhart & Vaast, 2012). Hence IB researchers may need to gather and analyze some of these texts, images, videos, and screenshots. Learning how to capture and code digital texts, including a screenshot, images, and videos, may require the researcher to learn new skills. In the case of research in Angola and Mozambique, pictures were used to capture what was happening in the field, as well as screenshots of data from social media sites. These screenshots were loaded into NVivo, and some of them included in the research, where sections of the image could be coded. This data was then triangulated with other sources of data, such as the in-country semi-structured interviews and WhatsApp data collected during communication exchanges during the interviewees during the data gathering phase of the research.

F. Consider posting a disclaimer informing participants of data collection intent and its purposes

For IB qualitative researchers, maintaining respondent confidentiality while presenting rich, detailed accounts of data gathered through social media presents unique challenges. These challenges are not appropriately addressed in the literature on research ethics and research methods. Consider posting a disclaimer informing participants of data collection intent and purposes, and carefully consider the population target for the research. One way of doing so is by re-envisioning the informed consent process, which allows qualitative researchers to avoid confidentiality dilemmas that might otherwise lead them not to report rich, detailed data. Given that qualitative studies often contain rich descriptions of study participants, confidentiality breaches via deductive disclosure are of particular concern to IB qualitative researchers. As such, qualitative researchers face a conflict between conveying detailed, accurate accounts of the social media domain and protecting the identities of the individuals who participated in their research (Kaiser, 2009).

G. Attempt to obtain consent whenever possible on sites where an expectation of privacy exists

Discussing confidentiality at the outset is necessary for acquiring informed consent and building trust with respondents (Crow et al., 2006). However, these exchanges occur without knowledge of the specific information subsequently shared by the respondent. Furthermore, discussions about informed consent and confidentiality are rarely ongoing; once the consent form is signed researchers lack a standardized way of returning to the issue of privacy and data use with respondents. Matters regarding confidentiality should be disclosed at the time of data collection when IB qualitative researchers should make assurances of confidentiality, typically via informed consent forms, but taking into consideration co-location factors such as in-country language, as well as meaning, culture, and values. Statements such as, “all identifying characteristics, such as occupation, city, and ethnic background, will be changed” should be considered (Sieber, 1992, p. 52).

H. Obtain permission to access social media area

Although approval may not be needed on some publicly available social media sites, on others the researcher may need to gain access to specific potentially restricted areas of the platform or website. In some cases, permission needs to be obtained from one or more gatekeepers to acquire the necessary level of access or pay for an upgrade to access “premium” services, such as was the case with LinkedIn in our experience.

I. If needed, become a temporary member of a restricted area

Similar to recommendation H, there may be a need to obtain access or permission, or even pay for upgraded services in a social media platform, to be able to access relevant users and data.

J. Consider creating and using an avatar(s)

When attempting to obtain the appropriate access rights to a restricted area at some social media sites, researchers might need to create and use an avatar, or a few (Schultze, 2010), which can be in human form or some other fantasy-based form. The avatar becomes the researcher's identity when conducting the research. Researchers might also have to choose an online name (Hagström, 2008), or *screen name*. When doing so, researchers should take into consideration the values or cultures of the participants when choosing an identity. For

example, the first author selected the screen name (MGRider) based on his first and last name initials and the fact he is an avid road cyclist. However, for some participants, they associated it with someone that could be *taking them for a ride*, an untrustworthy individual. In some sites, therefore, it may be prudent to consider creating an account with the researcher's real identity, but in others an assumed one might be fine. In some online worlds, such as SecondLife.com, it is possible for a researcher to create and use multiple avatars, which might enable an understanding of the research problem from various perspectives. Using various avatars is considered normal behavior in several of these online worlds and would not be seen as abnormal behavior.

K. Get acquainted with and socialized into the social media platform (e.g., memberships in LinkedIn, Facebook, Twitter, and so forth.)

The researcher must become acquainted with and socialized into the social media platform, which is similar to many other forms of qualitative research in which the investigator "immerses himself or herself in the life of the social group under study" (Myers, 1999, p. 4). In the case of social media, this might mean finding out who the most influential people are on the site, which may also mean negotiating with them to obtain access to specific restricted areas. Becoming familiar with the culture of a social media platform involves being able to understand visual cues, possibly learning the language of the social media platform, and perhaps discovering new types of behavior. As social media sites, LinkedIn is very different from Facebook, despite the fact one may notice a trend of convergence from both sides. Since a certain level of proficiency is required in LinkedIn, the only way to gain this expertise is to spend time conversing with members, becoming familiar with discussion topics on the site that may be related to the research at hand.

In preparing for the previous study, the first author invested over 2200 hours reading, posting and interacting with LinkedIn members. This increased familiarity with MNEs from Angola and Mozambique, key decision-makers, and influencers in the country, as well as the issues and concerns they were facing during the time of the research. This led to a somewhat surprising new type of behavior discovered during the fieldwork in Mozambique, that, despite the level of education, whether having a college degree or not, women there must obtain the consent of their parents, or spouses, to be able to work, to become a practitioner. The unearthing of these cultural practices provided more insight about the co-evolution of the moral fabric and behavior of female professionals in Mozambique, which caused us to modify the semi-structured interviews in that country. Furthermore, the first author was already fluent in the Portuguese language and somewhat familiar with Angolan and Mozambican culture, and this familiarity meant that the participants were willing to share their opinions, thus increasing our confidence in the findings (see also Ameripour et al., 2010).

L. Use mixed methods to triangulate various types of data

Given that participants on social media sites and platforms may be anonymous or use pseudonyms, questions might arise about the authenticity of the data. Researchers should consider using mixed methods to triangulate different types of data to increase the level of confidence in the findings or to provide additional insights. One way to do this is to triangulate different kinds of data within the social media sites and platforms. In Angola

and Mozambique, four different types of data including chat logs, discussion forum posts, online survey, and participant observation were used in triangulation. Each of the data types can confirm or challenge a different perspective on the same phenomenon under exploration. For instance, in our research in Africa, the first author matched the interview notes with text downloaded from LinkedIn discussion forums and chatted logs. Additionally, the researcher could experience the issues relating to the interview notes directly through his participant observations. These data sources, when combined, provided a more in-depth picture of the situation. Alternatively, a researcher could triangulate the data obtained from the social media platform with data collected external to the platform (e.g., websites or interviews).

M. Subject's identity should not be important when developing research questions

To overcome the problem of authenticity, aside from triangulating data, researchers may develop research questions where the identity of participants would not be necessary. In the Angolan and Mozambican study, knowing the personal identification of the participants was not essential for answering our primary research question, only their role in the MNEs.

N. Inquiry the group on whether they consider available information public or private

Social media has drastically changed the way people interact with their friends, associates, and family members. Although social media, like Twitter, LinkedIn, Facebook, Google+, YouTube, Snapchat, and FourSquare, play a significant role as a vast repository of unstructured data, it can also pose serious privacy risks. Hence, when IB researchers use social media as a tool for gathering data, it is essential to acknowledge and understand the privacy risks involved. If not clear, researchers should attempt to find out from group whether its members consider available information public or private.

O. Take steps to ensure the reliability and validity of research findings

The conclusions must be trustworthy, consistent, relevant and reliable to be beneficial to any reader and researcher. Reliability of data is directly correlated to consistency in which the research must produce the same results if replicated. In turn, validity relates to the accuracy, or correctness, of the findings (see also Golder et al., 2017).

P. Become familiarized with the various social media platforms and features (e.g., the use of avatars, filters, augmented reality, and so forth.)

With such a large percentage of the public using social media, qualitative researchers cannot afford to interact with research participants via channels like Facebook, Twitter, and other social media platforms. However, while using these social media platforms may enhance qualitative research, not all of them are created equal. It is essential for to use platforms where population targets are more likely to be reached and engaged. To choose the best social media channels, researchers must take some time to familiarize themselves with each platform, how it runs and what demographics use that platform.

Q. Apply a filtering or data mining technique

This recommendation is related to the challenge of collecting a large volume of data. It is essential to use data mining techniques to filter or reduce the vast amount of text. There is some disagreement, however, among researchers about how much data to code in

qualitative research (Saldaña, 2009). For instance, Lofland, Snow, Anderson, and Lofland (2006) recommend coding the entire dataset, while Seidman (2006) argues that only the most significant data should be coded. Our recommendation for using social media data in qualitative research is to follow Seidman's approach. With such a large data set, it is virtually impossible for a qualitative researcher to analyze the entire set, and in any case, much of the data may not be as relevant or at least not exceptionally useful for answering the research question(s). Although at this may feel like searching for a needle in a haystack, it may be more practical for the researcher to filter the data and exclude what is irrelevant to the study. For instance, research for the Angolan and Mozambican study focused attention only on those discussion forum posts on LinkedIn that discussed MNE's internationalization mode (since it was directly related to the research question). Since the internationalization mode often varied from MNE's type and country, discussion forum posts were filtered on LinkedIn by focusing only on the data related to that theme. Researchers could use other methods, such as screening for keywords, events, research question specific concepts, or theoretical sensitivity. Payton and Kvasny (2012) organized their blog posting data chronologically to create a timeline of events.

R. Use qualitative data analysis software

Similar to recommendation C, using qualitative data analysis software such as NVivo or Leximancer can be imperative. However, researchers need to be careful with the fact that discussion forum data contain many threads and posts about an unlimited number of topics, most of which might be irrelevant to the research project. Therefore, the researcher may need to filter the data before using the software. In the illustrative case, the NVivo analysis was producing many inappropriate and irrelevant results before filtering. It was essential to filter the data once again after the NVivo analysis. For instance, the data from LinkedIn discussion forums often had HTML tags embedded within them. As these tags are text, they were included in the results. NVivo contains a pre-defined set of stop words which are skipped over by the algorithm. Since it is possible to edit the stop word list, HTML tags were added to the list, and the algorithm was rerun. This iterative process was completed many times to remove words unnecessary for the analysis. Other words were also added to the stop word list based on the prior knowledge of the researcher.

VII. Conclusions and Research Limitations

The use of social media as a tool for gathering data in IB qualitative researchers holds great potential. The introduction of big data on the Internet, which for the most part is unstructured textual data, offers a potentially valuable new source of qualitative data, which has now become available for analysis. However, as discussed in this paper, few qualitative researchers in international business are currently using social media or utilizing the vast amounts of qualitative data it creates and makes available for analysis at its sites. Quantitative

researchers in information systems are making good use of qualitative data on social media platforms, but qualitative researchers in general, and in particular in IB are not. Hence, this paper has looked at the potential use of social media in qualitative research in IB, discussed some of the challenges of using social media, and made some recommendations.

The challenges of using social media as a data gathering tool, or the analysis of unstructured data generated by it, in IB qualitative research are many. These challenges are related to the large volume of unstructured data, the nature of digital texts, visual cues, and the several types of user's comportment on social media sites, the authenticity of the data, the level of access obtained, and the digital divide in some situations, especially in frontier markets research. To discuss these challenges, we have made eighteen recommendations for conducting qualitative research using social media as a tool for gathering data or data analysis.

While some of our recommendations are similar to those that might be made for more traditional kinds of research, others were entirely different. First, international business researchers might have to learn new skills to gather and analyze social media data. Even if an IB researcher, adopting the outside observer role, uses commercially available web scraping applications for collecting the data, the typically large volume of data that is available or generated will undoubtedly require the use of some data filtering technique. To this end, IB qualitative researchers will need to learn and use a qualitative data analysis application (QDAA) package. Since not all QDAA packages have the same capabilities and features, however, it might be necessary for researchers to become skilled in more than one application package.

Second, while it is common practice for IB field researchers to immerse themselves in a field site, make sure to learn any used new language/jargon and become accustomed to foreign culture's idiosyncrasies. Also, try to learn about new types of behavior, even though such efforts

will likely still be unsatisfactory for conducting IB qualitative research using social media, such as in the case when adopting the involved researcher role. For instance, the only way to gain access to relevant information about MNEs in Angola and Mozambique on LinkedIn was by paying the Premium membership, to be able to conduct broader and more advanced searches through a larger pool of potential enterprises and professionals to be investigated. Also, to be worth the investment for paying for Premium membership in LinkedIn, it was necessary to attain a reasonably high level of proficiency of using the Premium features available on LinkedIn. Without a certain level of competence, the first study would not have enabled interaction and observation for any of the activities of the professionals from Angola and Mozambique, potential targets of the research. Consequently, this process reinforces the argument that researching on social media platforms will likely require a specific set of skills that may take time and commitment to master.

Third, researching involving social media, as a data gathering tool or tapping onto already available data, tends to lead a broader use of visual images, which can often be useful and enhance the understanding of research themes. Many journal publishers, however, are unwilling to include colorful images, or too many graphs and images, in the print version of the journal, given the extended length of the paper, additional expenses, or both. Conceivably IB researchers using social media can consider creative ways of including images and animations in their article by, for instance, embedding QR codes, web links to video clips on YouTube, images on Pinterest, in their manuscripts or providing an online appendix. The use of images is not such a problem with purely electronic journals, however, but it can be a problem for print journals in IB and other fields.

Fourth, the use of social media introduces new international business ethical challenges across borders for qualitative researchers. It is often challenging to determine the appropriate balance between ensuring the plausibility, validity, and rigor of the research findings, while at the same time abiding by international business ethical principles, which tends to vary significantly due to local laws, customs, and religious beliefs. Such obstacles, for instance, can be a significant challenge for informed consents.

Research Limitations

Of course, this paper, and its conclusions has limitations. Foremost, while we believe we have mentioned most of the challenges that international business qualitative researchers are likely to face when using social media as a data gathering tool or for analysis of its vast unstructured data available, we do not claim to have covered all of them. On the contrary, we are confident that more challenges exist and that new ones will continue to surface in the IB qualitative research horizon. We also have included and discussed in this paper only one example of qualitative research using social media, and therefore the challenges and lessons we debated might not necessarily be generalizable to the international business research field, nor to all types of social media tools and sites available as of today on the Internet. Lastly, we acknowledge and understand that the recommendations offered in this paper will likely not apply to every given IB qualitative research endeavor. It will be up to each IB researcher to decide which ones may be applicable.

Future international business qualitative research methods papers focused on the use of social media as a tool, or as a source of data, it is hoped, should address some of the limitations of this paper, and perhaps explore new opportunities, such as the high potential for collaboration between qualitative and quantitative IB researchers. The motivation in writing this paper was just

to inspire international business qualitative researchers to consider and start using social media as a tool, or its data in their research efforts, and to propose strategies in carrying on such research efforts. As social media applications continue to evolve, unquestionably new opportunities and challenges for IB qualitative researchers are sure to emerge.

VIII. References

- Ackland, R. (2013). *Web social science: Concepts, data and tools for social scientists in the digital age*. Sage.
- Alexa. (2015). The top 500 sites on the web. Retrieved from <http://www.alexa.com/topsites> Google Scholar
- Alvesson, M., & Sköldböck, K. (2009). *Reflexive methodology: New vistas for qualitative research* (2nd ed.). London, England: Sage.
- Ameripour, A., Nicholson, B., & Newman, M. (2010). Conviviality of Internet social networks: An exploratory study of Internet campaigns in Iran. *Journal of Information Technology*, 25(2), 244–257.
- Anderson, E., & Gatignon, H. (1986). Modes of entry: A transaction cost analysis and propositions. *Journal of International Business Studies*, 17 (3), 1-26.
- Aviles, M & Bauman, S. (2017) *Working Hand in Hand: Qualitative MR and Social Media Research*, American Marketing Association, <https://www.ama.org/publications/eNewsletters/MarketingInsightsNewsletter/Pages/working-hand-in-hand-qualitative-mr-and-social-media-research.aspx>, last accessed on 02/01/2018.
- Best, P., Manktelow, R., Taylor, B. (2014). Online communication, social media and adolescent wellbeing: A systematic narrative review. *Children and Youth Services Review*, 41, 27-36. doi:10.1016/j.chilyouth.2014.03.001 Google Scholar
- Błachnio, A., Przepiórka, A., Rudnicka, P. (2013). Psychological determinants of using Facebook: A research review. *International Journal of Human-Computer Interaction*, 29, 775-787. doi:10.1080/10447318.2013.780868 Google Scholar
- Boyd, D. (2009). Social media is here to stay... Now what? Paper presented at the Microsoft Research Tech Fest. Retrieved from <http://www.danah.org/papers/talks/MSRTechFest2009.html>.
- Bradley, A. (2010). A new definition of social media. Gartner blog network. Retrieved from http://blogs.gartner.com/anthony_bradley/2010/01/07/a-new-definition-ofsocial-media/.
- Caers, R., De Feyter, T., De Couck, M., Stough, T., Vigna, C., Du Bois, C. (2013). Facebook: A literature review. *New Media & Society*, 15, 982-1002. doi:10.1177/1461444813488061 Google Scholar, Link
- Christopher, T. (2009). In-game identities and meatspace mistakes. In L. Cuddy, & J. Nordlinger (Eds.), *World of Warcraft and philosophy: Wrath of the philosopher king* (pp. 165–171). Chicago: Open Court Publishing.

- Coleman, G. (2010). Ethnographic approaches to digital media. *Annual Review of Anthropology*, 39, 487–505.
- Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Thousand Oaks, CA: Sage.
- Crow G, Wiles R, Heath S, Charles V. (2006) Research ethics and data quality: The implications of informed consent. *International Journal of Social Research Methodology*;9:83–95.
- Dhir, A., Buragga, K., Boreqqah, A. A. (2013). Tweeters on campus: Twitter a learning tool in the classroom? *Journal of Universal Computer Science*, 19, 672-691. doi:10.3217/jucs-019-05-0672 Google Scholar
- Diaz Andrade, A., Urquhart, C., & Arthanari, T. S. (2015). Seeing for understanding: Unlocking the potential of visual research in information systems. *Journal of the Association for Information Systems*, 16(8), 3.
- Duggan, M. (2013). Photo and video sharing grow online. Pew Research Center. Retrieved from <http://pewinternet.org/Reports/2013/Photos-and-ideos.aspx> Google Scholar
- Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). Social media update 2014. Pew Research Center. Retrieved from <http://www.pewinternet.org/2015/01/09/socialmedia-update-2014/>
- Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., Madden, M. (2015). Social media update 2014. Pew Research Center. Retrieved from <http://www.pewinternet.org/2015/01/09/social-media-update-2014/> Google Scholar
- Dunning, J. H. (1993). *The globalization of business*. London: Routledge.
- Dunning, J.H. (1981). Explaining the international direct investment position of countries: towards a dynamic or developmental approach. *Weltwirtschaft-liches Archiv*, 117(1), 30-64.
- Facebook. (2018). About Facebook: Basic info. Retrieved from <https://www.facebook.com/facebook/info>
- Fox, J., Warber, K. M., & Makstaller, D. C. (2013). The role of Facebook in romantic relationship development: An exploration of Knapp's relational stage model. *Journal of Social and Personal Relationships*, 30, 771–794. doi:10.1177/0265407512468370
- Fuchs, C. (2013). *Social media: A critical introduction*. London: Sage.
- Garcia, A., Standlee, A., Bechkoff, J., & Cui, Y. (2009). Ethnographic approaches to the Internet and computer-mediated communication. *Journal of Contemporary Ethnography*, 38, 52–84.
- Germonprez, M., & Hovorka, D. S. (2013). Member engagement within digitally enabled social network communities: New methodological considerations. *Information Systems Journal*, 23(6), 525–549.
- Golder, S., Ahmed, S., Norman, G., & Booth, A. (2017). Attitudes Toward the Ethics of Research Using Social Media: A Systematic Review. *Journal of Medical Internet Research*, 19(6), e195. <http://doi.org/10.2196/jmir.7082>
- Goncalves, M. & Cornelius Smith, E. (2017). Internationalization strategies of frontier Lusophone-African multinational enterprises: Comparative case studies of Angola and Mozambique. *Journal of Transnational Management*, 223(3), 203-232. 10.1080/15475778.2017.1335127.
- Green, S. (2014, September 4). Revolution And Retweets Social Media And Political Change. *Micro Mart*, (1327), 64.
- Greene, J. A., Choudhry, N. K., Kilabuk, E., & Shrank, W. H. (2011). Online social networking by patients with diabetes: A qualitative evaluation of communication with Facebook. *Journal of General Internal Medicine*, 26, 287–292. doi:10.1007/s11606-010-1526-3

- Hamm, M. P., Chisholm, A., Shulhan, J., Milne, A., Scott, S. D., Klassen, T. P., Hartling, L. (2013). Social media use by health care professionals and trainees: A scoping review. *Academic Medicine : Journal of the Association of American Medical Colleges*, 88, 1376-1383. doi:10.1097/ACM.0b013e31829eb91c Google Scholar, Crossref, Medline.
- Hanna, P. (2012). Using internet technologies (such as Skype) as a research medium: A research note. *Qualitative Research*, 12(2), 239–242.
- Harden, A., & Thomas, J. (2010). Mixed methods and systematic reviews: Examples and emerging issues. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (2nd ed., pp. 749–774). Thousand Oaks, CA: Sage.
- Hunt, N., & McHale, S. (2007). A practical guide to the e-mail interview. *Qualitative Health Research*, 17(10), 1415–1421.
- Jones, E., Sinclair, J. M. A., Holt, R. I. G., Barnard, K. D. (2013). Social networking and understanding alcohol-associated risk for people with Type 1 diabetes: Friend or foe? *Diabetes Technology & Therapeutics*, 15, 308-314. doi:10.1089/dia.2012.0327 Google Scholar
- Jowett, A., Peel, E., & Shaw, R. (2011). On-line interviewing in psychology: Reflections on the process. *Qualitative Research in Psychology*, 8, 354–369.
- Jukic, T., & Merlak, M. (2016). (Un)Exploited Potential of Social Networks in Public Organizations: The Case of Slovenia. *European Conference on E-Government*, 98.
- Kaiser, K. (2009). Protecting Respondent Confidentiality in Qualitative Research. *Qualitative Health Research*, 19(11), 1632–1641. <http://doi.org/10.1177/1049732309350879>
- Kaplan, A. M., Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53, 59-68. doi:10.1016/j.bushor.2009.09.003 Google Scholar
- Kim, Y., Huang, J., & Emery, S. (2016). Garbage in, Garbage Out: Data Collection, Quality Assessment and Reporting Standards for Social Media Data Use in Health Research, Infodemiology and Digital Disease Detection. *Journal of Medical Internet Research*, 18(2), e41. <http://doi.org/10.2196/jmir.4738>
- Kozinets, R. V. (2010). *Netnography. Doing ethnographic research online*. London: Sage Publications Ltd.
- Light, B., McGrath, K. (2010) "Ethics and social networking sites: a disclosive analysis of Facebook", *Information Technology & People*, Vol. 23 Issue: 4, pp.290-311, <https://doi.org/10.1108/09593841011087770>
- Lofland, J., Snow, D., Anderson, L., & Lofland, L. H. (2006). *Analyzing social settings: A guide to qualitative observation and analysis*(4th ed.). Belmont, CA: Thomson Wadsworth.
- Lovink, G. (2011). *Networks without a cause: A critique of social media*. Cambridge: Polity Press.
- Manca, S., Ranieri, M. (2013). Is it a tool suitable for learning? A critical review of the literature on Facebook as a technology-enhanced learning environment. *Journal of Computer Assisted Learning*, 29, 487-504. doi:10.1111/jcal.12007 Google Scholar
- Marschan-Piekkari, R., & Welch, C. (2004) *Qualitative Research Methods in International Business: The State of the Art*, in Marschan-Piekkari, R., and Welch, C., *Handbook of Qualitative Research Methods in International Business*, Cheltenham, Northampton: Edward Elgar, pp.5-24.

- McKenna, B., Gardner, L., & Myers, M. D. (2012). The co-evolution of the “social” and the “technology”: A netnographic study of social movements in virtual worlds. Paper presented at the International Conference on Information Systems, Orlando, FL.
- Meikle, G., & Young, S. (2012). *Media convergence: Networked digital media in everyday life*. Basingstoke: Palgrave Macmillan.
- Morgan, E. M., Snelson, C., & Elison-Bowers, P. (2010). Image and video disclosure of substance use on social media websites. *Computers in Human Behavior*, 26, 1405–1411. doi:10.1016/j.chb.2010.04.017
- Müller, O., Junglas, I., & vom Brocke, J., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems*, 25, 289–302.
- Myers, M. D. (2013). *Qualitative research in Business & Management*(2nd ed.). London: Sage Publications.
- Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, 17(1), 2–26.
- Nadkarni, A., Hofmann, S. G. (2012). Why do people use Facebook? *Personality and Individual Differences*, 52, 243-249. doi:10.1016/j.paid.2011.11.007 Google Scholar
- Oh, O., Eom, C., & Rao, H. R. (2015). Role of social media in social change: An analysis of collective sense making during the 2011 Egypt Revolution. *Information Systems Research*, 26(1), 210–223.
- Okoroafo, S. C. (1990). An assessment of critical entry factors affecting modes of entry substitution patterns in foreign product markets. *Journal of Global Marketing*, 3 (3), 87-104.
- Papacharissi, Z. (2002). The virtual sphere: The internet as the public sphere. *New Media & Society*, 4(1), 5–23.
- Payton, F. C., & Kvasny, L. (2012). Considering the political roles of Black talk radio and the Afrosphere in response to the Jena 6: Social media and the blogosphere. *Information Technology & People*, 25(1), 81–102.
- Payton, F. C., & Kvasny, L. (2012). Considering the political roles of Black talk radio and the Afrosphere in response to the Jena 6: Social media and the blogosphere. *Information Technology & People*, 25(1), 81–102.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. London: Sage.
- Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*(3rd ed.). New York: Teachers College Press.
- Shi, Z., Rui, H., & Whinston, A. B. (2014). Content sharing in a social broadcasting environment: Evidence from twitter. *MIS Quarterly*, 38(1), 123–142.
- Shirky, C. (2011). The political power of social media: Technology, the public sphere, and political change. *Foreign Affairs*, 90(1), 28–41.
- Sieber, J. (1992) *Planning ethically responsible research: A guide for students and internal review boards*. Newbury Park: Sage.
- Snelson, C. (2011). YouTube across the disciplines: A review of the literature. *MERLOT Journal of Online Learning and Teaching*, 7, 159-169. Retrieved from http://jolt.merlot.org/vol7no1/snelson_0311.htm Google Scholar
- Snelson, C. (2011). YouTube across the disciplines: A review of the literature. *MERLOT Journal of Online Learning and Teaching*, 7, 159–169. Retrieved from http://jolt.merlot.org/vol7no1/snelson_0311.htm

- Snelson, C. (2016). Qualitative and Mixed Methods Social Media Research: A Review of the Literature. *International Journal of Qualitative Methods*, 15(1), 1-15.
- Statista. (2015). Number of social network users worldwide from 2010 to 2018 (in billions). Retrieved from <http://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/> Google Scholar
- Tsang, P. P., Au, M. H., Kapadia, A., & Smith, S. W. (2010). BLAC: Revoking repeatedly misbehaving anonymous users without relying on TTPs. *ACM Transactions on Information and System Security*, 13(4), 31–33 (39).
- Twitter. (2018). Twitter milestones: A selection of memorable moments from our company. Retrieved from <https://about.twitter.com/milestones>
- Urquhart, C., & Vaast, E. (2012). Building social media theory from case studies: A new frontier for IS research. Paper presented at the Thirty Third International Conference on Information Systems (ICIS).
- Vaast, E., & Levina, N. (2015). Speaking as one, but not speaking up: Dealing with new moral taint in an occupational online community. *Information and Organization*, 25(2), 73–98.
- Vaast, E., Davidson, E. J., & Mattson, T. (2013). Talking about technology: The emergence of a new actor category through new media. *MIS Quarterly*, 37(4), 1069–1092.
- Vijayan, J. (2015, June 25). Solving the Unstructured Data Challenge. *CIO Magazine*. Retrieved from <http://www.cio.com/article/2941015/big-data/solving-theunstructured-data-challenge.html>.
- Vyas, A. N., Landry, M., Schnider, M., Rojas, A. M., & Wood, S. F. (2012). Public health interventions: Reaching Latino adolescents via short message service and social media. *Journal of Medical Internet Research*, 14, 1–14. doi:10.2196/jmir.2178
- Walsham, G. (1995). Interpretive case studies in IS research: Nature and method. *European Journal of Information Systems*, 4(2), 74–81.
- Walsham, G. (2006). Doing interpretive research. *European Journal of Information Systems*, 15(3), 320–330.
- Wang, Y., Min, Q., & Liu, Z. (2014). A meta-analytic review of social media studies. Paper presented at the Proceedings of the Pacific Asia Conference on Information Systems (PACIS), Chengdu, China.
- Wasko, M., Teigland, R., Leidner, D., & Jarvenpaa, S. (2011). Stepping into the Internet: New ventures in virtual worlds. *MIS Quarterly*, 35(3), 645–652.
- Whelan, E., Teigland, R., Vaast, E., & Butler, B. (2016). Expanding the horizons of digital social networks: Mixing big trace datasets with qualitative approaches. *Information and Organization*, 26(1), 1–12.
- Williams, S. A., Terras, M. M., Warwick, C. (2013). What do people study when they study Twitter? Classifying Twitter related academic papers. *Journal of Documentation*, 69, 384-410. doi:10.1108/JD-03-2012-0027 Google Scholar
- Wilson, R. E., Gosling, S. D., Graham, L. T. (2012). A review of Facebook research in the social sciences. *Perspectives on Psychological Science*, 7, 203-220. doi:10.1177/1745691612442904 Google Scholar
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks. YouTube. (2018). About YouTube. Retrieved from <https://www.youtube.com/yt/about/>
- Zimmer, M. (2010). “But the data is already public”: On the ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313–325.