SOTICS 2018

The Eighth International Conference on Social Media Technologies, Communication, and Informatics

ISBN: 978-1-61208-673-6

October 14 - 18, 2018

Nice, France

SOTICS 2018 Editors

Nitin Agarwal, University of Arkansas at Little Rock, USA

SOTICS 2018

Forward

The Eighth International Conference on Social Media Technologies, Communication, and Informatics (SOTICS 2018), held on October 14 - 18, 2018- Nice, France, was an event on social eco-informatics, bridging different social and informatics concepts by considering digital domains, social metrics, social applications, services, and challenges.

The systems comprising human and information features form a complex mix of social sciences and informatics concepts embraced by the so-called social eco-systems. These are interdisciplinary approaches on social phenomena supported by advanced informatics solutions. It is quit intriguing that the impact on society is little studied despite a few experiments. Recently, also Google was labeled as a company that does not contribute to brain development by instantly showing the response for a query. This is in contrast to the fact that it has been proven that not showing the definitive answer directly facilitates a learning process better. Also, studies show that e-book reading takes more times than reading a printed one. Digital libraries and deep web offer a vast spectrum of information. Large scale digital library and access-free digital libraries, as well as social networks and tools constitute challenges in terms of accessibility, trust, privacy, and user satisfaction. The current questions concern the trade-off, where our actions must focus, and how to increase the accessibility to eSocial resources.

We take here the opportunity to warmly thank all the members of the SOTICS 2018 technical program committee, as well as all of the reviewers. We also kindly thank all the authors who dedicated much of their time and effort to contribute to SOTICS 2018. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

We also gratefully thank the members of the SOTICS 2018 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success.

We hope that SOTICS 2018 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the area of social eco-informatics. We also hope Nice provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

SOTICS 2018 Steering Committee

Lasse Berntzen, University College of Southeast Norway, Norway
Nitin Agarwal, University of Arkansas at Little Rock, USA
Andrea Nanetti, School of Art, Design, and Media | Nanyang Technological University, Singapore

SOTICS 2018 Industry/Research Advisory Committee

Roman Shtykh, Yahoo Japan Corporation, Japan
Xin Shuai, Thomson Reuters, USA
Andrea Cimino, Institute for Computational Linguistics (ILC-CNR), Pisa, Italy

SOTICS 2018

Committee

SOTICS Steering Committee
Lasse Berntzen, University College of Southeast Norway, Norway
Nitin Agarwal, University of Arkansas at Little Rock, USA
Andrea Nanetti, School of Art, Design, and Media | Nanyang Technological University, Singapore

SOTICS Industry/Research Advisory Committee
Roman Shtykh, Yahoo Japan Corporation, Japan
Xin Shuai, Thomson Reuters, USA
Andrea Cimino, Institute for Computational Linguistics (ILC-CNR), Pisa, Italy

SOTICS 2018 Technical Program Committee

Witold Abramowicz, Poznan University of Economics and Business, Poland
Nitin Agarwal, University of Arkansas at Little Rock, USA
Swati Agarwal, BITS Pilani, Goa Campus, India
Ahmet Aker, University of Duisburg-Essen, Germany / University of Sheffield, UK
Sultana Lubna Alam, University of Canberra, Australia
Frédéric Amblard, IRIT | Université Toulouse 1 Capitole, France
Esther Andrés Pérez, ISDEFE / Technical University of Madrid, Spain
Mehdi Asgarkhani, Ara Institute of Canterbury, New Zealand
Liz Bacon, University of Greenwich, UK
Grigorios N. Beligiannis, University of Patras, Greece
Gerardo Berbeglia, Melbourne Business School, Australia
Lasse Berntzen, University College of Southeast Norway, Norway
Brian Blake, University of Arkansas at Little Rock / Harding University / Acxiom Corporation, USA
Christos Bouras, University of Patras | Computer Technology Institute & Press «Diophantus», Greece
Piotr Bródka, Wrocław University of Science and Technology, Poland
María Luisa Carrió Pastor, Universitat Politècnica de València, Spain
Ilknur Celik, Cyprus International University, Northern Cyprus
Subrata Chakraborty, University of Southern Queensland, Australia
Manoj K. Chinnakotla, Artificial Intelligence and Research (AI & R), Microsoft, India
Christina Christodoulakis, University of Toronto, Canada
Andrea Cimino, Institute for Computational Linguistics (ILC-CNR), Pisa, Italy
Francesco Corcoglioniti, Fondazione Bruno Kessler - Trento, Italy
Alejandro Cortiñas, University of A Coruña, Spain
Jay Daniel, University of Derby, UK
Gert-Jan de Vreede, University of South Florida, USA / Management Center Innsbruck, Austria
Arianna D'Ulizia, National Research Council - IRPPS, Italy
Sourav Dutta, Nokia Bell Labs, Ireland
Aviad Elyashar, Ben-Gurion University of the Negev, Beer Sheva, Israel
Larbi Esmahi, Athabasca University, Canada

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

Towards a Social Media Research Methodology: Defining Approaches and Ethical Concerns

James Baldwin,
Department of Computing,
C3RI, Sheffield Hallam University,
Sheffield, United Kingdom,
email: j.baldwin@shu.ac.uk

Dr Teresa Brunsdon &
Dr Jotham Gaudoin,
Department of Engineering and Mathematics, Sheffield Hallam University, United Kingdom
email: t.m.brunsdon@shu.ac.uk
j.gaudoin@shu.ac.uk

Dr Laurence Hirsch,
Department of Computing,
Sheffield Hallam University,
United Kingdom
email: l.hirsch@shu.ac.uk

***Abstract*— Social media research and suitable methodologies and ethical approaches for analysing social media data are still emerging. This paper presents a methodology for projects using social media data alongside consideration of ethics within the social media analysis context. Earlier stages of the methodology will be expanded to develop a strategy for examining ethics alongside consideration of the relevant analysis techniques that may be employed. This will provide a comprehensive methodology that will provide a springboard for the clear and ethically sound scrutiny of social media data. We aim to present the challenges of using social media data, while the inclusion of ethical and legal aspects in this paper aim to draw researchers' attention to the peculiarity issues involved with dealing with social media data.**

Keywords—social media; methodology; strategy; methods; ethics; legal; lifecycle.

I. INTRODUCTION

Since 2011, interest has grown in social media from both the academic and industrial perspectives [1]. For example, Law Enforcement Agencies substantially increased their usage of social media data, with policy changes being implemented to adapt to social media and its possible uses after the 2011 London riots occurred [2][3]. This interest has to some extent been driven by the rapid increase in usage of social media networks and of internet accessibility; the internet was used daily or almost daily by 82% (41.8 million) of UK adults, compared with 78% (39.3 million) in 2015 and 35% (16.2 million) in 2006 [4]. Organisations now have social media teams to monitor events and actively release information, quickly reacting to situations of widespread interest [1]. A great deal of research both has helped to shape the future of social media research, but this remains in its infancy. Examples of this inside the UK include the Government Social Researchers [1], a research team within the UK government "ensuring ministers and policy makers have the data to understand social issues [5] and evaluating the policy responses to them", the Economic and Social Research Council, Ipsos MORI [6] and the Centre for Analysis of Social Media - part of a cross-party charity-run think tank DEMOS [7]. Outside the UK there are such things the Big Boulder Initiative [8] located in the United States, which markets itself as the “first trade association for the social data industry” and European Citizen Science Association in Europe that is looks to "connect citizens and science through fostering active participation" whether that is using social media or other platforms [9].

The Big Data characteristics of social media data as regards their volume, velocity and scope has created a need for methodological innovations that are suited towards investigating social media data and their overall lifecycle and which apply both qualitative and quantitative approaches [10]. Quantitative methods seem to be the most popular in research to date, but analyses are certainly not restricted to this approach [10]. For example, new approaches in qualitative research are being formed in areas ranging from narrative analysis, to so-called *thick* data that document human behaviour and the context of that behaviour, to the analysis of non-verbal data such as sound and images, to combining and linking data - both text and interactions - from different platforms across times and contexts. Given this vast, expanding area of research, scholars will need to acquire new skills to explore, analyse and visualise their findings and situate them into their appropriate contexts [10], and will also need to be able to make appropriate ethical considerations for their research.

There is a need for further development of a clear methodology drawing together the already extant building blocks of good practice displayed both in researchers' papers [11] and by organisations, such as Canadian’s government "Social Media Data Stewardship" (SMDS) project, that reduce the bias and flaws in social media data analysis. SMDS focuses on the data management processes applied in the context of using social media data. In the methodology section, we will discuss difficulties that are encountered when trying to find a social media lifecycle that has a clear defined strategy from start to finish, as without a clear approach to follow, such research can be a difficult experience for scholars embarking on work in this field.

The ethical perspective of extracting and collecting social media data in particular demands further consideration [10]. This is very important as it ensures the public’s data are protected and are represented in a fair and respectful manner, whereby a tweet or post is not been taken out of context or used inappropriately. Ethics must be taken into consideration when going through each stage of the methodological framework. This paper will focus on the social media research methodology process,

 ISBN: 978-1-61208-673-6

while simultaneously considering the relevant ethical concerns.

The sections to be covered in this paper will be in accordance with the social media project lifecycle presented in Section 2. This will look to build upon existing methodological frameworks for social media research and, in particular, the GSR's social media lifecycle. This lifecycle was not originally designed for research purposes, and so must be modified to be fit for such a purpose, but it will be seen that it provides us with a good starting point from which to begin. Other approaches will also be considered and these will be merged in order to create a hybridised lifecycle that forms the essence of the methodology presented here. In Section 3, we will discuss the ethical concerns that can impact the social media research strategy and its lifecycle. In Section 4, conclusions will be drawn from the paper.

II. METHODOLOGY

Section 2 discusses a series of social media research strategies and how they are integrated into our social media lifecycle.

A. Social Media Research Strategy

Upon reviewing a wide range of papers, it was noted [11]-[13] that some provided an excellent, thorough description of the steps they took in their research. However, it was often found that the initial stages of the research that would be needed for a complete addressing of any research question were poorly defined. The available literature tends to be project specific in its approach and is therefore not immediately suitable for generalisation to other research - not unexpected, given that social media research methodology is a topic still in its infancy. From an early researcher's standpoint in particular, it may be difficult to know where to start in the area and to identify what decisions need to be taken to form a social media methodology for the project in question.

The research community and other organisations are trying to come up with better ways to express their social media strategies, such as the SMDS project, which *"focuses on studying practices behind and attitudes towards the collection, storage, use, reuse, analysis, publishing and preservation of social media data"* [14]. SMDS has produced a social media data process that aims to clarify for researchers the layout and order of each phase that may be required in a social media data project. SMDS focuses on the data management process of social media data and aims to help researchers to consider their attitudes towards the data they wish to work with [14]. What we aim to do in this paper is to identify a *complete* set of stages for any social media research project lifecycle to follow, including within this the SMDS insights into data management, as these touch on highly pertinent points within the overall process.

Having found the nascent SMDS data management paradigm, we continued the search for a full social media project lifecycle. While this proved impossible to source as no such lifecycle yet exists, we did encounter a somewhat developed social media research project lifecycle created by the UK Government Social Research (GSR) service. The GSR based its lifecycle on the Cabinet Office framework for data science projects, as it had *"numerous parallels here"* [1, p8]. This lifecycle has been tested on two social media projects within Government, namely, using Twitter to predict cases of Norovirus and assessing the experiences of the 20th Commonwealth games held in Glasgow, producing reports on the analysis of broadcast and online coverage. There is no publically available information on whether or not this social media lifecycle was in fact a success. However, GSR produced outcomes that may be a measure for potential successes. For example, the Commonwealth games on Twitter were in the top 10 highest sporting event hashtags of the year, generating a highly positive contribution to Scotland and Glasgow both internationally and within the rest of the UK [15]. Furthermore, GSR identified that between 14/06/14 to 06/08/14, there were 3.2 million mentions of the Commonwealth Games on social media in the English language. There were other positive outcomes, but what this allows GSR to do is to identify where future improvements can be made with the organisers in raising the profile for relevant cities and events [1] [15]. In the sequel, we shall aim to integrate aspects of the GSR service lifecycle and the SMDS data management process alongside our own insights into the social media project lifecycle.

B. Our integrated social media project lifecycle

The GSR social media project lifecycle [1] consists of seven stages: Stage 1: Rationale – Business/Citizen Need, Stage 2: Data, Stage 3: Tools and Output, Stage 4: Research Phase, Stage 5: Implementation/Publication/Action, Stage 6: Evaluation and finally Stage 7: Business as Usual. While this is a useful basic framework that will help to guide researchers through their social media projects, it still requires further development and refinement as the considerations outlined at each stage are given in little detail. Furthermore, this lifecycle is applied in a commercial and governmental context, which can make it difficult to know what to do at each step from a research perspective. Nevertheless, we have chosen to adopt this framework as a starting point as it proved itself helpful in structuring our own initial social media research project. The research we are conducting aims to enhance the analysis of social media in the context of public (dis-)order events. This investigates how social media data are stored (big data issues), collected, analysed (text mining and sentiment analysis) and then disseminated (to the police, to help predict when disorder may occur). This will form part of the creation of a model to analyse social media data to try to predict the escalation of such events and our research is presently ongoing. We will adapt the GSR lifecycle to suit the needs, aims and goals of research projects (as opposed to governmental projects), and a diagram showing the relevant adaptations is displayed in Figure 1.

 ISBN: 978-1-61208-673-6

Figure 1. Social media research project lifecycle

The steps in the lifecycle are explained below. We will outline the purpose of each step and show where modifications have been made to the GSR lifecycle. The lifecycle explained below will be informed by the pilot study we conducted, which has involved analysing Twitter data around the time of the Baltimore riots, with the aim of developing models to identify potential riots before they occur.

1) In [1], stage 1 (Rationale – Business/ Citizen Need) is described as a need to think about social media's attributes (e.g. speed, cost, real-time production). On the basis of these attributes, there are suggestions for the business or citizen's need to be based on: *"using insight to deliver a more timely service to the citizen with fewer resources through the support of social media analysis than would have been possible with traditional means."* [1, p9]. To measure if the project is delivering a timely and resource efficient service to the citizen can be difficult to determine in some cases without actually conducting the project. A rationale for the research must be established, as without this the project will likely lack focus and be too broad, weakening any results or insights obtained. This means that valuable resource that could potentially be better utilised elsewhere is being wasted. While nothing new has been added to this section compared to the GSR lifecycle, we have placed into the appropriate research context. This stage in our process is important, as one must have a question to drive the collection and analysis of data in research and, as outlined by [10], one should not let the data drive the researcher. Without a suitable research question, the project would lack purpose. The rationale for the project we carried is outlined above.

2) Stage 2 is a new step which has been introduced called "Selection of Potential Method(s)s". This step is required to help adapt this commercial lifecycle into a research context where consideration must be given as to which methods (for example, case study or archival research) will be applied in the research process. This must be decided early on in the process, so that the following stages can take this into account when making relevant decisions in the latter phases of the lifecycle. If this step is not undertaken explicitly in a research context then results may be obtained that are of a particular nature, without account having been taken of the fact that the nature of the methods employed is inextricably linked with one's research outputs. This may cause a loss of momentum in the stages ahead, where special account would have to be made for the method or methods employed. For our particular research, we selected a case study-based approach to allow us to work with particular disorder events immediately and then attempt to generalise these to the wider public order context.

3) "Data" is now stage 3 of the lifecycle. In [1, p9] it is emphasised that *"The primary purpose of this data is not for research so consideration should be given to representativeness, robustness and ethics."* This statement is confusing, as the same level of rigour would apply in a research context. In this section, the researcher must justify the datasets to be used in the project and examine any necessary ethical considerations regarding the use of the social media data in question in their research. The original purpose of this section remains the same as in the original GSR lifecycle. This phase considers which dataset(s) may be explored to answer the research questions of the project. There is extra emphasis on selecting the correct data as cost may well be an issue here, more so than for a government entity, depending on the size of dataset required for the research, given the finite nature of research grants in particular. This step is also useful in providing time to think carefully about the selection of data-sets. If the data are chosen without due care then this will impact the cleaning, analysis and output of the project, though given the emerging nature of social media technology, it can of course be difficult to fully understand the range of data and metadata that are available before one already has a sample to hand. To that end, collection of a small pre-sample of data can also be a useful initial substage here. The dataset used for the pilot study is based on collecting live data from the 2015 Baltimore riots, USA. This pilot study will help to inform the collection of further datasets, on which the pre-processing and data manipulation scripts developed for the Baltimore data can be re-run.

4) Stage 4, "Tools and Outputs" is named the same as in the original GSR lifecycle. In this phase, the use of specialised social media tools can help to make cleaning and analysis of the collected data easier for researchers. Furthermore, social media data may require manipulation to *"render it useful in a social research setting"* [1, p9]. The outputs from analysis of these data can range from traditional reports showing present findings to predictive models designed to solve real time problems. GSR's process for this step is kept, but

 ISBN: 978-1-61208-673-6

in addition to this, the researcher must outline their data collection strategy to show how relevant data in relation to any research questions will be obtained, as well as considering how those data will be stored and whether single or multiple platforms are to be used as this will have an effect on the tools chosen. There are a plethora of tools available for data acquisition, processing and analysis and the tools to be used must be selected with care to ensure that they are both suitably secure and efficacious for the data in question, otherwise, time will be invested in tools that are not appropriate for large scale data retrieval (not all return the same metadata, for example), cleaning and/or analysis. The tools selected will depend upon the platform from which data are to be extracted. In our case, since we are dealing with Twitter data we chose NVivo NCapture to extract a live sample of data from the Baltimore riots and used R for data manipulation. For the retrospective datasets that we collect in the future, we will instead be using DiscoverText for acquisition. This tool is widely used in the research community because it provides access to one of the cheapest ways to retrieve a complete historical record from Twitter's official provider GNIP. Even though the extraction and analytical tools are being selected at this stage, the actual techniques for analysis will be investigated in stage 5.

5) Stage 5 was originally named "Research Phase" in the GSR lifecycle [1], rather than "Analysis". Clearly, given that we are aiming to develop a full research lifecycle, the former name is no longer appropriate. This step emphasises that care must be taken regarding the representativeness of data to mitigate any bias in the analysis. Lastly, *"Care should be taken to ensure research generates a dataset of a size which can be handled by the subsequent analytics programs."* [1, p10]. This is an important aspect to consider, as the volume of data produced can be on a very large scale. This could break the confines of some analytical programs' constraints. Other Big Data characteristics (namely: variety, veracity, velocity and virtue) and the type of techniques applied by the researcher can have an influence on the choice of analytical tool adopted to achieve their aim(s) [10]. The naming of this section has been selected to align with its focus on preparing the data for the analysis, helping to identify whether the chosen analytical tools need to be changed to handle the dataset(s) in question and to establish which techniques (in our case, change point identification, sentiment analysis and machine learning) should be applied to analyse the data to assist in responding to a research aim and answering relevant research questions. The selection of techniques to analyse the data is a complex process that is dependent on the investigators' level of experience of the techniques in question while also ensuring that they will suit the dataset(s) chosen. For example, in our pilot study, the selection of sentiment analysis techniques for a newcomer to a developing field can be fraught with difficulties as different papers suggest different techniques to use and most do not provide a concrete path to understanding the basics before choosing what path to follow. Social media analysis is a developing area and at present one does wonder if the techniques available are effective enough for any given specific domain, whereas in other fields techniques may well have been tried and tested over many years. In our experience within the pilot study, this led to it taking a considerable length of time to make a decision, which is why it's appropriate for this consideration to have a stage of its own. Another consideration to make at this stage is whether the researcher has the appropriate equipment to process Big Data and explore the intricacies of the dataset chosen using the desired tools. For example, initially within our research, using the R language presented some issues when processing a large amount of data, as R Studio is single threaded. This meant the PC being used was inadequate and required an upgrade due to poor single threading performance. An assessment must be made early on as to whether the PC or Cloud selection has the processing power to analyse the data in a reasonable amount of time (or indeed at all if there are memory considerations).

6) Stage 6 was originally entitled "Implementation/ Publication/Action" and has been renamed to "Implementation" here. In [1], it is originally emphasised that social media research is in its infant stages and that the likelihood is that the work being carried out will be exploratory. Any successful *"outcome or otherwise should be communicated"* [1, p10] to the interested communities to build on this in future work, which is the same in business as in research. To assist in these steps the researcher can include the good practice from the SMDS approach on "publishing" to "reuse/sharing" and "preservation" [14]. Publication is one of the steps in this section as dissemination of research is clearly vital. The GSR lifecycle emphasises successful outcomes, but as this is now named "Implementation", there is a new focus, more appropriate for research, on making sure the project requirements and specifications as previously outlined above are implemented in practice so as to achieve the aims of the project. For example, in this step we extracted the data with NVivo NCapture, cleaned them and analysed them to detect the sentiment within each Tweet and identify significant changes of sentiment within the timeframe over which the data were collected by using R. It was appropriate that this all took place within this phase, as one step flowed to the next with purpose and direction to contribute to the aim of the project. In addition, to this, ethical consideration must be given further thought at this phase to how any data are shared and preserved, but this data management process will not be discussed in this paper, as we shall

 ISBN: 978-1-61208-673-6

focus on the legal and ethical considerations of social media data usage, which will look in particular at publication dilemmas. Publication is included in the last phase of the lifecycle instead as we must implement and (in particular) evaluate *before* we can publish within the research context. In our own context, had we attempted to include publication here alongside analysis, this stage would have become confused by the lack of evaluation. Furthermore, given the paucity of the quality of social media data, we required additional focus on relevant cleaning of the data and attempting to consider publishing at the same time would have resulted in a loss of momentum.

7) Stage 7 (Evaluation) is included in the lifecycle due to the immaturity of social media research compared with other more established research fields. In [1], there is a focus on the evaluation of exploring what value there is in social media research compared to traditional methods. It suggests that this stage will confirm whether not social media was specifically required "to respond to a business or citizen need" [1, p10]. This stage will remain the same as outlined in GSR's lifecycle but with a rather different focus. Where the GSR strategy considers whether or not there was value in the use of social media data, the researcher's focus will be on how effective the use of such data was in addressing the research aims and questions. A stage devoted to evaluation is important, as through evaluation we can identify whether our techniques have been effective in answering any research questions. For example, in our case, we aim to consider whether using a lexicon dictionary approach over machine learning for detecting sentiment provides a greater level of accuracy within the framework we have set. We have not yet completed this section of the lifecycle for our own work on social (dis-)order, but this stage of the pilot study has shown us which techniques are less effective (e.g. Latent Dirichlet Allocation) for this specific study and allowed us to apply a greater focus on others (e.g. Changepoint identification).

8) Stage 8 has been renamed from "Business as Usual" as it is in the GSR lifecycle [1] to "Knowledge Management" in order to fit the research context. The original purpose of this phase remains, but with the addition of publication to emphasise its importance in this context. This phase re-evaluates research techniques in order keep research up-to-date with any modern research techniques and to think how about how any knowledge gained about social media research methods themselves can be transferred to others to instil good practice. This stage can be commenced once a significant part of the cycle is completed. Publications are crucial way of sharing good practice within the research community and can then lead to subsequent further research after interactions with the community, leading us back to stage 1 to begin a new project and frame suitable new research questions. The pilot study's outcome has informed us that this original lifecycle with a series of changes can be placed into a research context that is effective in guiding social media projects. These findings will be shared in the form of publications and with other researchers through other means of communication such as conferences.

It is important to note the lifecycle is not only to be used as a single iteration. A researcher can go repeat stages to develop the project through one or many iterations. Furthermore, this lifecycle itself will be further evaluated when cycling through it again within the rest of our research project. Having outlined a possible lifecycle for social media research, in the next section, we discuss the ethical and legal considerations that must be made throughout the social media research lifecycle.

III. ETHICAL AND LEGAL CONSIDERATIONS

Technological advancements are outpacing developments in research governance and what is agreed as good practice. The ethical code of conduct that we rely on for guidance for collection, analysis and representation of data in this digital era is not up-to-date [16][17]. Social media is ethically challenging because of its openness in relation to the availability of data. The Terms and Conditions of these platforms (including Twitter, Facebook, YouTube, Weibo, Qzone, Reddit, LinkedIn and other global social media platforms) state that users' data is available for third parties, so in accepting these, users are giving legal consent for their data to be made available [18]. As [19] outlines *"Just because it is accessible doesn't mean using is ethical"*, which means that researchers must evaluate their positions carefully, as to whether using the data is or is not ethically sound.

Datasets with this scale of social interaction, speed of generation and level of access are unprecedented in the social sciences. This has led to many published papers that include complete tweets and/or usernames without informed consent [18]. This seems to have happened because of the openness of some social media platforms, thus leading to assumptions that these are 'public data' and that projects using such data therefore do not require the same level of scrutiny by an ethics panel as do studies using data collected by more standard methods, such as interview or questionnaires [18]. Some universities may have not caught up with the pace of technology and this is often reflected in their ethical policies and within their forms dealing with ethical considerations. Even where ethics panels have already scrutinised such data, they may still deem it to be 'public data' due to the lack of a suitable framework to evaluate the potential harm faced by those whose ostensibly public data is used in the research in question [17]. In some cases, ethical approval is not required per se, but it is suggested by a given university's policy that researchers consult resources, such as the Association of Internet Researchers, that can help to ensure that any social media data are used in an ethical fashion [17] [20].

Despite noting above that some ethical panels are not making much consideration about the ethical use of social media data, there is some evidence to suggest that a

 ISBN: 978-1-61208-673-6

number of universities are making strides towards updating their ethical guidelines with regards to social media data. As one such example, the University of Sheffield has a research ethics policy note that raises many important points that can be considered in other institutions [21]. This note indicates that research must have ethical approval *before* a dataset can be extracted. However, this may pose both a financial and a contemporaneity problem. If the researcher wants to use historical data that will in any case come at a cost then this will be the case with or without prior ethical authorisation. However, if the data cannot be extracted on-the-fly because ethical approval is taking time to obtain, then the institution's budget would have to be prepared to pay for those data in the long term. Furthermore, if the researcher is considering topics of current interest and wishes to amend their search criteria as data come in, it may not in fact even be *possible* to seek suitable a priori approval. Of course, planning in advance is well advised here, but there are times when one cannot predict the topics of research interest that will arise today, tomorrow or in many weeks' time, which makes it difficult to plan such requests in advance. This policy is thought provoking, as it makes the researcher think about the importance of ethics in the very early stages of their research and the requirement for ethical approval for social media research is clearly a step in the right direction towards ensuring high ethical standards. However, as noted above, it may be financial unviable, or prevent the collection of data required for some projects. To that end, we would recommend that perhaps there be a fast track ethical approval system for time-critical social media data projects so that on the one hand they receive suitable ethical scrutiny, while on the other they can also proceed in a timely manner, enabling researchers to react to current events of public interest.

According to a series of survey findings from [22] and [23], it appears that there is a disconnect between the practices of researchers in publishing content on social media posts and *"users' views of the fair use (includes accuracy) of their online communications in publications and their rights as research subjects."* [17]. The decision-making process in one's ethical approach to social media data must consider the expectations of social media users as regards their personal privacy. In addition to this, the researcher must review the nature of the information from a user on social media alongside its originally intended purpose.

Users on social media *"may not intend for their data to be used for their [researchers'] purposes"* [24] and have, therefore, not consented to it being used for research. Considerations must be given to possible risks to the users whose data are being employed in any research. We must recognise that social media research transcends the usual boundaries of geography and standard methodologies. This means that a scholar's research design must ensure that it satisfies the legal regulations and terms of service of each platform as well as those platforms' hosting countries' laws and the laws where the researchers are based. This also includes institutional guidelines, the privacy and expectations of users and their vulnerability from publications covering their activities, the reuse and publication of data and how users' contributions are anonymised [24]. The application of ethics must consider the concerns raised above. If researchers and organisations are not careful in their approach, the disconnect between researchers and users may grow further. A lack of action regarding such ethics could lead to a series of undesirable consequences, such as users calling on social media platforms for changes in their terms of service to restrict the use of their data. The impact of this may make it extremely difficult to use social media data for research designed for the public good.

Social media research ethics as specified above requires further development and awareness to ensure that the public's data are represented in their context in an accurate, respectful and fair way [10] [25]. Ethics of social media data analysis is of significant importance and is hotly debated in the research community (by organisations such as, the Social Research Association [26], the Academy of Social Sciences [27], and the New Social Media, New Social Science [28]) and outside of it, where improvements are continually being made to relevant ethical frameworks [10]. Ethics could be applied in the sense of one's own morality and standard of ethics, but the problem with this is that not everyone may have the same high ethical standards. Indeed, one may think that they have a high set of standards when in actuality their standards are lower than they believe and overall this is a slippery slope as it is open to suggestions of improper usage as there is no conformity to an agreed set of rules.

Current ethical guidelines are an ongoing area of development amongst research institutes and other organisations. There are a series of organisations that have produced a set of guidelines to follow, all of which support a high standard of ethical practice in social media research. Some examples of these organisations and efforts are provided below.

- A Canada Research Chair has emerged from a five-year partnership with SMDS. This project aims to address the concerns of incoherent and inadequate practice in social media research and suggests a set of guidelines for conducting large scale and aggregated analysis through social listening [14] on sensitive topics, such as medical and religious data [29].
- Ipsos MORI (funded by institutes such as the EPSRC, ESRC, CASM and DEMOS) is a market research organisation in the UK that is "curious about people, markets, brands and society", where they "deliver information and analyses by making it faster and easier to navigate our complex world and aid clients in making better decisions." IPSOS MORI produced a guide that examines and reviews the ethical, legal and regulatory framework for embedding ethics in social media research [30]. When considered alongside the SMDS framework, these provide a comprehensive set of user-driven principles to help manage all aspects of social media data in research, such as how to decide and handle the use of reproduced tweets - especially those that concern sensitive topics.

 ISBN: 978-1-61208-673-6

- The Economic and Social Research Council (ESRC) has an "ESRC Framework for research ethics", which contains a few social media guidelines [20] that can be put into practice. As social media ethics develops, we would suggest that the ESRC might wish to consider the addition of further guidance aimed towards helping social media researchers, particularly newcomers to the field, to navigate the uncertainty and confusion of this nascent field to help to ensure that they meet a high standard of ethics.
- The Government Social Research (GSR) team used a data science framework and incorporated a social media element into this directly. This report shows some pertinent core principles for the researcher that must be considered when conducting any social media research [1]. There are many important ethical considerations given, such as *"Core principle 4: Avoidance of personal and social harm"* [1, p20] and *"Core principle 5: Non-disclosure of identity"* [1, p20] which are straightforward and clear to understand.

The above ethical guidelines cover different areas, for example, the SDMS guidelines are focused on the actual conduct of social media analysis, the IPSOS MORI framework covers legal and regulatory issues, the ESRC guidelines are rather generic and do not yet constitute a concrete approach while the GSR team have simply appended to their current ethical framework a social media element, so that the framework is more specialised towards social media [25]. There are calls from [25] for institutions' ethics committees to integrate requirements into the approvals documentation (by specifying which ethical guidelines would be applied in one's research); as [25] suggests there is a low level of ethical awareness amongst researchers applying social media data mining in their studies.

Now, all these guidelines provide very important points, but their multiplicity creates difficulties for the researcher as there are still uncertainties around the ethics of social media research in part because these guidelines do not always agree. Of course, this area is still in its early stages of development. In addition, what makes this area even more difficult is the terms and conditions set by the individual social media platforms. These can be hard to interpret because of the legal terminology or may be otherwise ambiguous and different platforms have different terms of service, such that it can be difficult for multi-platform research to adhere to them all simultaneously. Moreover, the terms and conditions can create ethical concerns for publication - for example, Twitter will not allow tweets to be presented without usernames [10], which can make it difficult to protect participants from potential harm. If the data are highly sensitive and the username is published, then the effect of linking the user to these data and the research may cause an effect within the public sphere. For example, the subject may receive positive responses, thereby boosting their reputation, or, perhaps more seriously, highlighting negative tweets may damage the mental or physical wellbeing of those mentioned within them.

Online research poses a greater risk to upholding confidentiality than does protecting offline research [18]. One reason for this is that at present there is a permanent record of what has been posted online. For instance, any quotation used can lead directly back that user in question with the use of a search engine [18]. This raises concerns over the anonymity of data. For example, as noted above, Twitter's data sharing licensing policy allows the sharing of Tweet IDs only, to ensure the data collection process is reproducible [18]. Using the identification ID provides a way to obtain the same dataset from Twitter's API. These IDs are unique and are easily searchable on the web to locate each tweet. This can be a cause for concern as it makes it easier to de-anonymise the data, so if the data are highly sensitive then a choice has to made as to whether that ID should be excluded from being shared if it causes an ethical concern [18]. Furthermore, the anonymisation techniques we can apply now may become easier to deanonymise in the future due to technological advancement.

In the UK, we must also take other laws into consideration, such as the Data Protection Act 1998, as researchers need to comply fully with the data protection principles laid out therein. Section 33 of the Data Protection Act 1998 allows exemptions to be made in accordance with principles 2 and 5 of the act for personal data used in research [31]. Recent developments within the UK Government suggest they are looking to form a council of data ethics *"to address the growing legal and ethical challenges associated with balancing privacy, anonymisation of data, security and public benefit."* [32] and also to implement the General Data Protection Regulation on the 25th of May 2018 [33]. Researchers will have to take these developments into account in their future practice as it may impact their social media research. Even after Brexit, the General Data Protection Regulation (which includes similarities with the existing UK Data Protection Act 1998) will be adopted into UK law [33]. It is essential that researchers keep abreast of any legal developments and keep up-to-date with good practise in their relevant area so as to make the best possible ethical use of social media data.

The concerns outlined above regarding the ethical challenges of using social media data can make for a difficult challenge for the social media researcher. The best course of action the researcher community can take is to address concerns and difficulties on case-by-case basis, thereafter trying to update guidelines and frameworks to deal with such cases. Genuine mistakes might have been made in the research community, which both individual researchers and the community as a whole can learn from. If a researcher has made a genuine ethics-related mistake in their work and has demonstrated remorse, then we as a community need to forgive and look to further strengthen the ethical standards and frameworks available to us. Indeed, ethical concepts are not just hoops to jump through in the early phases of research, but concepts requiring ethical inquiry [18], which may in itself take time. Mistakes may not be recognised until well after they have occurred and numerous judgements are possible, which can provide uncertainty and ambiguity, but this is likely to apply to any research [18]. Ethical

 ISBN: 978-1-61208-673-6

considerations will be in a constant state of assessment throughout any project and each case that arises during the research process can be worked through using a set of context-specific decisions. In addition to this, researchers must be guided by core ethical principles set by their employing organisations and external bodies, while also employing an appropriate mixture of the frameworks as laid out above, to ensure that the highest ethical standards are followed in any research.

There is a need to improve ethical assessment and one way to do this is to create a value-based ethical culture and practices in the research community and within other organisations for the development and deployment of intelligent systems both within the UK and elsewhere. This is known as Value Based Design (VBD) [34]. To do this, one must identify, enhance and ultimately embrace management strategies and social processes that facilitate value-based ethics within their design process. This could be included as an additional step in the framework in a future development, as it may provide a way to ensure a higher standard of ethical practice in the future.

IV. CONCLUSION

This paper has taken an existing methodology, the GSR lifecycle, and created from it a new social media lifecycle suitable for the research context. This was illustrated via a new diagram (Figure 1) that contains steps adapted to incorporate changes that are required for use in a research context. Alongside this, a number of ethical concerns have been explored and we have highlighted a series of pertinent points to consider in any future social media research project. Overall, this paper has sought to provide an easier way for researchers to enter the domain of social media research and then conduct relevant research, while providing an insight into the importance of the relevant ethical considerations in this area. Future research directions could include widening the framework beyond the UK, to other domains such as the wider European Union, the United States and Canada in a more detailed fashion and further thought could be given to how to expand the framework to include VBD.

REFERENCES

[1] "Using social media for social research: An introduction", *Assets.publishing.service.gov.uk*, 2016. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524750/GSR_Social_Media_Research_Guidance_-_Using_social_media_for_social_research.pdf. [Accessed: 09- Aug- 2018].

[2] J. Crump, "What Are the Police Doing on Twitter? Social Media, the Police and the Public", *Policy & Internet*, vol. 3, no. 4, pp. 1-27, 2011.

[3] M. Downes, "UK Police Forces using Social Media: Twitter, facebook, YouTube, Google+ and Hangouts On Air - The January 2013 Survey by Mike Downes", *Whatsinkenilworth.com*, 2013. [Online]. Available: http://www.whatsinkenilworth.com/2013/01/uk-police-forces-using-social-media.html. [Accessed: 08- Aug- 2018].

[4] "Internet access – households and individuals, Great Britain - Office for National Statistics", *Ons.gov.uk*, 2016. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2016. [Accessed: 08- Aug- 2018].

[5] "Government Social Research Service | Civil Service Fast Stream", *Faststream.gov.uk*, 2018. [Online]. Available: https://www.faststream.gov.uk/government-social-research-service. [Accessed: 22- Aug- 2018].

[6] "Global market and opinion research specialist | Ipsos MORI", *Ipsos.com*, 2018. [Online]. Available: https://www.ipsos.com/ipsos-mori/en-uk. [Accessed: 22- Aug- 2018].

[7] "Big Boulder Initiative, Big Boulder Conference Big Boulder Initiative", *Bbi.org*, 2018. [Online]. Available: http://www.bbi.org/. [Accessed: 08- Aug- 2018].

[8] "Demos", *Demos.co.uk*, 2018. [Online]. Available: https://www.demos.co.uk/. [Accessed: 22- Aug- 2018].

[9] "About us", *European Citizen Science Association (ECSA)*, 2018. [Online]. Available: https://ecsa.citizen-science.net/about-us. [Accessed: 08- Aug- 2018].

[10] A. Quan-Haase and L. Sloan, *The SAGE handbook of social media research methods*. Los Angeles: Sage Publications, pp 1-9, 2017.

[11] D. Ruths and J. Pfeffer, "Social media for large studies of behavior," *Science*, vol. 346, no. 6213, pp. 1063–1064, 2014.

[12] R. Procter, F. Vis and A. Voss, "Reading the riots on Twitter: methodological innovation for the analysis of big data", *International Journal of Social Research Methodology*, vol. 16, no. 3, pp. 197-214, 2013.

[13] Y. Theocharis, W. Lowe, J. van Deth and G. García-Albacete, "Using Twitter to mobilize protest action: online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements", *Information, Communication & Society*, vol. 18, no. 2, pp. 202-220, 2014.

[14] "About – Social Media Data Stewardship", *Socialmediadata.org*, 2017. [Online]. Available: http://socialmediadata.org/about/. [Accessed: 09- Aug- 2018].

[15] "Analysis of XX Commonwealth Games Host Broadcast Coverage, Online Media and Official Digital Channels", *Gov.scot*, 2015. [Online]. Available: https://www.gov.scot/Resource/0048/00482015.pdf. [Accessed: 09- Aug- 2018].

[16] M. Williams, P. Burnap and L. Sloan, "Crime Sensing with Big Data: The Affordances and Limitations of using Open Source Communications to Estimate Crime Patterns", *British Journal of Criminology*, vol. 57, no. 2, pp. Pages 320–340, 2016.

[17] "Ethics Resources – Social Data Science Lab", *Socialdatalab.net*, 2018. [Online]. Available: http://socialdatalab.net/ethics-resources. [Accessed: 08- Aug- 2018].

[18] K. Beninger. Social Media Users' Views on the Ethics of Social Media Research. In: A. Quan-Haase and L. Sloan. (eds.) *The SAGE handbook of social media research methods*. Los Angeles: Sage Publications, p.53–73, 2017.

[19] D. Boyd, "Privacy and Publicity in the Context of Big Data", *Danah.org*, 2010. [Online]. Available: http://www.danah.org/papers/talks/2010/WWW2010.html. [Accessed: 08- Aug- 2018].

[20] "ESRC Framework for research ethics Updated January 2015", *Esrc.ac.uk*, 2015. [Online]. Available: http://www.esrc.ac.uk/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2015. [Accessed: 08- Aug- 2018].

[21] "The University of Sheffield Research Ethics Policy Note no. 14, Research Involving Social Media Data", *Sheffield.ac.uk*, 2017. [Online]. Available: https://www.sheffield.ac.uk/polopoly_fs/1.670954!/file/Research-Ethics-Policy-Note-14.pdf. [Accessed: 08- Aug- 2018].

[22] M. Williams, "Towards an ethical framework for using social media data in social research", *Socialdatalab.net*, 2015. [Online]. Available: http://socialdatalab.net/wp-content/uploads/2016/08/EthicsSM-SRA-Workshop.pdf. [Accessed: 07- Aug- 2018].

[23] K. Beninger, A. Fry, N. Jago, H. Lepps, L. Nass and H. Silvester, "NatCen Social Research, Research using Social Media; Users", *Natcen.ac.uk*, 2014. [Online]. Available: http://www.natcen.ac.uk/media/282288/p0639-research-using-social-media-report-final-190214.pdf. [Accessed: 07- Aug- 2018].

[24] A. Bruns, "Challenges in Social Media Research Ethics | Snurblog", *Snurb.info*, 2017. [Online]. Available: http://snurb.info/node/2227. [Accessed: 07- Aug- 2018].

[25] J. Taylor and C. Pagliari, "Mining social media data: How are research sponsors and researchers addressing the ethical challenges?", *Research Ethics*, vol. 14, no. 2, pp. 1-39, 2017.

 ISBN: 978-1-61208-673-6

[26] "The SRA | Home of the Social Research Community", *The-sra.org.uk*, 2018. [Online]. Available: http://the-sra.org.uk/. [Accessed: 03- Oct- 2018].
[27] "Academy of Social Sciences", *Acss.org.uk*, 2018. [Online]. Available: https://www.acss.org.uk/. [Accessed: 03- Oct- 2018].
[28] New Social Media New Social Science, "#NSMNSS", *Nsmnss.blogspot.com*, 2018. [Online]. Available: http://nsmnss.blogspot.com/. [Accessed: 03- Oct- 2018].
[29] E. Yom-Tov, D. Borsa, I. Cox and R. McKendry, "Detecting Disease Outbreaks in Mass Gatherings Using Internet Data", *Journal of Medical Internet Research*, vol. 16, no. 6, p. e154, 2014.
[30] H. Evans, S. Ginnis and J. Bartlett, *Ipsos.com*, 2015. [Online]. Available: https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Publications/im-demos-social-ethics-in-social-media-research-summary.pdf. [Accessed: 07- Aug- 2018].
[31] W. Ahmed, "Ethical Challenges of Using Social Media Data In Research", *YouTube*, 2017. [Online]. Available: https://www.youtube.com/watch?v=VeFMqL4Hj60. [Accessed: 07- Aug- 2018].
[32] "Government agree to set up 'Council of Data Ethics", *UK Parliament*, 2016. [Online]. Available: https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/news-parliament-2015/big-data-dilemma-government-response-15-16/. [Accessed: 07- Aug- 2018].
[33] "Overview of the General Data Protection Regulation (GDPR)", *Ico.org.uk*, 2017. [Online]. Available: https://ico.org.uk/media/for-organisations/data-protection-reform/overview-of-the-gdpr-1-13.pdf. [Accessed: 05- Aug- 2018].
[34] "Methodologies to Guide Ethical Research and Design", *Standards.ieee.org*, 2018. [Online]. Available: https://standards.ieee.org/develop/indconn/ec/ead_methodologies_research_v2.pdf. [Accessed: 05- Aug- 2018].

 ISBN: 978-1-61208-673-6

The Perceived Psychological Empowerment of Women Using Mobile Dating Applications: The Case of Tinder

Mandlakazi Ndlela, Maureen Tanner
Department of Information Systems
University of Cape Town
Cape Town, South Africa
email: ndlman015@myuct.ac.za , mc.tanner@uct.ac.za

***Abstract*— Modern romance and dating are currently experiencing a 'digital revolution' that is powered by online dating platforms. Different forms of social media and social networks, such as dating apps, are gradually tapping into the Internet's acclaimed ability to empower users. Social media has been described a tool which plays a significant role in enabling women to gain control of and improve the quality of their lives. Literature suggests that online dating is of particular importance to women who seek to gain more control of their dating lives. While online dating apps and literature may cite women empowerment as the overall goal of dating apps such as Tinder, there is a need to empirically investigate the extent to which the use of dating apps empower women. The purpose of this study was therefore, to explore how women's use of mobile dating apps promote or hinder their psychological empowerment. This study applied a qualitative, single case-study strategy of the Tinder dating app. A deductive approach was taken, as the study relied upon a conceptual model to illustrate Tinder's affordances and features alongside the components of psychological empowerment. The research findings confirm that the use of mobile dating app affordances and features, to some extent, promote as well as hinder particular psychological empowerment components. The identified findings could further contribute to the understanding of how the use of specific online dating app affordances and features influence the psychological empowerment of women. An enhanced understanding of how women use these affordances and features, may encourage mobile dating app developers to rethink the design of their apps, in the context of women and their psychological empowerment.**

Keywords- Online dating; Mobile dating apps; Tinder, Women; Psychological Empowerment; Affordances; Features

I. INTRODUCTION

Modern romance and dating are currently experiencing a 'digital revolution' [1]. This revolution is powered by online dating sites and applications (apps). Online dating platforms have claimed the role of conventional matchmakers, which used to include families, friends, and dating agencies [1].

Online dating platforms, such as Tinder, are often referred to as social media applications or social networking sites [2] [3]. Stavrositu et al. [4] claim that social media and social networks are gradually tapping into the Internet's claimed ability to empower its users. According to Hamid et al. [5], social media plays a significant role in specifically enabling women to gain control, and improve the quality, of their lives, which is consistent with the definition of Psychological Empowerment [PE] provided by Tahir et al. [6].

Literature suggests that online dating is particularly interesting to women who wish to gain more control of their dating lives, as it presents an opportunity for them to partake in activities that cannot be achieved by traditional means of dating [7]. Similar sentiments regarding a woman's role in her dating life are shared by the creators of Mobile Dating Apps (MDAs), such as Tinder. Tinder has stressed that one of its overall goals is to empower women [3]. While online dating apps and literature may cite women empowerment as the platform's overall goal, there is a need to empirically investigate the extent to which the use of dating apps empower women.

It is important to specifically investigate the empowerment of women, as they have been subject to a greater level of disempowerment when compared to their male counterparts [8]. Although development initiatives are still in the process of bridging this gap, scholars have stressed that third parties are unable to directly empower, but they can attempt to create empowering conditions [9] [10]. While significant attention has been given to women's economic empowerment, there is a need to address their PE, since the construct has been disregarded in the past [8].

To better understand how women's use of dating apps influenced their perceived PE, this study applied a case study research strategy of the Tinder MDA and collected data from women who participated in semi-structured interviews. The data analysis process was guided by the proposed conceptual model, along with a thematic analysis approach [11]. It is hoped that this research will provide future scholars, as well as dating app designers, insight into how women may use MDA affordances and features to influence their perceived PE.

In Section 2, this research paper will discuss a literature review on MDA affordances, features, and PE. Section 3 will describe the methodology used. The findings will be presented and discussed in Section 4, followed by a conclusion in Section 5.

 ISBN: 978-1-61208-673-6

II. LITERAURE REVIEW

The following section will discuss the literature reviewed on MDA affordances, features, and PE. The reviewed affordances include mobility, immediacy, proximity, visual dominance, and multimediality. The features are discussed in terms of swiping and un-matching, followed by an explanation of the intrapersonal, interactional, and behavioural components of PE.

A. Affordances of Mobile Dating Applications

Schrock [12] emphasises that an individual has agency in deciding how a tool should be used. In this study, agency refers to a woman's capacity to make personal choices [47]. It is therefore, important to note that this study acknowledges the role of one's agency by asking the primary research question: How does women's use of mobile dating apps influence their perceived psychological empowerment?

According to Willemse et al. [13] affordances are "the perceived and actual properties of an object, primarily the functional properties that determine just what and how the object could possibly be used" (p. 2). While affordances are not always visible [14], they structure the way in which one can interact with an object by either allowing or prohibiting various actions [15].

In discussing the affordances of MDAs, it is necessary to first discuss the role of mobile media in dating apps. Mobile media fall under a category of mobile technologies, such as smartphones and tablets, which are used to run mobile applications [12]. Users therefore, interact with MDAs by using mobile media [16].

Mobile media consists of four generic communicative affordances: portability, availability, locatability, and multimediality [12]. MDAs depend on all four communicative affordances [16], but also consist of affordances which are specific to dating apps. This study will consider the following MDA affordances: mobility, immediacy, proximity, visual dominance, [15] and multimediality [17]. Although literature has not identified multimediality as a specific MDA affordance, the Tinder app has made changes to its design, which now allows for multimediality [17]. After noticing that users were manually adding their Instagram page links to their accounts, Tinder decided to incorporate the feature into their design [17]. As this study focuses on the Tinder app, it will include the multimediality affordance.

1) Mobility: Tinder's mobility affordance is consistent with the generic 'portability' communicative affordance of mobile media [16]. Mobility encourages users to interact with MDA in various locations [15]. The affordance implies that having Tinder on a smartphone or tablet gives one the freedom to use it wherever one desires [18]. There are no restrictions as to where the app can be used, which allows for interaction in private, public, and semi-public spaces [16].

2) Immediacy: MDAs aim to provide immediate social interactions [19]. Tinder enables interactions between users to take place quicker, by alerting them to new matches (indicating mutual interest between two users) or messages [15].

3) Proximity: Proximity is associated with the 'locatability' communicative affordance. Both proximity and locatability are provided by Tinder's location-based services [20], as users are required to indicate the preferred distance between them and potential matches [3].

4) Visual dominance: Visual dominance refers to how Tinder encourages one to review users based on their profile photos which tend to take up a large proportion of a phone's screen [18].

5) Multimediality: The multimediality affordance allows users to connect their other social media profiles, such as Instagram, to their Tinder accounts [16]. The linking of different social media profiles allows women to further present themselves to other users [16].

B. Features of Mobile Dating Applications

For this study, the swiping and un-matching features were considered. For a match to be made, users must both swipe right on each other's profiles [17]. Users may also use the un-match feature which allows them to remove previously matched people from their list. The person who has been unmatched will not be able to contact the user that initiated the action [21].

C. Psychological Empowerment

Psychological empowerment has been described as, "a mechanism by which people gain mastery of their affairs" (p. 144) [22]. It is important to note that PE is not static, as it changes over time [10]. There are several studies, which have conceptualised PE in different manners, however, they mainly rely on the Zimmerman [10] PE conceptual model [23]. Zimmerman [10] warns that in order to holistically investigate the PE of an individual, the intrapersonal, interactional, and behavioural components must all be considered. It is also stressed that when applying the PE nomological framework, studies should only consider the measures which are context appropriate [10].

1) Intrapersonal PE

The intrapersonal component refers to an individual's beliefs about their ability to have an impact on their environment and to achieve their personal goals [11]. Zimmerman [10] lists motivation to control, perceived control, self-efficacy, and mastery as sub-components of intrapersonal empowerment. Table I defines each intrapersonal PE factor in the context of this study.

2) Interactional PE

The interactional component emphasises an individual's choice to act in a manner that will help them achieve their goals. In order to have control in a particular setting, one must be aware of one's options [10]. The interactional component also includes an individual's intellectual comprehension of their social environment and whether they possess the necessary resources and knowledge to effect change [10]. These factors include critical awareness,

understanding causal agents, resource mobilisation, decision-making skills, and problem-solving abilities [10]. Table II defines each interactional PE factor in the context of this study.

TABLE I. INTRAPERSONAL PE DEFINITIONS

PE Factor	Definition
Motivation to control	Motivation to control is the extent to which a woman using an online dating app wants to be in control of her self-presentation [17] [24].
Perceived control	A woman's belief about the level of control she has over her self-presentation when participating in online dating [17].
Self-efficacy	Self-efficacy is a woman's belief about her ability to reach her online dating goals [25]. One's self-efficacy influences the extent to which one will take part in certain behaviours [25].
Mastery	Mastery relates to how competent one finds oneself to be [6]. A woman's self-efficacy will determine her mastery of online dating behaviours [25].

TABLE II. INTERACTIONAL PE DEFINITIONS

PE Factor	Definition
Critical awareness	The degree to which a woman is aware that a dating app may be used as a resource to create close relationships [26].
Understanding causal agents	Causal agents are authoritative figures that can exert control over situations [27]. It is important to understand these agents, their role and involvement in the relevant issue and the factors which may affect their decision-making [27]. Acknowledging these factors is important as they can hinder or promote one's attempt to gain control in one's environment [10]. This study will consider the users of dating apps as causal agents.
Resource mobilisation	Resource mobilisation refers to being able to manage an acquired resource to achieve a goal [10]. In this study, resource mobilisation will refer to a woman's ability to manage a dating app.
Decision-making	A woman's ability to decide whether she wants to interact with another dating app user [28].
Problem-solving	Problem-solving refers to how women deal with problems related to online dating [20].

3) Behavioural PE

The behavioural component of PE is described as the actions which are necessary for a desired result [29]. The interactional component of PE prepares individuals to carry out behaviours which are necessary for achieving their desired goals [10] Table III defines three behavioural PE factors: community involvement, participation, and coping behaviours.

III. RESEARCH METHODOLOGY

The following section describes the research methodology that was followed for this study. The research model is presented first, followed by a discussion on the nature of this study and its research questions. The methodology then describes the case, followed by a discussion regarding sampling, data collection and analysis, ethical considerations, validity and reliability.

TABLE III. BEAHVIOURAL PE DEFINITIONS

PE Factor	Definition
Community involvement	The degree to which one can openly participate in interactions with other users, beyond the one-on-one style usually allowed on dating apps [30].
Participation	The extent to which a woman engages in dating app activities [31].
Coping	Coping describes the, "behaviours and thoughts which are consciously used by an individual to handle or control the outcome of dealing with responsibilities or problems successfully or in a composed manner" (p.6) [11]. This study will consider the coping behaviours of women using dating apps.

A. Research Model

The conceptual model which was used for this study is presented in Figure 1. The model has been adapted from Osman et al. [11] by replacing telecentre components with MDA affordances and features. The affordances and features appear alongside each PE component.

B. Nature of study

The nature of this study was exploratory, interpretive, qualitative, deductive, and applied a case study research strategy. Osman et al. [11] suggest that a case study strategy allows for a modern phenomenon to be empirically explored and thoroughly understood in its real-life context. A single case study approach, therefore, assisted in answering the primary research question, 'How does women's use of mobile dating apps influence their perceived PE?'

C. Research Questions

The following Secondary Research Questions (SRQ) aided in answering the primary question:

- How are women using mobile dating apps? (SRQ1)
- How does women's use of mobile dating app affordances and features promote their perceived psychological empowerment? (SRQ2)
- How does women's use of mobile dating app affordances and features hinder their perceived psychological empowerment? (SRQ3)

D. Case Study: Tinder

Tinder is an MDA which was released in October 2013 [19]. The app aims to connect people who do not know each other but live near each other [20]. The app is often referred to as the female counterpart of the all-male user app 'Grindr' which connects gay and bisexual men to each other [3]. Unlike 'Grindr', Tinder is not restricted to users of specific sexual identities and has accumulated a growing number of heterosexual females since its launch [20]. The app asks a user to select their gender, age, sexual preferences and the distance of other user's profiles they would like to view [3]. Once signed in, users can indicate their interest by swiping right or their disinterest by swiping left. Only after a match is made can a conversation be initiated [1].

 ISBN: 978-1-61208-673-6

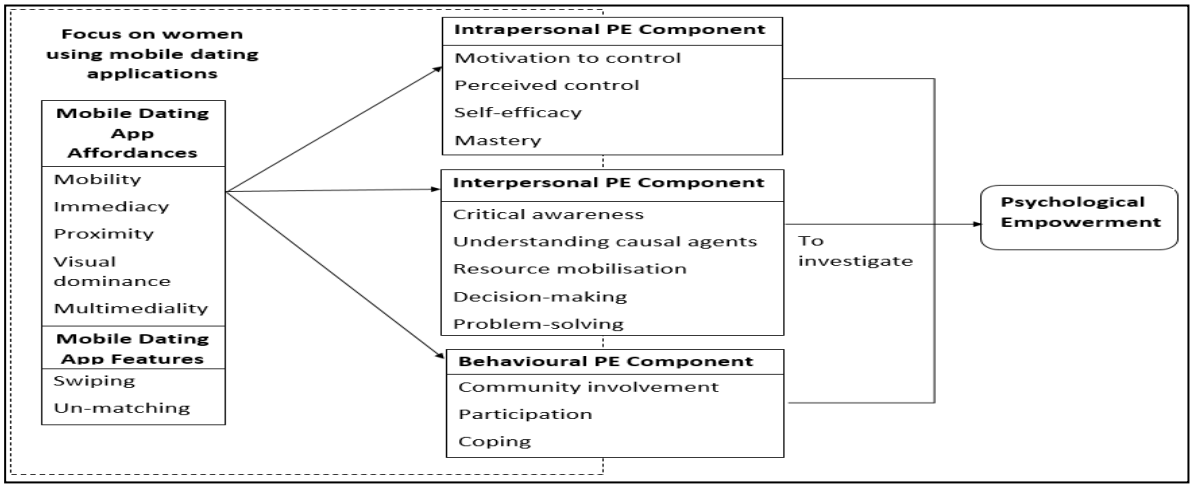

Figure 1. The PE of women using mobile dating apps, adapted from Osman and Tanner [11]

E. Sampling, Data collection and analysis

The sample consisted of ten women. While there is no widely accepted method for choosing an appropriate number of interviews to be conducted, it is suggested that a sample size of seven to ten participants is sufficient for qualitative studies [45]. A purposive and convenience sampling strategy was used. Women who were over the age of 18, living in South-Africa, and had ever registered as Tinder users, served as the unit of analysis and participated in audio-recorded, semi-structured interviews (Appendix A). To make sense of the transcribed data, a thematic analysis approach was used. To generate initial codes, the researcher made use of the conceptual model of PE for women who use mobile dating apps (Figure. 1), while thematic analysis was used to search, define, and name the crucial themes [11]. Saturation was reached at 10 interviews. At this point, there were no new themes that emerged from the data collected.

F. Ethical considerations,Validity and Relaibility

This research was conducted in an ethical manner, as all of the participants were informed about the purpose of the research and were asked to sign a consent form. Validity and reliability is of utmost importance in qualitative research [40]. Validity refers to the extent to which a study accurately reports on the views expressed by its participants [41]. Leung [42] describes reliability as the consistency in a study's procedures and its findings. While Noble et al. [40] argue that there is no unanimously agreed upon criteria to evaluate the validity and reliability of qualitative studies, they suggest several strategies to ensure the credibility of findings. To enhance the credibility of findings, this study adopted verification strategies proposed by Noble et al. [40] as well as Morse et al. [43]. Table IV illustrates each verification strategy along with a description of how it was applied to this study.

TABLE IV. VALIDITY & RELIABILITY VERIFICATION STRATEGIES

Verification Strategy	Application
Methodological coherence [43]	Methodological coherence is the extent to which the research question is consistent with the different elements of the research method used [43]. The primary research question, 'How does women's use of mobile dating apps influence their perceived PE?' was matched with an exploratory, interpretive, and qualitative research method. The method ensured that the question was answered according to the accounts given by the women interviewed.
Appropriate sample [43]	The sample used was appropriate for this study. Purposive sampling ensured that each of the participants were selected based on their ability to answer the research questions. The sample consisted of ten women who had ever had a Tinder account.
Reflexivity [40]	A reflective journal was kept throughout the duration of the study. The journal was mainly used to take note of interesting points made by participants. The decision to add to or amend the interview questions was partly based on the points highlighted in the journal.
Representation of the findings [40]	The audio recorded, semi-structured interviews used in this study allowed for the researcher to revisit the collected data to identify new themes. The flexibility of the interviews also ensured that participants were allowed the opportunity to provide an in-depth account of their Tinder experiences.
	To support the findings, 'rich' and 'thick' quotes were taken verbatim. This allowed the study to maintain the true nature of each participant's account.
Collecting and analysing data concurrently [43]	Similar to reflexivity, simultaneously collecting and analysing data allowed the researcher to identify issues that had been addressed and those which still needed to be further investigated.

I. FINDINGS AND DISCUSSION

The following section will present and discuss the findings of this study. Findings will be discussed in relation to each MDA affordance, feature, and various PE constructs.

A. Mobility

Findings reveal that women's use of mobility influences the intrapersonal, interactional, and behavioural components of their PE.

1) Intrapersonal and Interactional PE

While literature mainly discusses the 'motivation to control' component when referring to a user's desire to control their self-presentation [17] [24], findings revealed that women were motivated to control aspects beyond self-presentation. Due to Tinder being available on their mobile devices, five of the participants were motivated to control the privacy of their dating lives, as well as where and when they could engage in mobile dating. Participants expressed that the mobility of the app allowed for private use in public spaces, which would not be possible on a computer. Respondent 1 explained that she does not want people to know that she uses Tinder and that using the app on her phone, as opposed to a computer, allows her to keep her dating life private. She added that it would be embarrassing to interact with the app on a computer, as the larger screen would make it obvious to others that she participates in mobile dating.

"I feel like it's more privatised on my phone because…sometimes I don't want people to know that I'm using Tinder." (Respondent 1)

a) Perceived control

Findings revealed that women's need to control the privacy of their social interactions was satisfied by the control they felt they had gained. Respondent 7 felt that by using the app on her phone, she controlled the extent to which people saw her taking part in mobile dating. The findings which describe the motivation to control one's privacy, as well as the perceived control felt over privacy, are supported by literature. Studies show that there are no restrictions as to where dating apps can be used, which allows for interaction in private, public, and semi-public spaces [16]. Studies also suggest that there is a stigma attached to online dating, as its users may be perceived as being desperate [19].

b) Resource mobilisation (interactional PE)

Tinder's mobility affordance enhanced five of the participants' ability to use the app as a resource for meeting new people while travelling or moving to a new city. This finding is supported by Lean et al. [34], who explained that women use the Tinder app as a travel tool.

2) Behavioural PE

a) Participation

Findings revealed that being able to interact with the Tinder app on a mobile phone enhanced the extent to which the majority (9) of the respondents participated on the Tinder app. Timmerman et al. [15] support this finding by stating that Tinder's mobility increases how often users are able to use the app. Respondent 9 did, however, mention that her participation was negatively affected by only being able to access the app on her phone. Due to the considerable amount of time she spends on her laptop, she would prefer a desktop version of the app. Table V summarises the occurrence of each mentioned theme.

TABLE V. MOBILITY THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Intrapersonal	Motivation to control	Motivation to control privacy	5	1, 4, 7, 8, 10
	Perceived control	Control over privacy	5	1, 4, 7, 8, 10
Interactional	Resource mobilisation	Travel/ New city	5	1, 2, 3, 5, 8
Behavioural	Participation	Encouraged participation	9	1, 2, 3, 4, 5, 6, 7, 8, 10
		Limited participation	1	9

B. Immediacy

Findings reveal that women's use of the immediacy affordance influenced their resource mobilisation (interactional PE component). A new finding indicates that while dating apps claim to offer immediacy, four participants reported that the Tinder app did not deliver or notify them of any new messages from users. This hindered the extent to which some women could manage the app to interact with their matches. Table VI summarises the occurrence of each mentioned theme.

TABLE VI. IMMEDIACEY THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Interactional	Resource mobilisation	Faulty notification system	4	3, 4, 5, 9

C. Proximity

Emerging findings indicate that women's use of the proximity affordance influences the intrapersonal, interactional, and behavioural components of their perceived PE.

1) Intrapersonal PE

a) Motivation to control

Five of the participants reported that it was crucial for them to have control over their maximum distance settings, to ensure that they would only interact with users within a specific distance. The proximity affordance of the app enhanced the extent to which women wanted to have control over their maximum distance settings.

"… when I first got Tinder I had the location settings very narrow, because I was like, why do I want to swipe for people that I'm not ever going to be in contact with?" (Respondent 5)

b) Perceived control

When asked how they felt about the proximity affordance, four participants expressed that being able to adjust their maximum distance settings allowed them to have some control over the matching process and possibly meeting with nearby users.

"..you get people closer to you and it increases the probability of you meeting the person if they're closer to you." (Respondent 9)

An emerging finding suggests that a few women's feelings of control over the matching process and their safety were hindered. For example, three participants reported that although a preferred distance would be set, the app would return user profiles outside of the participant's location boundary. It is important to draw special attention to this finding, as Tinder stressed that the app was developed in a way which considered women's issues, such as 'proximity and control' [3].

2) Interactional and Behavioural PE

a) Critical awareness and resource mobilisation

By indicating the distance between women and other users, participants were made more aware of how Tinder could be used as a resource to locate potential matches. Tinder's proximity affordance enhanced women's ability to use the app as a resource for specifically meeting users who were nearby. Participants were able to eliminate users who were too far to meet up with.

"I love that because then you know who's close by and who isn't." (Respondent 6)

This finding is supported by literature which explains that the proximity affordance encourages users to match, exchange messages, and meet face-to-face with matches that are nearby [15] [16] [19]. The enhanced perceived control, critical awareness, and resource mobilisation also contributed to women's participation (behavioural PE) on Tinder. This finding is partially supported by Zimmerman [10], who stated that there is an association between perceived control and an individual's participation. Table VII summarises the occurrence of each mentioned theme.

D. Visual dominance

Findings reveal that visual dominance influences the interactional component of women's perceived PE

1) Interactional PE

a) Decision-making

Five women expressed that their decision to interact with a potential partner is influenced by that person's appearance. Tinder's visual dominance contributed to women's ability to review several users' looks and decide on who they wanted to interact with.

"So, I feel like that is like good about Tinder..I can choose how the people look that I am talking to." (Respondent 3)

Respondent 8, however, suggested that it was slightly difficult to decide on a potential partner based on their profile photo, instead of considering their personality. This finding is supported by Hess et al. [35] who criticise visual dominance (as well as the swiping feature) for forcing users to make hasty judgements based on images. Table VIII summarises the occurrence of each mentioned theme.

TABLE VII. PROXIMITY THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Intrapersonal	Motivation to control	Motivation to control maximum distance settings	5	1, 5, 6, 7, 9
	Perceived control	Enhanced control over matching process	4	1, 9, 6, 7
		Incorrect location boundary	3	4, 5, 9
		Less control over safety	2	8, 10
Interactional	Critical awareness	Awareness of potential matches	5	1, 5, 6, 7, 9
	Resource mobilisat-ion	Locating nearby potential matches	5	1, 5, 6, 7, 9
Behavioural	Participatio n	Encouraged participation	5	1, 5, 6, 7, 9

TABLE VIII. VISUAL DOMINANCE THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Interactional	Decision-making	Assists in choosing a potential match	5	3, 4, 5, 6, 9
		Difficulty in choosing a potential match	1	8

E. Multimediality

Findings reveal that women's use of the proximity affordance influences the intrapersonal and interactional components of their perceived PE.

1) Intrapersonal PE

a) Motivation to control

Four participants expressed that they needed to be in control of how much other Tinder users knew about them. Respondent 10 revealed that if she was not given the option to link her Tinder profile to other social media sites, such as Instagram and Facebook, she would not have signed up to use the app. Respondent 6 similarly expressed that by being presented with the option to apply multimediality, she was motivated to control how much other users knew about her. Three other participants also revealed that multimediality added to their motivation to control how they were presented to other users. Respondent 7 shared that she wanted to present herself as someone who has both masculine and feminine characteristics, and used the social media linking option to do so.

David et al. [17] support these findings by explaining that the "motivation to control how one is seen" (p. 4) is apparent in the manner in which users learn how to present a particular version of themselves on online dating apps. This presentation may include using pictures from social media.

b) Perceived control

For some participants, the motivation to control self-presentation was followed by the experience of perceived control. Respondent 9 reported that she was able to provide users with a "more rounded view" of who she is. This finding is supported by Ward [24], who claims that individuals that operate in online platforms experience more control over self-presentation.

Other participants felt that their level of control over privacy was limited, as signing up with a Facebook account or linking one's Instagram account created an opportunity for strangers to easily contact them on those platforms. Participants explained that while Tinder only allowed users to communicate if there was consent from both parties, linking one's Facebook or Instagram account to their Tinder profile encouraged other users to initiate unwanted contact.

"And then another guy invited me on Facebook and Tinder, ag (I mean) Instagram in like a day of me like just like swiping around in his profile. And I was just like, 'that is creepy'." (Respondent 3)

This finding is supported by a study conducted by Pond et al. [36], in which other women also described the unwanted contact via other social media sites as 'creepy'. This type of privacy is referred to as social privacy which includes receiving unwanted and inappropriate friend requests on social media platforms [16]. Emerging findings from the study reveal that while enhanced control through self-representation was an attractive outcome for some women, others were concerned about how multimediality compromises the control they have over their social privacy.

2) Interactional PE

a) Understanding causal agents

Mutimediality allowed half of the interviewed women to gain a better understanding of potential matches. This was due to being able to review the social media profiles of other users. Respondent 6 revealed that checking a match's other social media profiles gives her 'a more holistic' perspective of who they are.

"I've gone into someone's Instagram and gone as far as looking at Twitter for instance and see what they post and what they say." (Respondent 6)

This finding is confirmed by Pond et al. [36] who reported that it is a common practice for women using dating apps to visit a user's other social media accounts in an effort to "learn more about them" (p. 15).

b) Decision-making

Findings revealed that to some extent, multimediality contributed to a woman's ability to decide whether she would interact or meet face-to-face with another user. Four participants decided to interact with other users based on Tinder's mutual friend feature, which identifies common Facebook friends between two users. A study conducted by Timmermans et al. [15] supports this finding by explaining that Tinder users often use the 'mutual friends' indicator to choose which other users they want to interact with. Table IX summarises the occurrence of each mentioned theme.

TABLE IX. MOBILITY THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Intrapersonal	Motivation to control	Motivation to control privacy	4	3, 5, 6, 10
		Motivation to control self-presentation	3	1, 7, 9
	Perceived control	Limited control over social privacy	2	3, 8
		Control over self-presentation	3	1, 7, 9
Interactional	Understanding casual agents	Learning about other users	5	1, 2, 3, 6, 7
	Decision-making	Social Interaction choices	4	1, 4, 5, 7

F. Swiping

Findings show that women's use of the swiping feature affects the intrapersonal, interactional, and behavioural components of their perceived PE.

1) Intrapersonal PE

a) Perceived control

The swiping feature played a critical role in promoting women's perceived control. While existing literature has emphasised a dating app user's control in terms self-presentation [24], swiping allowed the majority (8) of women to feel a sense of control over other aspects of mobile dating. By allowing women to indicate their interest or disinterest, the use of the swiping feature improved their sense of control regarding who they matched up with.

"I think it's because you literally are in control, you choose who you like, so if you haven't liked anyone back there is no way they can match with you. So, you've sort of eliminated them." (Respondent 10)

This finding is supported by Timmermans et al. [15] who revealed that using the swiping feature enhances a user's sense of control. Participants also conveyed that the swiping feature allowed them to control who had the ability to directly send them messages. Respondent 7, who had previously used the online dating website OKCupid, explained that she preferred Tinder, as when she was last on OKCupid, she did not have control over which users could interact with her. Timmerman et al. [15] support this finding by stressing Tinder's need for mutual interest and consent in allowing one user to interact with another, unlike online dating websites which allow users to freely contact one another. In December of 2017, OKCupid made changes to its messaging system in an effort to reduce unwanted messages

from other users [44].

2) Interactional PE

a) Resource mobilisation

It is interesting to note that Tinder was used for more than just a resource for finding romantic partners. For example, respondents 3 and 7 used Tinder as a way to improve their self-confidence, by swiping and reviewing the list of users who were interested in them. Sumter et al. [37] support this finding by adding that both men and women used Tinder's swiping (and matching) features to validate their self-worth.

b) Decision-making

Respondents 5 and 7 felt empowered by the swiping capability as it encouraged them to be more decisive in terms of choosing which users to interact with.

"It allows us to be decisive and that's empowering, and it encourages us too...Here in this space you have no choice, you have to put either left or right there isn't a maybe." (Respondent 7)

Literature supports the role of swiping in decision making by explaining that the swiping gesture demands a "firm, decisive, micro-action" from its users (p. 7) [17]. It is, however, necessary to consider the extent to which the quick swiping action improves a woman's decision-making capability. Similar to visual dominance, a participant acknowledged that swiping limited her criteria to make decisions by requiring a quick action while emphasising the appearance of a user. Respondent 8, who also had issues with the app's visual dominance, felt that the fast-paced nature of the decision-making led her to place greater value on a user's looks, as opposed to considering their possible personality traits. This finding is supported by a study in which users felt that the decision-making process is fast paced and at times involuntary [17]. While online dating apps insist that users make a decision, it is important to question the quality of the decisions made. There is no indication of whether urging a user to decide to swipe left or right affects their ability to make a successful and effective choice [38].

3) Behavioural PE

a) Coping behaviours

When asked about how it feels when another Tinder user does not reciprocate a right swipe (interest), three respondents felt disappointed while four others exhibited a way of thinking which helped them cope with the feeling of being rejected. These four women were able to develop coping behaviours to deal with unfavourable circumstances.

"You're allowed to say no, you're allowed to have preferences, you're allowed to like who you like or not like who you like." (Respondent 6)

Due to Tinder only allowing users to view their match list, users are encouraged to focus on people who are interested in them, as opposed to those who have rejected them [39]. Table X summarises the occurrence of each mentioned theme.

TABLE X. SWIPING THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Intrapersonal	Perceived control	Control over matching	8	1, 2, 4, 5, 7, 8, 9, 10
		Control over social interactions (messaging)	2	6, 7
Interactional	Resource mobilisation	Improving self-confidence	2	3, 7
	Decision-making	Interaction choices	2	5,7
		Difficulty in choosing a potential match	1	8
Behavioural	Coping behaviours	Acceptance of being rejected	4	2, 6, 7, 8
		Disappointment in being rejected	3	3,10, 9

G. Un-matching

Findings indicate that women's use of the swiping feature affects the intrapersonal and behavioural components of their perceived PE.

1) Intrapersonal PE

a) Perceived control

When asked about being able to un-match someone on Tinder, most (9) participants expressed that they felt they had control over who they interacted with, while two others also emphasised control over their safety

"I think it is extremely important to have things like that...I have been in a situation where I am very grateful that you can block people off all of your social medial accounts Like, especially as women, you have to be aware of your safety." (Respondent 3)

When asked about how it feels to be un-matched, the majority (7) of participants reported that they were unaffected by the feature, as they did not notice that they had been unmatched, or they soon forgot about it. Two of the ten participants felt that they were no longer in control of their Tinder interactions. The two participants explained that the same platform which initially improved their sense of control, by allowing them to indicate interest in potential partners (swiping), deteriorated that control by allowing other users to un-match them without an explanation as to what they had done wrong. Respondent 4 did not speak to issues of control but felt that her level of confidence was negatively affected by being un-matched.

2) Interactive PE

a) Problem-solving

Un-matching provided participants with a way to solve problems related to social interactions on Tinder. Respondent 1 controlled the outcome of her conflicts by un-matching users who harassed her. She explained, *"so, the thing about Tinder is, right, if you harass me online and I unmatched*

 ISBN: 978-1-61208-673-6

you, I will never ever get to see you again". Harassment was a common theme mentioned by seven other participants. Unlike the majority of the participants, respondent 5 interestingly said that she preferred to ignore users who harassed her and only un-matched those who she had accidentally swiped right for. Respondent 7 did not feel the need to un-match people due to being harassed. She added, *"I think it's mostly because I interact with women. Women aren't like that. I haven't experienced a woman who's been vulgar, obsessing or…"*

b) Coping behaviours

Respondents 8 and 9 used the un-matching option to cope with men who they felt fetishised black women.

> *"Ja (yes), so if white guys are just like, mentioning things like you can be my chocolate queen…I can be your white man, ja. Ooh, I'm going to un-match with you, I cannot entertain that bulls**t."(Respondent 8)*

These findings are partially supported by Richey [21] who concluded that being able to use the un-matching feature allows one to manage the process of dealing with inappropriate users. Table summarizes the occurrence of each mentioned theme.

TABLE XI. UN-MATCHING THEMATIC ANALYSIS

PE Component	Theme	Sub-theme	Count	Resp.
Intrapersonal	Perceived-control	Control over social interactions	9	1, 2, 3, 4, 6, 7, 8, 9, 10
		Control over safety	2	3, 8
Interactional		Loss of control (being-unmatched)	2	8, 10
	Problem-solving	Unaffected (being-unmatched)	7	1, 2, 3, 5, 6, 7, 9
		Harassment	7	1, 2, 3, 4, 6, 8, 9
	Coping-behaviours	Accidental Swipe	1	5
		Fetishisation of Black women	2	8, 9

II. CONCLUSIONS & FUTURE WORK

The purpose of this study was to investigate the extent to which women's use of mobile dating apps influenced their perceived PE. The research strategy was a case study of the Tinder app. Data were collected from women who participated in audio-recorded, semi-structured interviews. To answer the primary research question, this study analysed how women's use of each affordance (mobility, immediacy, proximity, visual dominance, multimediality) and feature (swiping, un-matching) hindered or contributed to components of their perceived PE.

It is important to note that although this study presented its findings in a manner which identified the thematic count for each investigated mobility or feature, the goal of qualitative research is not to necessarily produce generalisable findings [46]. The aim is to thoroughly explore issues related to a phenomenon and to provide more understanding [46].

The first secondary research question for this study was: how are women using mobile dating apps? Findings revealed that women use MDAs not just for seeking romantic partners, but for making friends, meeting new people while travelling, and improving their self-confidence. Two additional secondary research questions asked: how does women's use of mobile dating app affordances and features promote (SRQ2) or hinder (SRQ3) their perceived PE? Findings suggest that women found the mobility and proximity affordances to be empowering, while the immediacy affordance hindered their sense of PE due to Tinder's faulty notification system. Findings also revealed that women's use of the majority of affordances and features resulted in an ambiguous effect on their PE. Affordances, such as visual dominance and multimediality, along with features, such as swiping and un-matching, were found to both promote and hinder women's PE. Women's sense of control can be highlighted as the most common theme throughout this study. Women expressed that using certain affordances and features influenced the control they had over their safety, privacy, social privacy, app usage, self-presentation, and interactions.

It must be emphasised that although women's use of online dating apps hindered or contributed to certain PE components, there was no indication of whether participants were able to achieve PE outside of the Tinder platform. Participants highlighted that certain actions, such as anonymously communicating one's interest or disinterest in a potential romantic partner (by using the swiping feature), or permanently disallowing someone from communicating with them (by using the un-matching feature), cannot be replicated outside of the mobile dating space. It can, however, be argued that there may be a few PE factors, such as coping-behaviours, which women may apply to their everyday lives. It is also important to note that most of the findings were only partially supported by literature, as there is a lack of studies which have specifically highlighted women's use of dating apps in relation to PE components. Existing studies have mainly focused on users (both men and women), as opposed to differentiating between the two.

The contribution of this research is twofold. First, it proposes a conceptual model which can be used by the wider research community to investigate the PE of women who use mobile dating apps. The model provides a way to holistically evaluate PE by taking into account the affordances and features of mobile dating apps, as well as the relevant PE constructs. Second, the findings of this study are addressed to those involved in mobile dating app development. This research provides app developers with more insight into the experiences of women using dating apps, while shedding light on critical issues, such as security and social privacy. Engaging with the findings of this study will encourage developers to address these issues by evaluating the design of their apps.

The exploratory nature of this study allowed for each participant to openly share their Tinder experiences. It is, however, necessary to conduct further studies in which

certain themes can be explored with more depth. Future studies should consider drawing special attention to the role of certain features, such as swiping and their role in PE, as there seemed to be a degree of ambiguity regarding the influence of swiping on women's decision making. A possible research question would be: How does the swiping feature on mobile dating apps influence a woman's decision-making? It would also be interesting to conduct a similar study which includes both men and women. This would allow for comparisons to be made between the two groups.

APPENDIX A

General Opening Questions

1. How long have you been using Tinder?
2. Why did you download Tinder?
3. Do you still use Tinder? (Why or why not)
4. In what ways does Tinder allow you to achieve the goals you just mentioned?
5. Is there anything challenging about using the app?
6. How did you overcome this challenge?
7. What does empowerment mean to you?
8. Why do you choose to use Tinder over traditional manners of meeting people?
9. Would you say you're a person who likes to have control in situations?
10. Does that extend to your dating life or the relationships you have with others?
11. When you use the Tinder app, do you feel that you have or don't have control over any parts of your life?
12. In what ways does Tinder give you more or less control over certain parts of your life?

Immediacy

13. How do you feel when you receive a match notification from the app?
14. What do you do once you receive the notification?

Proximity

15. What do you think about the location settings?
16. How do you feel about being able to pick a maximum distance?

Visual Dominance

17. How did you set up your profile?
18. How did you choose your pictures?
19. Do you ever feel commodified or objectified because of your profile?

Mobility

20. How do you feel about being able to use the app through your mobile device instead of having to use a computer?

Swiping

21. How do you feel about being able to swipe left or right on someone's profile?
22. How do you feel about people swiping your profile left or right? (Do you ever feel like you're a commodity?)

REFERENCES

[1] M. Hobbs, S. Owen, and L. Gerber, "Liquid love? Dating apps, sex, relationships and the digital," Journal of Sociology, vol. 53, pp. 271-284, Sept. 2016, doi:1440783316662718

[2] F. MacKee, "Social media in gay London: Tinder as an alternative to hook-up apps," Social Media+Society, vol.2, pp. 1–10, July-Sept. 2016, doi: 10.1177/2056305116662186.

[3] C. L. Mason, "Tinder and humanitarian hook-ups: the erotics of social media racism," Feminist Media Studies, vol. 16, 822-837, Feb. 2016, doi: 10.1080/14680777.2015.1137339.

[4] C. Stavrositu and S. Sundar, "Does blogging empower women? Exploring the role of agency and community,". Journal of computer-Mediated Communication, vol. 17, pp. 369-386, July 2012, doi:10.1111/j.1083-6101.2012.01587.x.

[5] N. A. Hamid, S. I. Mohd, and Y. Norhafezah, "Assessing validity and reliability of social media as an empowerment tool for a group at risk in Malaysia," Jurnal Komunikasi, Malaysian Journal of Communication, vol. 32, pp. 193-207, 2016, doi: 10.17576/JKMJC-2016-3201-09.

[6] M. A. Tahir and S. A. Rana, "Balochistan: an evaluation of the level of Psychological Empowerment and well-being among Baloch and other ethnic groups," The Dialogue, vol.8, pp. 85-97, Jan-March 2013.

[7] S. McWilliams and A. E. Barret, "Online dating in middle and later life: gendered expectations and experiences," Journal of Family Issues, vol. 35, pp. 411-436, Feb. 2014, doi: 10.1177/0192513X12468437.

[8] P. X. Francina and M. V. Joseph, "Women Empowerment: The Psychological dimension," Rajagiri Jojurnal of Social Developmet, vol. 5, pp. 163-176, Dec. 2013, ISSN: 0973-3086.

[9] S Mosedale, "Assessing women's empowerment: towards a conceptual framework," Journal of international development, vol.17, pp. 243-257, Feb. 2005, doi: 10.1002/jid.1212.

[10] M. A. Zimmerman, "Psychological Empowerment: Issues and Illustrations," American Journal of Community Psychology, vol. 23, pp. 581-599, 1995, doi:10.1007/BF02506983.

[11] M. A. Osman and M. Tanner, "The influence of telecentre components on the psychological empowerment of underserved community members in the Western Cape, South Africa," Electronic Journal of Information Systems in Developing Countries (EJISDC), vol. 81, pp.1-29, Dec. 2017, doi: 10.1002/j.1681-4835.2017.tb00596.x.

[12] A. R. Schrock, "Communicative affordances of mobile media: Portability, availability, locatability, and multimediality," International Journal of Communication , vol.9, pp.1229-1246, 2015, ISSN: 1932-8036.

[13] J. J. Wilemse and V. Bozalek, " Exploration of the affordances of mobile devices in integrating theory and clinical practice in an undergraduate nursing programme," Curationis, vol. 38, pp. 1-10, Sept. 2015, doi: 10.4102/curationis.v38i2.1510.

[14] D. Norman, "The design of everyday things: Revised and expanded edition," Basic Books (AZ), 2013, ISBN 978-0-465-05065-9.

[15] E. Timmermans and C. Courtois, "From swiping to casual sex and/or committed relationships: Exploring the experiences of Tinder users," The Information Society, vol. 34, pp. 59-70, March 2018, doi: 10.1080/01972243.2017.1414093.

[16] G. Ranzini and C. Lutz, "Love at first swipe? Explaining Tinder self-presentation and motives," Mobile Media & Communication, vol. 5, pp. 80-101, Jan. 2017, doi: 10.1177/2050157916664559.

[17] G. David and C. Cambre, "Screened Intimacies: Tinder and Swipe Logic," Social Meda + Society, vol. 2, pp. 1-11, April 2016, doi: 10.1177/2056305116641976.

 ISBN: 978-1-61208-673-6

[18] S. L. Chan, "Who uses dating apps? Exploring the relationships among trust, sensation-seeking, smartphone use, and the intent to use dating apps based on the Integrative Model," Computers in Human Behavior, vol. 72, pp. 246-258, Feb. 2017, doi: 10.1016/j.chb.2017.02.053.

[19] E. March, R. Grieve, J. Marrington, and P. K. Jonason, P, "Trolling on Tinder® (and other dating apps): Examining the role of the Dark Tetrad and impulsivity," Personality and Individual Differences, vol. 110, pp. 139–143, Feb. 2017, doi: 10.1016/j.paid.2017.01.025.

[20] S. Duguay, "Dressing up Tinderella: interrogating authenticity claims on the mobile dating app Tinder," Information, Communication & Society, vol. 20, pp. 351-367, 2017, doi: 10.1080/1369118X.2016.1168471.

[21] L.A Richey, " 'Tinder Humanitarians': The Moral Panic Around Representations of Old Relationships in New Media," Javnost-The Public, vol. 23, pp. 398-414, Nov. 2016, doi: 10.1080/13183222.2016.1248323.

[22] A. M. Browning, "CNE article: moral distress and psychological empowerment in critical care nurses caring for adults at end of life," American Journal of Critical Care, vol. 22, pp. 143-151, March 2013, doi: 10.4037/ajcc2013437.

[23] Z. Li, "Psychological empowerment on social media: Who are the empowered users?" Public Relations Review, vol. 42, pp. 49-59, March 2016, doi: 10.1016/j.pubrev.2015.09.001.

[24] J. Ward, "What are you doing on Tinder? Impression management on a matchmaking mobile app," Information, Communication & Society, vol. 20, Nov. 2016, pp. 1644-1659, doi: 10.1080/1369118X.2016.1252412.

[25] J. L. Gibbs, N. B Ellison, and C. H. Lai, "First comes love, then comes Google: An investigation of uncertainty reduction strategies and self-disclosure in online dating," Communication Research, vol. 38, pp. 70-100, Feb. 2011, doi: doi.org/10.1177/0093650210377091.

[26] L. Frischlich, D. Rieger, T. Dratsch, and G. Bente, "Meet Joe Black? The effects of mortality salience and similarity on the desire to date in-group versus out-group members online," Journal of Social and Personal Relationships, vol. 32, pp. 509-528, June 2015, doi: 10.1177/0265407514536305.

[27] A. Zimmerman, "Empowerment Theory: Psychological, Organizational and Community Levels of Analysis," Handbook of Community Psychology, Springer, pp. 43-63, 2000, doi: doi.org/10.1007/978-1-4615-4193-6_2.

[28] D. Zytko, S.A. Grandhi, and Q. Jones, "Impression management struggles in online dating," The 18th international conference on supporting group work, Association for Computing Machinery (ACM), pp. 53-62, Nov. 2014, doi: 10.1145/2660398.2660410.

[29] N. A. Peterson, C. H. Peterson, L. Agre, B. D. Christens, and C. M. Morton, "Measuring youth empowerment: Validation of a sociopolitical control scale for youth in an urban community context ," Journal of Community Psychology, vol. 39, pp. 592-605, June 2011, doi:10.1002/jcop.20456.

[30] C. Masden and W.K. Edwards, "Understanding the role of community in online dating," The 33rd annual ACM conference on human factors in computing systems, Association for Computing Machinery (ACM), pp. 535-544, April 2015, doi: 10.1145/2702123.2702417.

[31] M. R. Lopes and C. Vogel, "Women's perspective on using Tinder: a user study of gender dynamics in a mobile device application," The 35th ACM International Conference on the Design of Communication, Association for Computing Machinery (ACM), pp. 1-10, Aug. 2017, doi: 10.1145/3121113.3121220.

[32] M. Tanner and M. L Pan, "The Impact of Impromptu User Story Modifications on the Project, Customers and Team Members During a Sprint," The MakeLearn and TIIM Joint International Conference, ToKnowPress, Jan. 2015, ISSN: 2232-3309.

[33] J. Buchan, M. Bano, D. Zowghi, S. MacDonell, and A. Shinde, "Alignment of Stakeholder Expectations about User Involvement in Agile Software Development," The 21st International Conference on Evaluation and Assessment in Software Engineering, Association for Computing Machinery (ACM), pp. 334-343, June 2017, doi: 10.1145/3084226.3084251.

[34] . Lean and J. Condie, "The Curious Case of Tinder Tourism," Critical Tourism Studies Proceedings, Digital Commons. [Online]. Available from: digitalcommons.library.tru.ca/cts-proceedings/vol2017/iss1/69 [retrieved: Aug. 2018].

[35] A. Hess and C. Flores, "Simply more than swiping left: A critical analysis of toxic masculine performances on Tinder Nightmares," New media and society, vol. 20 , pp. 1-18, Dec. 2016, doi: 1461444816681540.

[36] T. Pond and P. Farvid, " 'I do like girls, I promise': Young bisexual women's experiences of using Tinder," Psychology of Sexualities Review, vol. 8, pp. 6-24, Nov. 2017, ISSN: 2047-1475.

[37] S. R. Sumter, L. Vandenbosch, and L. Ligtenberg, "Love me Tinder: Untangling emerging adults' motivations for using the dating application Tinder," Telematics and Informatics, vol. 34, pp. 67-78, Feb. 2017, doi: 10.1016/j.tele.2016.04.009.

[38] K. L. Modecki, M. J. Zimmer-Gembeck, and N. Guerra, "Emotion Regulation, Coping, and Decision Making: Three Linked Skills for Preventing Externalizing Problems in Adolescence," Child Development, vol. 88, pp. 417-426, Feb. 2017, doi:10.1111/cdev.12734.

[39] G. Orosz et al., "The personality, motivational, and need-based background of problematic Tinder use," Journal of behavioral addictions, vol. 7, pp. 1-16, April 2018, doi: 10.1556/2006.7.2018.21.

[40] H. Noble and J. Smith, "Issues of validity and reliability in qualitative research," Evidence-Based Nursing, vol. 0, pp. 1-2, Feb. 2015, doi: 10.1136/eb-2015-102054.

[41] K. Yilmaz, "Comparison of Quantitative and Qualitative Research Traditions: epistemological, theoretical, and methodological differences," European Journal of Education, vol. 48, pp. 311-325, May 2013, doi: 10.1111/ejed.12014.

[42] L. Leung, "Validity, reliability, and generalizability in qualitative research," Journal of Family Medicine and Primary Care, vol. 4, pp. 324-327, Jul-Sep. 2015, doi: 10.4103/2249-4863.161306.

[43] J. M. Morse, M. Barret, M. Mayan, K. Olson, and J. Spiers, "Verification strategies for establishing reliability and validity in qualitative research," International journal of qualitative methods, vol. 1, pp. 13-22, June 2002, doi: 10.1177/160940690200100202.

[44] okCupid. *Why OkCupid is changing how you message.* [Online]. Available from: https://theblog.okcupid.com/why-okcupid-is-changing-how-you-message-f14d492e7853

[45] T. R. Osborne et al., "Improving the assessment of quality of life in the clinical care of myeloma patients: the development and validation of the Myeloma Patient Outcome Scale (MyPOS)," BMC cancer, vol. 15, pp. 1-12, April 2015, doi: doi.org/10.1186/s12885-015-1261-6.

[46] C. Mackintosh, "Protecting the self: A descriptive qualitative exploration of how registered nurses cope with working in surgical areas," International Journal if Nursing Studies, vol. 44, pp. 982-990, Aug. 2007, doi: 10.1016/j.ijnurstu.2006.04.009.

[47] F. C. I. Yang, "Remediating Japanese Dramas: Exploring the Politics of Gender, Class, and Ethnicity in Loser ‑ Dog Queen in Taiwan," The Journal of Popular Culture, vol.46, pp. 1070-1091, Oct. 2013, doi: 10.1111/jpcu.12067.

 ISBN: 978-1-61208-673-6

Understanding Digital Ethnography: Socio-computational Analysis of Trending YouTube Videos

Muhammad Nihal Hussain[1], Kiran Kumar Bandeli[1], Serpil Tokdemir[1], Samer Al-khateeb[2], Nitin Agarwal[1]
[1]University of Arkansas at Little Rock (UALR)
Little Rock, Arkansas, USA
{mnhussain, kxbandeli, sxtokdemir, nxagarwal}@ualr.edu
[2]Creighton University
Omaha, Nebraska, USA
sameral-khateeb1@creighton.edu

Abstract— **The online video sharing website - YouTube, which was launched in February 2005 to help people share videos of well-known events, has rapidly grown to be a cultural phenomenon for its massive user-base. According to Alexa, the web traffic monitoring tool by Amazon, YouTube is the second most popular website globally. Still there is a lack of systematic research - both qualitative as well as quantitative - focusing on the video-based social networking site as compared to other social media sites. Video comments serve as a potentially interesting data source to mine implicit knowledge about users, videos, categories, and community's interests. In this research, we studied top 200 YouTube videos trending daily for a 40-day period separately in the United States of America (USA) and the Great Britain (GB) regions. We collected data for 7,998 videos trending in the USA throughout the 40-day time period and 7,995 videos trending in the GB regions. We studied content engagement behavior of users in the USA and GB regions by analyzing views, likes, dislikes, and comments on the set of trending videos. The study helped us glean some of the digital ethnographic behaviors of the users in these two regions. This paper presents highlights of the similarities and differences observed in such behaviors between the USA and GB regions.**

Keywords-YouTube; digital ethnography; USA; GB; social network analysis

I. INTRODUCTION

YouTube, an online video sharing website was launched in early 2005 to help people share videos [1]. Since then sharing video content has become a cultural phenomenon. According to Alexa, the web traffic monitoring service owned by Amazon, YouTube accounts for 15.3% of traffic from search on Internet [2]. YouTube is also the second most popular website globally with over 300 hours of videos uploaded every minute and 5 billion videos watched every single day [3]. YouTube overall, and even YouTube on mobile alone, reaches more 18–34 and 18–49-year-olds than any cable network in the USA. YouTube has launched local versions of its platform in more than 88 countries. One can navigate YouTube in 76 different languages covering 95% of the Internet population [3].

Several investigations have reported digital communication tools especially social media plays an important role in how people interact, communicate, and share information. As evident in the mass protests and cyber campaigns, social media platforms help individuals to coordinate, spread messages, organize, and mobilize support for their efforts. While YouTube and other social media sites have helped frame and change the nature of online discourse, prolific linking of the content across multiple social media sites has helped spread narratives at lightning fast speeds. Hence, it is imperative to examine the emerging video-based social media platforms.

While significant body of work exists that analyze Twitter and other such social media platforms, there is a lack of systematic research focusing on video-based social networking sites. A few studies shed insights into the dynamics of online discussions on YouTube [4]. In addition to easy publishing and distribution of nearly any kind of video content, YouTube provides various features to engage with the video content, such as liking or disliking a video, commenting on a video, replying to a comment, liking or disliking a comment, or posting a video response. Comments on the videos may also be studied to extract insights into audience reactions to important issues or towards particular videos. Comments serve as a potentially interesting data source to mine implicit knowledge about the video's content, viewers, regions, and community interests. While some studies have focused on specific genre of videos (e.g., childbirth, coming out for lesbian, gay, bisexual, transgender and queer – LGBTQ – communities) [5] to understand the society's pulse, others have focused on the type of information in these videos [6] and the threats these videos pose to the societies [7].

In this research, we study the content engagement and consumption behaviors on YouTube that further helps develop a digital ethnographic mapping of user behaviors in terms of likes, comments, sentiments, and cross linking with other social media channels – compared across the United States of America (USA) and the Great Britain (GB). We analyzed top 200 YouTube videos trending daily for a 40-day time period separately in the USA and the GB regions. The dataset was obtained from Kaggle [8] and includes title, category, URL, ID, comments, views, likes, and dislikes for each video. Further, we enhanced the dataset by extracting associations with other social media platforms from the YouTube channels of these trending videos. This enhanced dataset was used to conduct a digital ethnographical

 ISBN: 978-1-61208-673-6

mapping of users' video content generation, sharing, and consumption behaviors. We found interesting correlations between various ways users engage with videos for both USA and GB regions. We also found multiple similarities and stark differences in behavior of users of these regions. We further studied how prolific YouTubers leverage other social media platforms to disseminate their content more widely and gain audience. Detailed analysis and implications of the findings are presented in Section III. Next, we present literature review in Section II.

II. LITERATURE REVIEW

Steady rise in YouTube's popularity has attracted a surge of research. This section summarizes studies on YouTube users' watching preferences, the video recommendation system, users commenting behaviors and deviant behaviors.

For online video watching in general, a study of USA internet users in 2009 [9] focused on viewer's preferences of categories. The study found that 50% of adults had watched funny videos, 38% had watched educational videos, 32% had watched TV shows or movies, and 20% had viewed political videos. In terms of common content categories in YouTube, music videos are a significant presence in YouTube, probably accounting for about a quarter of videos with entertainment, comedy and sports categories [10]. 60% of videos are watched at least 10 times during the first day in which they are posted [11]. One of the very first studies also showed that videos that did not attract many viewers within the first few days of publication were unlikely to grow an audience later on [11]. Even so, once a viewer is on the YouTube website, by showing related videos (or recommended videos) to him/her, YouTube attempts to increase the time spent by individuals on the site itself. Davidson et. al [12] explain the video recommendation system of YouTube. YouTube uses association rule mining (videos watched in within same session) to generate sets of related videos. To compute personalized recommendations for a user, user's activities or interactions like videos watched, liked, rated, or added to playlist, are used to build a seed set of recommended videos. Videos related to these seed set are selected as relevant videos but to enhance user experience, a subset of these videos with/from diverse categories are ultimately recommended to the user. Zhou et. al [13] argue YouTube's recommendation system has significant impact on views and virality of videos.

Although there have been some large-scale quantitative investigations into YouTube [4], few have focused on discussions in comments. Most YouTube research seems to be small-scale and qualitative that give insights into how discussions can occur around videos without giving broad overall patterns that are of use. However, the study [14] was an exception, which identified patterns in user types that can be used to predict users' likely behaviors. Yet, a little is known about YouTube discussions/comments in general including the role of sentiments. YouTube comments are textual and much research has investigated the limitations and peculiarities of the electronic text. Early studies were particularly concerned that the absence of the nonverbal channel in textual communication would lead to widespread misunderstanding, particularly in short message formats, such as mobile phone texting [15]. In response, however, a number of conventions have emerged to express sentiment in short informal text, such as emoticons and deliberate non-standard spellings [16]. Work on sentiment classification and opinion mining such as [17][18] deals with the problem of automatically assigning opinion values (viz., positive, negative, or neutral) to documents or topics using various text-oriented and linguistic features. Recent works in this area also use SentiWordNet [19] to improve sentiment classification performance.

The rise in usage and popularity of social media sites, such as YouTube, has made them particularly vulnerable for abusive behaviors from bots and troll accounts that can post spam comments in large volumes [20]. According to O'Callaghan et. al. [21], bot-posted spam comments are often associated with orchestrated campaigns and can be detected by assessing the similarity in their comments. Authors noted that these cyber campaigns remain active for long periods of time and popular videos are usually targets of these spam comments. Although, it is unknown how far user comments in this context may promote hate speech or which videos are thought to deliver their aimed potential to users.

However, the problem setting in these papers differs from ours as we conduct a comparative behavioral analysis of users' engagement and consumption of content on YouTube in terms of likes, comments, sentiments, and cross linking with other social media platforms across the USA and GB regions.

III. METHODOLOGY

Here, we describe the datasets used for the research, overall methodology and discuss results from our analysis.

A. Data Collection

We obtained the dataset consisting of a list of top trending videos on YouTube from Kaggle [8]. YouTube provides top 200 videos daily that are trending or popular. A dataset of these top 200 videos trending daily in USA and GB regions from September 13, 2017 to October 22, 2017 was obtained from Kaggle [8]. There were 7,998 videos trending in the USA throughout the 40-day time period with 2,395 unique videos and 7,995 videos trending in the GB with 1,769 unique videos. Out of the 4,164 videos, 770 videos were trending in both the countries. The original dataset from Kaggle has the following attributes: *URL of the video*, *video ID*, *title of the video*, *title of the channel* that published the video, *category* in which the video belongs to, *number of views*, *number of likes*, *number of dislikes*, *number of comments* the video received at the time data was collected, and the *date the video was trending*.

 ISBN: 978-1-61208-673-6

We enhanced the dataset obtained from Kaggle by adding the description of the video, date the channel was created, and the number of subscribers of the channel. It is a common practice among prominent YouTubers to associate their various social media accounts with their YouTube channel. Using Web Content Extractor (WCE) [22], we collected these social media associations and used them to study the cross-media integration. Due to the noisy nature of the data, several data processing steps such as data standardization, noise elimination, and data formatting were performed.

In Table I, we present high-level statistics of the attributes used from USA and GB regions.

TABLE I. YOUTUBE DATA STATISTICS FOR THE USA AND GB REGIONS.

Attributes (total number of)	USA	GB
Trending videos	7,998	7,995
Unique trending videos	2,364	1,736
Views	1,652,652,122	1,105,805,449
Likes	109,231,965	91,084,832
Dislikes	5,249,389	3,763,875
Comments	6,528,503	4,894,060
Commenters	3,824,862	3,083,770
Likes on comments	45,816,147	35,514,209
Replies on comments	2,112,200	1,605,248

B. Analysis and Findings

We conducted comparative analysis on the trending videos. First, we examined differences and similarities between users' interests in USA and GB regions by analyzing the categories of the trending videos in each region. Videos in the *Entertainment*, *Music*, and *People & Blogs* categories tend to trend more for both USA and GB regions (see Figure 1). *Comedy* videos are watched more in USA, while *Sports* videos trend in GB. Further, *Political* videos trend more in USA than in GB.

Next, we study the lifespan of trending videos in the USA and GB regions. We analyzed the number of days a video would trend on YouTube in the USA and GB regions (see Figure 2). We found that in both the regions there is a sharp decline in the popularity of videos after a certain number of days, i.e., for USA, videos trend for the first 4 days, where their popularity increases gradually after which their popularity declines sharply. There was no video that was popular for more than 8 days, i.e., the longest lifespan. For GB region, a similar trend is observed; however, a sharp decline in popularity occurs after the 6th day. From Figure 2, we can observe that videos in the GB region have longer lifespan as compared to USA videos. The longest lifespan for videos trending in USA region was 8 days, while for GB region the longest lifespan was 13 days. This dissimilarity between the two regions raises multiple questions, do users in GB tend to watch trending videos more repeatedly? And/or does the information stays longer in the social networks in GB than USA? In other words, it takes longer for information to become obsolete in GB than USA, suggesting information propagation is either slower or more long-lasting in GB than USA. We investigate this phenomenon further and reflect on the possible reasons later in the paper.

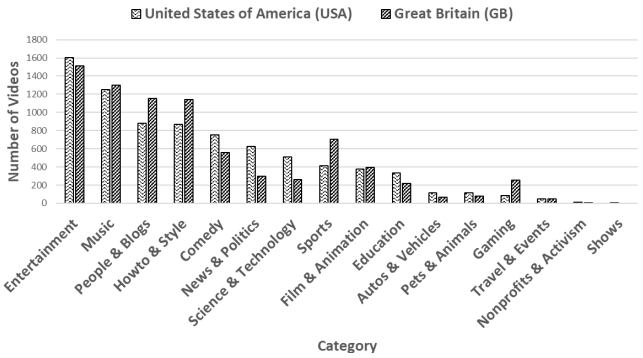

Figure 1. Number of trending videos per category for USA and GB regions.

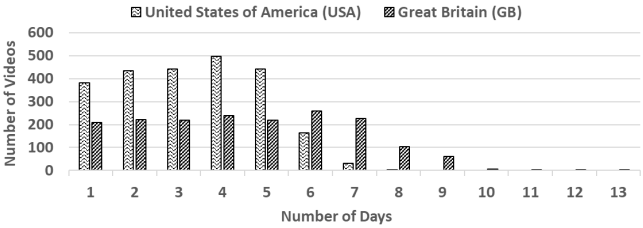

Figure 2. Lifespan of trending videos in USA and GB regions.

Next, we examine the role of integration of other social media platforms with YouTube for the trending videos in both regions. From the enhanced data, we know the various social media platforms that are associated with the trending videos. Top social media platforms that were found to be linked to the trending videos in USA and GB regions are Twitter, Facebook, Instagram, Google Plus, and Tumblr, in that order. In the USA region, most videos are affiliated with 5 social media sites, while most views were obtained by videos that were affiliated with 10 social media sites. This implies having 5 social media platforms associated with the video increases its possibility of going viral but associating more social media platforms increases the probability of video getting more views. Having other social media platforms associated with the video certainly brings more visibility through shares or other cross-media

 ISBN: 978-1-61208-673-6

information spillovers. This might be one of the reasons for weak correlation, in Figure 3, between number of views and number of social media sites affiliated. A similar finding was observed for the GB region (Figure 4), where most videos are affiliated with 5 social media sites, while most views were obtained by videos that were affiliated with 20 social media sites.

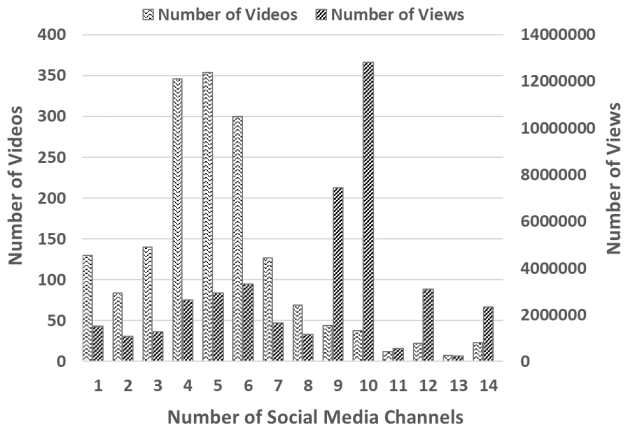

Figure 3. Social media affiliation of the trending videos in USA.

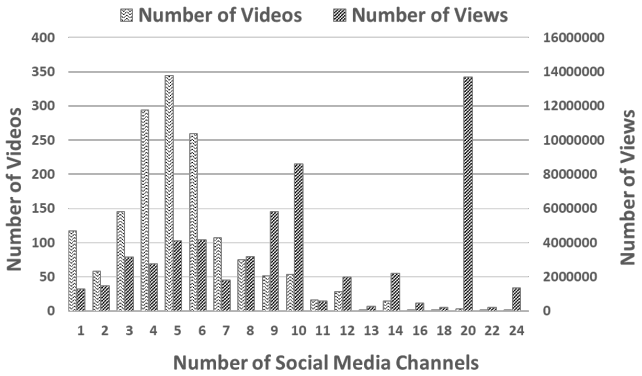

Figure 4. Social media affiliation of the trending videos in GB.

Next, we look at the network of various social media affiliations for the trending videos in both regions. The network map presented in Figure 5 helps us understand the integration of other social media sites with YouTube in the USA region. Different colors denote different clusters. Clustering analysis is done by Gephi network visualization using modularity algorithm. The network shows us the clusters of the most commonly used social media websites and how those social media platforms connect/relate and interact with each other. We found that Facebook, Instagram, Twitter, Google Plus, Pinterest, and Tumblr are the most popular social media sites. Same set of social media sites are also more frequently associated with the trending videos in the GB region (Figure 6). It is clear from the USA region's social media map that social media platforms related to the Music category (green colored nodes in top left corner of the network) clustered independently compared to other social media sites. This implies that the videos posted in the Music category were, quite understandably, shared or affiliated to completely.

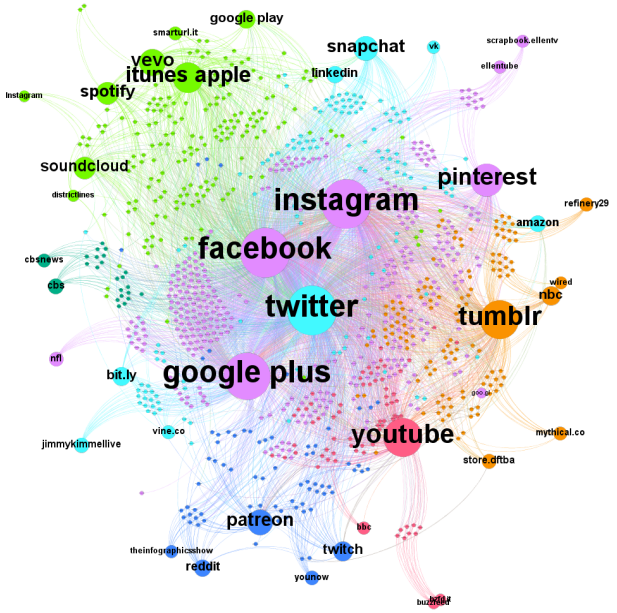

Figure 5. Social media map of the trending videos in USA

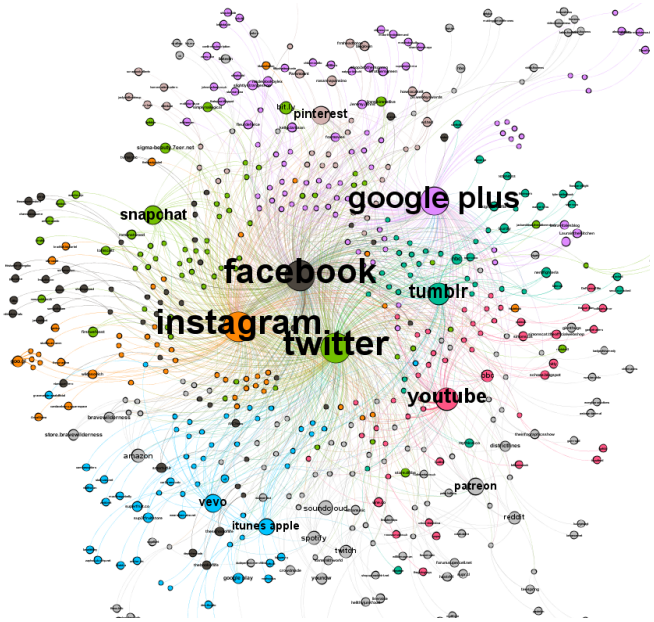

Figure 6. Social media map of the trending videos in GB

different family of social media sites. However, the media map of the GB region seems more well-connected than the USA region implying that the trending videos in GB region tend to have a richer cross-media integration on their YouTube channels that could further help explain the longer lifespan of trending videos in GB as compared to USA

 ISBN: 978-1-61208-673-6

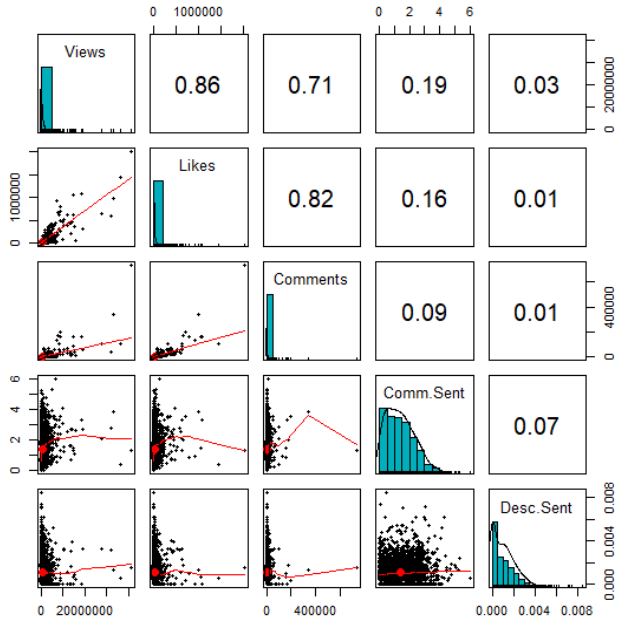

Figure 7. Correlation analysis of the trending videos in USA.

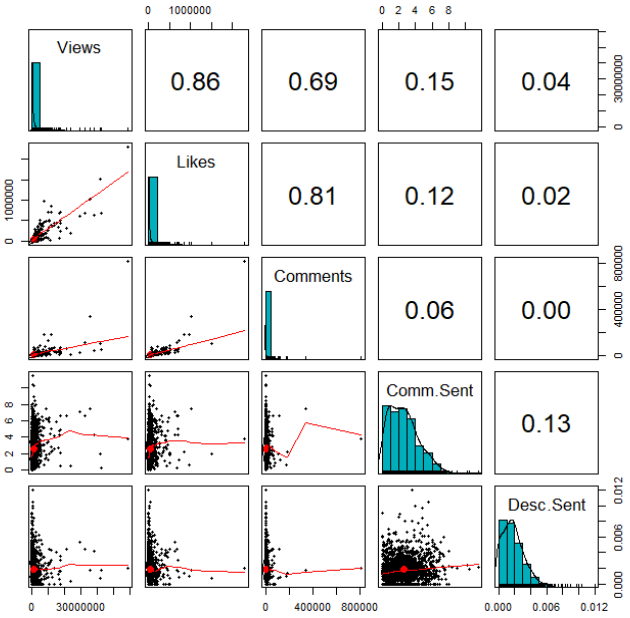

Figure 8. Correlation analysis of the trending videos in GB.

To examine users' content engagement behavior on YouTube, next, we conduct correlation analysis among number of views, number of likes, number of dislikes, number of comments, sentiments of comments, and sentiment of video description for the trending videos in both regions. The correlation charts for the USA region are shown in Figure 7. We observed that the more comments a video has the more neutral the sentiments are. On the other hand, fewer comments a video has more polarized the sentiments are (column 3 and row 4). A similar behavior was observed between sentiments of comments and number of views on videos, i.e., the more views a video has the more neutral the sentiments are, and the fewer views a video has more polarized the sentiments are (column 1 and row 4).We observed a strong correlation (0.82) between the number of comments and number of likes (column 2 and row 3) as well as between number of comments and number of views (column 1 and row 3 with correlation coefficient of 0.71). A weak correlation (0.07) was observed between sentiments from comments and sentiments from video description (column 4 and row 5). Further research is needed to investigate the reasons for a relatively weak correlation between sentiments from the comments and video description, which might be attributed to unrelated comments generated by bots or troll accounts. A similar analysis is conducted for the trending videos in GB region. The correlation charts for the GB region are shown in Figure 8. Comparing the correlation analysis of the trending videos in both regions resulted in similar findings. In both the regions, number of likes and number of views show a stronger correlation as compared to number of dislikes and number of views. This implies that if a viewer watches the whole video he/she is more likely to like the video than dislike it. A weak correlation (0.07) was observed between sentiments from comments and sentiments from video description (column 4 and row 5). Further investigations are required to test the causal relationship of these correlations.

IV. Conclusion and Future Work

Although YouTube is the second most popular website globally with over 300 hours of videos uploaded every minute and 5 billion videos being watched every single day, there is a lack of systematic research focusing on the video-based social networking site as compared to other social media sites. There are a few studies that shed insights into the dynamics of online discussions on YouTube. Video comments serve as a potentially interesting data source to mine implicit knowledge about users, videos, categories, and community interests.

In this research, we studied top 200 YouTube videos trending in the USA and GB regions. We did a comparative study of users' video consumption behaviors and engagement to find stark difference in preferences of users of these regions. Trending videos in each region provided a glimpse of the interests of the viewers. For instance, viewers in USA are more interested in Comedy and Politics, while viewers in GB are more interested in Sports. Additionally, videos in the GB region tend to have longer lifespan compared to the USA region. Furthermore, videos in GB region are shared on more social media platforms than in USA.

We found correlation between various aspects of user engagement, such as number of comments a video received, and the polarity of these comments was inversely correlated, number of views of a video and its comments were strongly correlated. We did not test the causality of this relationship and intend to collect more data to do the same. We also

 ISBN: 978-1-61208-673-6

found correlation coefficient for above mentioned relationships are different for USA and GB regions. Since the data for this analysis was collected only on a 40-day period, more data samples are needed for validation. Although we observed a strong correlation between number of comments and number of likes and number of views, a weak correlation was observed between sentiment of comments and sentiments of video description. The fact that commenting behavior is somewhat unrelated to the videos' content, warrants further analysis of a likely presence of spam comments from bots or troll accounts.

This research has attempted to shed light on content consumption and engagement behaviors on one of the most prominent video-based social media platforms, i.e., YouTube, in USA and Great Britain regions. As exciting as these findings are, the analysis is limited to the 40-day study period and result might vary for different time-periods. We plan to investigate the reasons for the observations in more depth and identify the ethno-digital cultural factors that help manifest the differences in the content consumption and sharing behaviors. We envision this study will help open doors for innovative and foundational work in analyzing digital behaviors of our societies (good, bad, and the ugly) as the Information and Communication Technology (ICT) landscape evolves. For instance, we can study complex questions such as, what is the role of the platform and its users in disseminating misinformation? Which societies or strata within a society are more vulnerable to misinformation? More broadly, our study motivates the need for development of methodologies to diagnose novel pathologies of online social media.

ACKNOWLEDGMENT

This research is funded in part by the U.S. National Science Foundation (IIS-1636933, ACI-1429160, and IIS-1110868), U.S. Office of Naval Research (N00014-10-1-0091, N00014-14-1-0489, N00014-15-P-1187, N00014-16-1-2016, N00014-16-1-2412, N00014-17-1-2605, N00014-17-1-2675), U.S. Air Force Research Lab, U.S. Army Research Office (W911NF-16-1-0189), U.S. Defense Advanced Research Projects Agency (W31P4Q-17-C-0059) and the Jerry L. Maulden/Entergy Fund at the University of Arkansas at Little Rock. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

REFERENCES

[1] J. Hopkins, "Surprise! There's a third YouTube co-founder - USATODAY.com." [Online]. Available: https://usatoday30.usatoday.com/tech/news/2006-10-11-youtube-karim_x.htm. [Retrieved: Aug 19, 2018].

[2] "Youtube.com Traffic, Demographics and Competitors - Alexa." [Online]. Available: https://www.alexa.com/siteinfo/youtube.com. [Retrieved: August 19, 2018].

[3] S. Brain, "YouTube Company Statistics," *2017 Statistic Brain Research Institute*, 01-Sep-2016. [Online]. Available: https://www.statisticbrain.com/youtube-statistics/. [Retrieved: August 16, 2018].

[4] P. Gill, M. Arlitt, Z. Li, and A. Mahanti, "Youtube traffic characterization: a view from the edge," in Proceedings of the 7th ACM SIGCOMM conference on Internet measurement, 2007, pp. 15–28.

[5] K. Thorson, B. Ekdale, P. Borah, K. Namkoong, and C. Shah, "YouTube and Proposition 8: A case study in video activism," Information, Communication & Society, vol. 13, no. 3, pp. 325–349, 2010.

[6] P. L. Steinberg, S. Wason, J. M. Stern, L. Deters, B. Kowal, and J. Seigne, "YouTube as source of prostate cancer information," Urology, vol. 75, no. 3, pp. 619–622, 2010.

[7] S. P. Lewis, N. L. Heath, J. M. St Denis, and R. Noble, "The scope of nonsuicidal self-injury on YouTube," Pediatrics, p. peds. 2010-2317, 2011.

[8] M. J, "Trending YouTube Video Statistics and Comments," *Kaggle*, Oct-2017. [Online]. Available: https://www.kaggle.com/datasnaek/youtube/data. [Retrieved: August 18, 2018].

[9] K. Purcell, "The State of Online Video," *Pew Research Center: Internet, Science & Tech*, 03-Jun-2010. [Online]. Available: http://www.pewinternet.org/2010/06/03/the-state-of-online-video/. [Retrieved: August 18, 2018].

[10] X. Cheng, J. Liu, and C. Dale, "Understanding the characteristics of internet short video sharing: A youtube-based measurement study," IEEE Transactions on Multimedia, vol. 15, no. 5, pp. 1184–1194, 2013.

[11] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "Analyzing the video popularity characteristics of large-scale user generated content systems," Ieee/Acm Transactions On Networking (Ton), vol. 17, no. 5, pp. 1357–1370, 2009.

[12] J. Davidson et al., "The YouTube video recommendation system," in Proceedings of the fourth ACM conference on Recommender systems, 2010, pp. 293–296.

[13] R. Zhou, S. Khemmarat, and L. Gao, "The impact of YouTube recommendation system on video views," in Proceedings of the 10th ACM SIGCOMM conference on Internet measurement, 2010, pp. 404–410.

[14] M. Maia, J. Almeida, and V. Almeida, "Identifying user behavior in online social networks," in Proceedings of the 1st workshop on Social network systems, 2008, pp. 1–6.

[15] J. B. Walther and M. R. Parks, "Cues filtered out, cues filtered in," Handbook of interpersonal communication, vol. 3, pp. 529–563, 2002.

[16] D. Derks, A. E. Bos, and J. Von Grumbkow, "Emoticons and online message interpretation," Social Science Computer Review, vol. 26, no. 3, pp. 379–388, 2008.

[17] K. Denecke, "Using sentiwordnet for multilingual sentiment analysis," in Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on, 2008, pp. 507–512.

[18] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 2002, pp. 79–86.

[19] M. Thomas, B. Pang, and L. Lee, "Get out the vote: Determining support or opposition from Congressional floor-debate transcripts," in Proceedings of the 2006 conference on empirical methods in natural language processing, 2006, pp. 327–335.

[20] A. Sureka, "Mining user comment activity for detecting forum spammers in YouTube," arXiv preprint arXiv:1103.5044, 2011..

[21] D. O'Callaghan, M. Harrigan, J. Carthy, and P. Cunningham, "Network Analysis of Recurring YouTube Spam Campaigns.," in ICWSM, 2012.

[22] Newprosoft, "Web Content Extractor," *Newprosoft Web Data Extraction Software*, 2004. [Online]. Available: http://www.newprosoft.com/web-content-extractor.htm. [Retrieved: August 18, 2018].

 ISBN: 978-1-61208-673-6