

Bridging the Gap: A Reflection on an Interdisciplinary Approach to Social Media Research

Pablo Aragón
David Laniado

Jessica G. Neff
Ricard Ruiz de Querol

Andreas Kaltenbrunner
Carlos Ullod

Karolin Kappler
Yana Volkovich

Fundació Barcelona Media – Social Media Research Group
Avinguda Diagonal 177, Planta 9, 08018 Barcelona, Spain
{name.surname}@barcelonamedia.org

ABSTRACT

In this paper we review the recent research efforts of the Social Media (SM) Research Group at Fundació Barcelona Media. The group was formed with the objective of combining qualitative and quantitative methodologies to generate new knowledge about the interplay of social behavior and new ‘social media’, as well as devising methods, tools and technologies that facilitate the acquisition and application of this knowledge in a variety of enterprise and social environments. The SM group conducts empirical research on data mining, the characterization and modeling of user behavior on social networks and social media, and motivations underlying the behavioral patterns that we uncover. The team’s current work includes analysis, modeling and prediction of temporal, social and structural interaction patterns in social media websites, analysis of conversations and the social architecture in social media, enhanced ranking and community detection algorithms, and user profiling based on interaction logs.

1. INTRODUCTION

To illustrate the approaches different disciplines take to the analysis of social networks and social media data, perhaps it is best to start with an anecdote.

A sociologist and a computer scientist are brainstorming about a new research project to study an online social network. The sociologist expresses concerns regarding the size of the sample, wanting to have at least a couple of hundred users included in the study. The computer scientist responds, “Either we study the whole network of 10 million users, or it doesn’t make sense to study it at all!”

This anecdote is a true story from our research team, and it serves to illustrate that communication and collaboration between disciplines is not necessarily evident. Different disciplines may share the same

terminology, the same research population, and sometimes even the same research objectives, but they come from different worlds, follow different approaches and view the world through different lenses.

2. THE SOCIAL MEDIA RESEARCH GROUP AT FUNDACIÓ BARCELONA MEDIA

The Social Media (SM) Research Group at Fundació Barcelona Media is an interdisciplinary team that strives to reconcile these two distinct approaches to research. Our research group brings together a diverse mix of international researchers with cross-disciplinary profiles, including physicists, mathematicians, computer scientists, sociologists, and communications scholars. As an interdisciplinary research team we are constantly grappling with the challenge of how to leverage the unique array of methodological approaches and theoretical perspectives that our team members bring to the table in order to produce cutting edge research.

The SM group was formed with the objective of combining qualitative and quantitative methodologies in order to generate new knowledge about the interplay of social behavior and new ‘social media’, as well as devising methods, tools and technologies that facilitate the acquisition and application of this knowledge in a variety of enterprise and social environments. We conduct empirical research on data mining, the characterization and modeling of user behavior on social networks and social media, and the motivations underlying the behavioral patterns that we uncover. Our current work includes analysis, modeling and prediction of temporal, social and structural interaction patterns in social media websites, analysis of conversations and the social architecture in social media, enhanced ranking and community detection algorithms, and user profiling based on interaction logs.

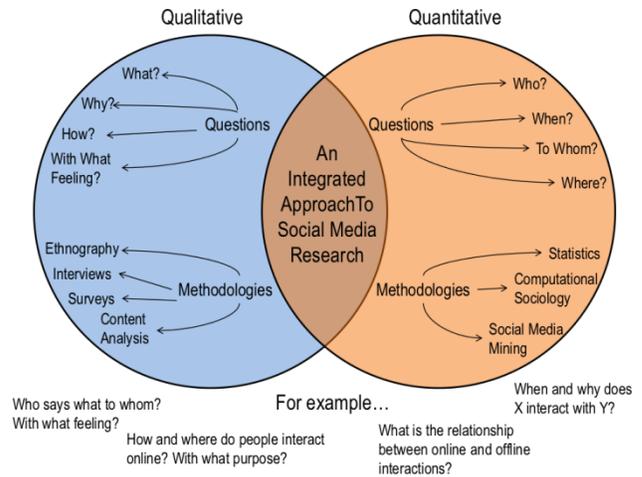
3. AN INTERDISCIPLINARY APPROACH TO SOCIAL MEDIA RESEARCH

Extant research on social media has largely been conducted within two distinct theoretical and methodological approaches. On the one hand there are theories and methodologies whose aim is to explicitly describe and measure variables, answering the following questions: Who is talking? When? Where? To whom? Disciplines like computer sciences, engineering, and mathematics offer their tools and approaches for this endeavor. On the other hand, social scientists strive to understand the motivations underlying structural patterns, aiming to answer questions such as What is being said? Why? How? With what feeling? As the results are more implicit, they often seem to be more “speculative” than observable. Only by answering both sets of questions posed by the different disciplines can a comprehensive picture of social media be drawn, and a deeper understanding reached. However, it is rare to see interdisciplinary research teams tackling these questions in a collaborative way. More often we see siloed research communities, rarely interacting with one another. In an effort to address this shortcoming, our group has set forth an interdisciplinary research agenda that incorporates: 1.) quantitative approaches, 2.) qualitative approaches, and 3.) integrated approaches, which incorporate methods and results from both types of research. We do not always succeed in our attempts to reconcile the different approaches. However, we are making a constant effort and are learning many valuable lessons along the way.

3.1 Quantitative Research Line

In recent years the work of the quantitative branch of the SM group has mainly focused on the analytical study of forum conversations. The work was initiated with an in-depth analysis of temporal patterns in Slashdot (Kaltenbrunner et al., 2008), where homogeneous patterns in the reaction time of the Slashdot-community to new news stories were discovered. These patterns can be used to predict the total amount of comments a news story on Slashdot will receive (using only a very short snapshot of the initial activity) as shown in Kaltenbrunner, Gómez and López (2007).

Later the structure of nested forum conversations was analysed (Gómez, Kaltenbrunner, & Lopez, 2008), introducing a balanced depth measured inspired by the h-index (Hirsch, 2005). The measure accounts for the complexity of the discussion.



This index was also applied to the discussions (about Wikipedia articles) on Wikipedia Talk pages (Laniado, Tasso, Volkovich, Kaltenbrunner, 2011). Other variants of the index (not requiring nested discussions) have been introduced in Kaltenbrunner, Bondia, Banchs (2009) to rank users and forum topics. An interactive tool to grasp the shape of nested discussion has been introduced by Pascual-Cid & Kaltenbrunner (2009) and a theoretical model that allows for encoding of the growth process behind such discussions with only a few parameters was also developed (Gómez, Kappen, Kaltenbrunner, 2011).

Furthermore, the behavioral social networks generated by the interactions in the discussion have been analyzed for Slashdot (Gómez, Kaltenbrunner, López 2008) and for discussion in a Digg-like social news website (Kaltenbrunner, Gonzalez, Ruiz De Querol, Volkovich, 2011). The later study also compares this behavioral social network with an explicitly declared follower network among the same users. The two networks show important structural differences indicating fundamental distinctions between the two types of social relations.

The question of measuring user generated content on social news sites has been raised in Volkovich, & Kaltenbrunner (2011), where the authors proposed to quantify interestingness and attractiveness with respect to the original source via an index.

For the Spanish government entity Red.es we have performed a large scale quantitative study to understand usage of social networks in Spain. In

particular, we have analyzed the invitation-only Spanish social networking service Tuenti, often referred to as the “Spanish Facebook”, and Spanish social news sharing platform Menéame. We have presented a detailed demographic and geographical study, evaluated participation and interaction patterns, discussed various approaches for definition of influential users, and, finally, designed a number of visualizations of the studied social networks.

Furthermore, we are currently analyzing the relationships between interaction strength, spatial distance and structural position of ties between members of the Tuenti network. Our results (Volkovich et al., 2012) suggested that spatial constraints on online social networks are intimately connected to structural network properties. In the same time, the amount of interaction is fairly uncorrelated to spatial distance, which is a different outcome in comparison to studies of email networks.

3.2 Qualitative Research Line

The Social Sciences and Humanities research branch of the SM-group is driven by the following sorts of research questions: Why do people use social media? What sorts of personal and contextual factors drive user patterns of behavior? How can social media be most effectively leveraged in order to support processes such as collaboration, and coordination, and knowledge management? In this context, we study user motivation and user profiles, user patterns of interaction and collaboration, specific interaction patterns, trust in social media sites, offline-online interactions and effects, and emotional patterns in Social Media communication. We have also begun to explore the idea of social media “with a purpose” - that is, how social media might be leveraged to enable processes such as social innovation.

In one of our studies, based on a qualitative and grounded theory-guided analysis of several focus groups, we explored web users’ social media practices, attitudes and motivations, and make a first effort to describe different user-profiles. Four main categories emerged: homo oeconomicus vs. homo gaudens, on one side, and homo individualis vs. homo socialis, on the other side. The first pair of categories is characterized by the difference between a pragmatic versus impulsive driven use of social media. The second pair is based on the individualistic and egocentric vs. participative and collaborative use of social media. Within these broader categories we

identified 7 subtypes of social media users: homo socialis, homo utilitatis, homo habitus, homo narcisista, homo ludens, homo patologensis, and homo romantico. In a second step, we specifically analyzed and focused the interactive/interpassive dimension of these user profiles, parting from Lacan’s and Zizek’s philosophical and psychoanalytical terminology and guiding us towards a reflection on the transformations of the social subject in the realm of the new 2.0 media. (Kappler & Ruiz de Querol, 2011).

In another study we analyzed the evolution of social media, including the social and cultural emergence of social media after the .com crash at the end of the 1990s (Ruiz de Querol & Kappler, 2010). Using different Social Theory-frames (e.g. Lipovetsky & Serroy, 2009), we studied ongoing trends, considering different cultural axioms. For example, Hypercapitalism: representing the driving force of globalization and embodied by the homo oeconomicus; Hypertecnification: defining the digital era, with individuals living an abstract life, cloistered in their new technologies, while they stay at home; Hyperindividualism, or life à la carte: centered on premises such as self-realization, subjective autonomy, hedonism, and following a narcissistic tendency; and finally, Hyperconsumerism, favoring comfort and convenience above any other thing, but leading to growing disorientation in the hypermodern societies.

In one of our current research projects we are exploring current and future trends within the social media sector. We are surveying industry experts in order to obtain a deeper understanding of: 1.) the value of information derived from social media, 2.) monetization, 3.) the role of big data, 4.) the impact of social media on the workplace, 5.) transparency, interactions between the offline and offline worlds, and 7.) the adoption and acceptance of social media across various sectors.

3.3 Integrated Research Line

3.1.1 Discussion Structure and Sentiment

In the direction of interdisciplinary studies, Slashdot discussions have been segregated by topic (Kaltenbrunner, Gonzalez-Bailon & Banchs, 2009; Gonzalez-Bailon, Kaltenbrunner & Banchs, 2010), which revealed important structural differences between topics. Political discussions were found to be more deliberative than discussions with a less

ideological topic. These differences are also confirmed by the model parameters extracted for the different topic categories (Gómez, Kappen & Kaltenbrunner, 2011).

The content of forum discussions has been analyzed (Gonzalez-Bailon, Banchs & Kaltenbrunner, 2012) using the Affective Norms for English Words (ANEW) (Bradley & Lang, 1999) on USENET discussion over political topics during prolonged times (1999-2005). It was found that the emotional time-series obtained from the discussions can help to predict the approval rate of the US-president, providing evidence for the close relation between emotion and political opinion.

3.1.2 Interactions on Wikipedia

Wikipedia is a natural test bench for our integrated approach, as one of the largest existing collaborative projects, providing complete information about user interactions. In particular, Wikipedia discussions are a place where different points of view have to be combined to find a consensus, and encyclopedic content is negotiated by editors according to community-defined policies. In Laniado, Tasso, Volkovich & Kaltenbrunner (2011) we have analyzed structural patterns of these discussions; we are currently extending this study to investigate their temporal evolution and the relation with offline events and internal community dynamics. Furthermore, to gain a deeper understanding of social dynamics and to uncover their emotional dimension, in a current project we apply state-of-art sentiment analysis techniques to Wikipedia discussions, focusing on characterizing different categories of editors and the interactions among them, and finding significant differences according to the gender and to the experience and status of the users.

Another ongoing project aims at understanding the interplay between different political identities and a common community identity (Gould et al, 2012). We address this question by examining patterns of interaction among self-declared Republicans and Democrats within the Wikipedia community. We use network analysis to map patterns of user interaction, and sentiment analysis to examine the emotional valence of interactions. We are performing a more fine grained qualitative analysis of the content of cross- and within-party discussion, and are also examining user identity and representation through qualitative analysis of user pages.

3.1.3 Political Interactions on Twitter

We are also examining political interactions within Twitter. In a current study we are analyzing patterns of interaction among users who were actively tweeting around the topic of the presidential election in Spain on November 20, 2011. We constructed several social networks defined by interactions (retweets, replies and mentions) among the Twitter users assigned to the major political parties. Also, we identify profiles that belong to political activists and profiles from traditional mass media to categorize the impact of their influence in Twitter. We observed important structural differences, which can be explained by the differences in the types of social campaigns run by these parties. Using this knowledge, we are working to define a framework for measuring the most important and influential political party members within this Twitter network. Possible next steps include conducting sentiment analysis of the tweets.

In another project we are examining the large scale protests that took place in Spain on October 15th, 2011 (Aragón et al., 2012). On this date, the Spanish 15-M movement called for a global demonstration to express and join the different local and national indignant-movement around the world. We are conducting a mixed methods analysis based on data collected from Twitter and 'traditional' media sources. We are performing comparisons of the different forms of political participation according to their geographic location, study the differences in emotional valence between Social Media and traditional mass media communication, classifying the different hashtags used around the 15O-demonstrations according to their content and valence, and studying the local and global networks that emerged around this social movement. Through our analysis we aim to provide deeper insights into the complex interplay between online and offline political participation.

3.1.4 Social Media vs. Traditional Media

In this study we will investigate the degree to which the relationship between political USENET discussions and the approval rates of US-presidents as shown in Gonzalez-Bailon, Banchs, Kaltenbrunner (2012) also exists for political news in traditional media.. Furthermore we will also analyze the possible existence of a relationship between traditional and social media using the same dataset.

4. CHALLENGES FOR INTERDISCIPLINARY RESEARCH

Research on the interplay of social behavior and new 'social media' can be greatly enriched by a multi-disciplinary approach that incorporates a range of different perspectives and methodologies. However, there are a number of challenges inherent in interdisciplinary research.

Theoretical frameworks differ greatly across disciplines, as do the methodological approaches to research. The lack of a shared frame of a reference can present a challenge to interdisciplinary research, as can the resulting lack of a shared vocabulary. For example, social scientists accustomed to utilizing qualitative research methods may find it difficult to interpret and understand results of complex modeling performed by computer scientists and mathematicians. At the same time, quantitative researchers may not fully appreciate how to draw meaningful conclusions based on interview and survey data.

Another challenge is that different terms can have very different meanings across disciplines, so discussion of, and agreement on, a common working vocabulary is important. For example, a robust sample in the context of conducting interviews might be 20, while for analysis of certain quantitative metrics of social network analysis it might be necessary to have access to the entire dataset (see introduction). Also a jump from a sample of 20 to 200 000 might imply no increase in the workload for the quantitative researcher (and only a little bit more of computer time).

Selection of publication outlets for papers emerging out of interdisciplinary research poses yet another challenge. Computer scientists and mathematicians often publish in proceedings of technical conferences. However, social sciences conferences rarely publish proceedings, and conference presentations are viewed as an opportunity to receive feedback on papers that are being prepared for publication in journals. Therefore, it can be challenging to reconcile these divergent disciplinary perspectives on publishing, and to publish work in venues that are valuable for the careers of all contributors.

Treatment of previous theory and literature also differs quite greatly across disciplines. In the social sciences, new research must be grounded very carefully in previous theoretical and empirical work,

and the theoretical mechanism(s) underlying a proposed relationship must be clearly explicated. In disciplines such as computer science, it is often sufficient to cite related work in a few paragraphs.

5. FUTURE DIRECTIONS

Moving forward, the goal of the SM team is to continue to develop mixed-methods, interdisciplinary research projects that can leverage the unique skill sets of all of our team members. We also aim to develop interdisciplinary collaborations with other institutions outside of Fundació Barcelona Media to pursue the following objectives:

5.1 Development of Tools to Enable Sociological Studies of Big Data

For this purpose we plan to develop (in intense collaboration between the two research lines) intelligent tools, which enable researchers from social sciences to explore big datasets such as collections of tweets or online discussions with several thousands of comments (possibly distributed among several platforms). The tools will automatically propose important users, tweets, topics, forum threads, etc. We later plan to extend these tools to enable their use by non-researchers, such as corporate users who want to gain deeper insights about big data related to their business, or individuals, or government agencies, or NGOs who want to obtain an overview of the topics discussed and the main actors in large online debates.

5.2 Design and Develop a Methodology for Social Media Experiments

We plan to design and develop a methodology for social media experiments, where users are engaged in a task unrelated to the research objective of the experiment. The aims of the experiments are twofold. First, we want to provide deeper insights into psychological and sociological aspects of human behavior. Second, we want to extract information about social connections and personal preferences valuable for marketing and customer service departments of companies.

5.3 Enterprise Social Media

With the increase in the implementation of enterprise social media solutions, there is a growing need to understand how and why workers are using these technologies, and how they might be improved. We plan to conduct mixed-methods research to gain an understanding of user behaviors and motivations with respect to this emerging set of tools.

5.4 Social Media for a Cause

In collaboration with an interdisciplinary group of researchers, our team has drafted a manifesto on social media for social innovation, which promotes the possibility of using social media as a platform to effectively support the processes of social innovation, overcoming its limitations of speed and scale to become an alternative to currently established institutional mechanisms (Kaletka et al., 2012). Using the manifesto as a starting point, we will develop field experiments that will explore how social media tools can be used both to develop innovative solutions to social challenges and to facilitate their adoption.

We hope to engage other research institutions to these objectives. If you are interested feel free to contact one of the authors of this document.

6. SELECTED REFERENCES

- Aragón, P., Gould, J., Kaltenbrunner, A., Kappler, K., Laniado, D., & Volkovich, Y. (2012), "Real counter power or a digital bubble? A case study of the 15-O Movement", Paper to be presented at the ISS'12.
- Bradley, M., & Lang, P. (1999), "Affective norms for english words (ANEW): Instruction manual and affective ratings," (Tech. Rep.). University of Florida: The Center for Research in Psychophysiology.
- Gómez, V., Kaltenbrunner, A., & López, V. (2008), "Statistical analysis of the social network and discussion threads in Slashdot", In WWW'08.
- Gómez V., Kappen H. J., Kaltenbrunner A. (2011), "Modeling the Structure and Evolution of Discussion Cascades", Hypertext'11.
- Gonzalez-Bailon S., Kaltenbrunner A., Banchs R. E. (2010), "The Structure of Political Discussion Networks: A Model for the Analysis of E-Deliberation", Journal of Information Technology, 25, 2010, pp 230-243.
- Gonzalez-Bailon, S., Banchs, R. E., & Kaltenbrunner, A. (2012), "Emotions, public opinion and U.S. presidential approval rates: A 5 year analysis of online political discussions", Human Communication Research, 38 (2), 121-143.
- Gould, J., Kaltenbrunner, A., Kappler, K., Laniado, D., Aragón, P., & Volkovich, Y. (2012), "Jointly they edit: Examining the impact of political identity and community attachment in Wikipedia", presented at INSNA Sunbelt'12.
- Hirsch, J. (2005), "An index to quantify an individual's scientific research output", Proceedings of the National Academy of Sciences of the United States of America, 102 (46), 16569.
- Jaimes, A., Lalmas, M., Volkovich, Y. (2011), "Report on the First International Workshop on Social Media Engagement (SoME'11)", SIGIR Forum 45(1) : 56-62
- Kaletka, C., Kappler, K., Pelka, B., Ruiz de Querol, R. (2012), "Challenges at the Intersection of Social Media and Social Innovation: a Manifesto", in Challenge Social Innovation. Springer. (Forthcoming 2012).
- Kaltenbrunner, A., Gómez, V., & López, V. (2007). Description and prediction of slashdot activity. LA-WEB '07.
- Kaltenbrunner, A., Gómez, V., Moghnieh, A., Meza, R., Blat, J., & López, V. (2008), "Homogeneous temporal activity patterns in a large online communication space", IADIS International Journal on WWW/INTERNET , 6 (1), 61-76.
- Kaltenbrunner, A., González, G., Volkovich, Y., Ruiz de Querol, R. (2011), "Comparative Analysis of Articulated and Behavioural Social Networks in a Social News Sharing Website", New Review of Hypermedia and Multimedia, 17 (3), 243-266.
- Kaltenbrunner A., Bondia E., Banchs R. E. (2009), "Analyzing and ranking the Spanish speaking MySpace community by their contributions in forums", WWW '09.
- Kaltenbrunner A., Gonzalez-Bailon S., Banchs R. E. (2009), "Communities on the Web: Mechanisms Underlying the Emergence of Online Discussion Networks", WebSci'09.
- Kappler, K., & Ruiz de Querol, R. (2010), "Redes conversacionales en la blogosfera: Un análisis cualitativo". Red de Cátedras Telefónica: Social Media. June 2010.
- Laniado, D., Tasso R., Volkovich Y., Kaltenbrunner A. (2011), "When the Wikipedians Talk: Network and Tree Structure of Wikipedia Discussion Pages", ICWSM'11.
- Lipovetsky, G., and Serroy, J. (2010), "La cultura-mundo. Respuesta a una sociedad desorientada", Barcelona, Anagrama.
- Pascual-Cid, V. & Kaltenbrunner, A. (2009), "Exploring Asynchronous Online Discussions through Hierarchical Visualisation", IV'09.
- Ruiz de Querol, R. & Kappler, K. (2010), "La emergencia de los 'social media': Retos y oportunidades", Cuaderno Red de Cátedras Telefónica: Social Media. January 2010.
- Volkovich Y. & Kaltenbrunner A. (2011), "Evaluation of Valuable User Generated Content on Social News Web Sites", WWW '11.
- Volkovich, Y., Scellato, S., Laniado, D., Mascolo, C., Kaltenbrunner, A. (2012), "The length of bridge ties: structural and geographic properties of online social interactions", ICWSM'12.