Making Sense of Microposts (#Microposts2016) Computational Social Science Track

Katrin Weller*
GESIS Leibniz Institute for the Social Sciences
Cologne, Germany
katrin.weller@gesis.org

Aba-Sah Dadzie*
Knowledge Media Institute, The Open University
Milton Keynes, UK
aba-sah.dadzie@open.ac.uk

Danica Radovanović*
Basic Internet Foundation / UNIK
Oslo, Norway
danica@basicinternet.org

ABSTRACT

For the second time, the #Microposts workshop features a track to highlight social science perspectives on micro communication structures in online environments. This paper introduces the *#Microposts2016 (Computational) Social Science Track*, which all contribute to connecting research methods and approaches in computer science and social science. By providing a forum for closer interaction between the two fields, the track is becoming a platform for interdisciplinary projects and new ideas to combine different methodologies and theories. For this year's special track we see the trend of relating Microposts to external demographics or survey data as a way to better understand Microposts in their broader contexts.

CCS Concepts

•Information systems → World Wide Web; •Applied computing → *Enterprise computing;*

Keywords

Microposts, Social Science, Web Science, Computational Social Science, Internet science, Internet research, social media, user-generated content, online communication

1. BACKGROUND

Accelerated development of Information-Communication Technologies (ICT) has a profound impact on socio-cultural relationships and processes; this presents a challenge for researchers from multiple disciplines and backgrounds. In an interconnected world of information, different forms of communication and social dynamics are formed, referring to the socio-technological processes that take place online. New technologies shape information, communication and collaboration dynamics in different environments while contributing to persistent interdisciplinarity.

In academia there is increasing tendency toward interdisciplinary work between computer science and other technical sciences [11, 21] and the social sciences [8], increasingly referred to as *Computational Social Science*. Different social media tools provide an expressive medium for sharing with others – both acquaintances and the general public – feelings, needs, current status, or simple statements [16]. This provides solid ground for forming *phatic expressions*, which we also refer to as *Microposts*. A Micropost constitutes a small, brief message, theme or a single thought, quick and easy to publish, and that, posted from a variety of platforms and by very large numbers of individuals with as many viewpoints and interests, collectively provide a rich source of information and opinion about a range of topics.

The first micro-sized social media posts were exchanged using text; while the term Micropost has evolved with advances in technology the term now rarely needs definition. Text is still a popular means of expression, e.g., in a tweet, status update or a comment in a news feed. The term is however now used even outside the workshop to refer not just to brief text posts but as a means of sharing also other multimedia information – in photos, streaming and pre-recorded video and audio, with the caveat that the post remains small. Popular options include as an Instagram photo (with a hashtag), a Snapchat quick video message or a three second Instagram Boomerang video.

The brevity of Microposts makes them a convenient, low-cost means for sharing information in the moment and on the go, from a variety of personal devices, from the myriad new apps and portals built each year. The utility of Microposts is seen in that where direct access to the Internet is not available, for instance in remote, off-grid or underserved areas, apps exist to allow the next most convenient communication means, including mobile networks and even radio, to be used to transport Microposts. On the other hand, where traditional communication channels and media outlets are overwhelmed or cut off, typically during disasters and emergencies, or during uprisings and mass demonstrations, social media access via Internet access may provide a lifeline or "connector" to emergency and information services, family and friends [4, 10, 14]. The Ushahidi platform, for instance, has been used for crowdsourced translation and relay of information in a number of crises, using SMS with social media, notable particularly following the 2010 Haiti earthquake. *LINE*, one of the most popular social media tools in Japan, was born in the wake of the 2011 earthquake to aid communication [9, 12].

*All authors made equal contributions

Published as part of the #Microposts2016 Workshop proceedings, available online as CEUR Vol-1691 (http://ceur-ws.org/Vol-1691)

#Microposts2016, Apr 11th, 2016, Montréal, Canada.

2. TOWARD MORE INTERDISCIPLINARY APPROACHES FOR MAKING SENSE OF MICROPOSTS

The #Microposts workshop [1] aims to provide an interdisciplinary forum for Computer Science researchers and practitioners to collaborate with specialists from other fields, including Information, Web, Social and Computational Sciences, to discuss the generation and analysis of Micropost data and promote effective application of its knowledge content in different contexts and situations, including emergency response, crowd and event tracking, mass communication and marketing, opinion mining and sentiment analysis. The track also looks at research examining ethics, legal and privacy issues in the analysis and reuse of data which, although typically published in public or semi-public fora, often includes sensitive, identifiable information about individuals. The social and communication dynamics that result from the use of Micropost-based services are sometimes impacted by cultural, generational and regional differences [9], often seen in data journalism and citizen reporting of civil action and politics as in the use of Microposts in the Arab Spring, and in the use of the forum for education and citizen empowerment.

Making a significant contribution to today's "big data", the research and analytical activity necessary in making sense of Microposts rely in large part on techniques and tools for large-scale information extraction and knowledge acquisition, in order to handle the very high rate at which Microposts are published, and increasingly using multi-media. However, automated analysis alone cannot decipher the sublanguage necessary to capture complete messages in such small packets, or the nuance often used to aid encoding of these very small snippets of information, and in public or semi-public fora, load message content. Key to unpacking Micropost content is the human in the analysis loop, to identify and interpret nuance that cannot be picked up by automated methods, to better understand why end users employ this medium and in what circumstances it may be seen as a preferred means of communication. In addition to large-scale computational analysis it may also be necessary to look more closely at selected cases – and combine this with other quantitative and, importantly, qualitative approaches for analysis. It is also often necessary to combine different sources of data which complement each other. Further, from a social science perspective, approaches that help to relate online communities to wider online and offline populations are particularly interesting.

3. THE #MICROPOSTS2016 (COMPUTATIONAL) SOCIAL SCIENCE TRACK

While the #Microposts workshops have always sought to bring together researchers with different disciplinary backgrounds, the need to place an even stronger focus on contributions from the broader field of social sciences became apparent. #Microposts2015 thus featured the first dedicated Social Science track in the #Microposts series, and the programme committee was broadened to assemble even more social scientists and researchers from related disciplines engaged in the study of Microposts and online communication. Based on the success in its first year, which saw two accepted submissions, the Social Science track was continued in 2016. One modification was to the track name, to *(Computational) Social Science*. The new name better reflects the intent to include contributions from the social sciences as additional disciplines on the one hand, and also contributions that clearly apply multidisciplinary approaches and bridge the social sciences and computer science on the other. Submissions accepted in 2015 demonstrated very well how cross-boundary submissions enrich the overall workshop scope, and those in 2016 reinforced the value in such work to making sense of Microposts.

3.1 Track Sponsor: GESIS

For the second time *GESIS*, the Leibniz Institute for the Social Sciences [2] is our sponsor for this special track. As a research infrastructural organisation for the social sciences, GESIS offers support and services for different phases in the research process. The established services for supporting traditional social science research have recently been complemented with a GESIS department in Computational Social Science, which focuses on algorithms and theories for studying social phenomena based on Web data. This work is in many ways closely related to research on making sense of Microposts. Support from GESIS confirms the relevance of the special track; by helping to raise the profile of the track, the Microposts workshop continues to attract participation from authors and attendees who would not normally participate in a conference targeted predominantly at Computer Science.

3.2 Topics of Interest

In 2016 three submissions were received for the (Computational) Social Science track. A fourth submission submitted to the main track was also included in the review process for this track as it clearly crossed interdisciplinary boundaries, with two reviewers from each track's committee assigned to it. Out of these submissions, one was accepted for full paper presentation and two as poster presentations.

All three accepted submissions contribute valuable perspectives to the understanding of Microposts, using different disciplinary approaches to bridge the gap between theory and practice. However, after a second review round a decision was made not to award the best paper prize (Computational) Social Science track, as the standard for doing so was not fully reached. Honourable mention however goes to the paper *Comparing Social Media and Traditional Surveys Around the Boston Marathon Bombing* [6]. The authors, Cody Buntain, Erin McGrath, Jennifer Golbeck and Gary LaFree, address one of the key challenges in bridging traditional social science methods with new computational approaches. In a case study about the bombings during the Boston Marathon in April 2013 Buntain *et al.,* combine survey data (from a panel survey and an experimental setup) with digital trace data from social media. They thus placed Microposts, on Twitter, in the broader context of opinions examined through survey data. Of value especially to computer scientists is how this paper shows capability to gain insight into the value of survey data, as a way to enrich online communication data. Social scientists on the other hand will find information that helps them to judge the costs (temporal and financial) of collecting new types of data in comparison to standard approaches employed in the field. As the authors conclude, applied together, survey and social media data can show a more complete picture of public

opinion.

Both poster presentations in this track present studies on a specific group of Twitter users, and in both cases refer to these users as some sort of elite. The poster *Studying the Role of Elites in U.S. Political Twitter Debates* by Sebastian Stier [20] concentrates on "elites" that dominate political conversations: politicians, other government representatives and news media representatives. Alex Jeongwoo Oh & Pramuan Bunkanwanicha in *CEOs on Twitter* [13] focus on "elite" users in different business sectors. They analyse the tweets of executives of major firms to detect differences in style and practice and relate them to demographics on the one hand and economic performance of the respective companies on the other.

Studying Microposts based on selected groups of users instead of specific topics (e.g. as expressed through hashtags) enables different forms for understanding the ongoing communication. Studying "elite" users such as politicians and CEOs has another direct advantage: once an account is verified as belonging to a certain elite user it typically also becomes possible to work with additional information about this person, such as demographics like age and location. Even more detailed biographical information is usually available from verified external sources for such elites. For instance, information about politicians may include political programmes, information on CEOs may include wages and different positions held during their career. On the other hand, when studying regular users (non-public figures) one is usually bound to the information these users provide on public (Web) platforms when filling in user profiles. Such profile information is often limited, incomplete, ambiguous or even deliberately misleading. On Twitter, for example, user profiles do not include information on gender or age and there is no standardised form for entering geo-location. This has inspired research on methods to automatically infer demographic information or other user attributes from online activities or user networks [5]. This is not without its challenges, also because there is still ongoing debate in the Internet research ethics community about how much information may be collected or inferred about regular Web users, who may not have explicitly given "informed consent" on how and if their data may be used for research or other purposes [3, 22, 23]. Focusing on elite users and their public roles through official accounts on Web platforms reduces much of the ethical concerns in social media research, as it becomes possible to work with publicly available, official information from external, verified sources. Further, Microposts from such accounts are usually official statements rather than private statements or opinion. This has enabled, therefore, publication of some of the underlying research data [19] utilised in the study of political elites by Stier [20] in #Microposts2016. This serves to provide both a valuable contribution to other research on this and related topics, and is in line with public policy toward increasingly open data, to enrich research and lead to improvements in use of the collective intelligence gathered through media such as Microposts.

We observe that some topics from the 2015 workshop were picked up again but from new perspectives, applying new foci and methods. Political communication clearly is a persistently popular topic in this field (though the focus on elite users as in the 2016 case study is possibly a niche area). Dealing with crises through social media and understanding the role of Micropost messages in dealing with unexpected events is another popular topic, having seen practical use in crises, environmental disasters and emergencies. For example, #Microposts2015's Social Science track also featured a case study that investigated communication around a crisis event – the 2015 Charlie Hebdo Shooting in France [7]. The 2016 paper on the Boston Marathon bombing [6] does not delve as deeply into understanding the polarisation of different groups reacting to the terrorist attacks; it rather reminds us that Microposts can be further enriched with additional data. As we see in this workshop, from a social science perspective, this may be opinions polled via surveys as well as demographics collected for specific user groups.

3.3 Track Committee

The proposal to include again the (Computational) Social Science track was strengthened by the programme committee, who work in the Social Sciences, Computer Science and in Business Administration in Higher Education in Europe, the Middle East and North and South America. Together, our committee comprise a wealth of research that was exhibited in the the informative reviews for all submissions, whether accepted or not. This feedback resulted in final papers that should provide a good contribution to the literature on Making Sense of Microposts and the broader fields of related research.

Acknowledgments

Katrin Weller is senior researcher and team leader at the Department of Computational Social Sciences, GESIS Leibniz Institute for the Social Sciences in Cologne, Germany. Aba-Sah Dadzie is researcher in Data Science at KMi, the Open University, working on the EU project EDSA (no. 643937). Danica Radovanović is a senior visiting researcher at UNIK, Kjeller, Norway and a Digital Equality Advisor at the Basic Internet Foundation headquartered in Oslo, Norway.

4. REFERENCES

[1] #Microposts2016 website. http://microposts2016.seas.upenn.edu.

[2] GESIS. http://www.gesis.org.

[3] S. Alim. An initial exploration of ethical research practices regarding automated data extraction from online social media user profiles. *First Monday*, 19(7), 2014.

[4] N. Antoniou and M. Ciaramicoli. Social media in disaster cycle – useful tools or mass distraction? In *64th International Astronautical Congress*, 2013.

[5] V. Beretta, D. Maccagnola, T. Cribbin, and E. Messina. An interactive method for inferring demographic attributes in twitter. In *Proc., 26th ACM Conference on Hypertext & Social Media (HT '15)*, pages 113–122, 2015.

[6] C. Buntain, E. McGrath, J. Golbeck, and G. LaFree. Comparing social media and traditional surveys around the boston marathon bombing. In Preoţiuc-Pietro et al. [15], pages 34–41.

[7] F. Giglietto and Y. Lee. To Be or Not to Be Charlie: Twitter hashtags as a discourse and counter-discourse in the aftermath of the 2015 Charlie Hebdo shooting in France. In Rowe et al. [18], pages 33–37.

[8] R. H. Heiberger and J. R. Riebling. Installing computational social science: Facing the challenges of new information and communication technologies in social science. *Methodological Innovations*, 9, 2016.

[9] F. E. Jandt. Dimensions of culture. In *An Introduction to Intercultural Communication Identities in a Global Community*, pages 160–190. SAGE Publications, Inc, 8th edition, 2016.

[10] J.-Y. Jung. Social media use and goals after the great east japan earthquake. *First Monday*, 17(8), 2012.

[11] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne. Computational social science. *Science*, 323(5915):721–723, 2009.

[12] H. McCracken. How Japan's LINE app became a culture-changing, revenue-generating phenomenon. *Fast Company Magazine*, Mar 2015.

[13] A. J. Oh and P. Bunkanwanicha. CEOs on Twitter. In Preoţiuc-Pietro et al. [15], pages 46–48.

[14] B. D. M. Peary, R. Shaw, and Y. Takeuchi. Utilization of social media in the East Japan earthquake and tsunami and its effectiveness. *Journal of Natural Disaster Science*, 34(1):3–18, 2012.

[15] D. Preoţiuc-Pietro, D. Radovanović, A. E. Cano-Basave, K. Weller, and A.-S. Dadzie, editors. *Proceedings, 6th Workshop on Making Sense of Microposts (#Microposts2016): Big things come in small packages, Montréal, Canada 11th of Apr 2016*, 2016.

[16] D. Radovanović and M. Ragnedda. Small talk in the digital age: Making sense of phatic posts. In Rowe et al. [17], pages 10–13.

[17] M. Rowe, M. Stankovic, and A.-S. Dadzie, editors. *Proceedings, 2nd Workshop on Making Sense of Microposts (#MSM2012): Big things come in small packages, Lyon, France, 16 April 2012*, April 2012.

[18] M. Rowe, M. Stankovic, and A.-S. Dadzie, editors. *Proceedings, 5th Workshop on Making Sense of Microposts (#Microposts2015): Big things come in small packages, Florence, Italy, 18th of May 2015*, May 2015.

[19] S. Stier. Elite actors in the U.S. political Twittersphere. GESIS datorium. http://dx.doi.org/10.7802/1178, 2016.

[20] S. Stier. Studying the role of elites in U.S. political twitter debates. In Preoţiuc-Pietro et al. [15], pages 43–45.

[21] M. Strohmaier and C. Wagner. Computational social science for the world wide web. *IEEE Intelligent Systems*, 29(5):84–88, Sept 2014.

[22] K. Weller and K. E. Kinder-Kurlanda. I love thinking about ethics: Perspectives on ethics in social media research. In *Selected Papers of Internet Research (SPIR), Proc., ir15 – Boundaries and Intersections*, 2014.

[23] M. Zimmer. "but the data is already public": on the ethics of research in facebook. *Ethics and Information Technology*, 12(4):313–325, 2010.