Hall, Margaret et al.

Conference Paper

Following User Pathways: Cross Platform and Mixed Methods Analysis in Social Media Studies

CHI EA '16 Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems

Suggested Citation: Hall, Margaret et al. (2016) : Following User Pathways: Cross Platform and Mixed Methods Analysis in Social Media Studies, CHI EA '16 Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, ISBN 978-1-4503-4082-3, ACM, New York City, pp. 3400-3407, https://doi.org/10.1145/2851581.2856500

This Version is available at:
http://hdl.handle.net/11108/282

Following User Pathways: Cross Platform and Mixed Methods Analysis in Social Media Studies

Margeret Hall
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
hall@kit.edu

Athanasios Mazarakis
Isabella Peters
Kiel University / ZBW
Duesternbrooker Weg 120
24105 Kiel, Germany
a.mazarakis@zbw.eu
ipe@informatik.uni-kiel.de

Martin Chorley
Cardiff University
Queens Buildings,
Cardiff, CF24 3AA, UK
chorleymj@cardiff.ac.uk

Simon Caton
National College of Ireland
Mayor Street, IFSC
Dublin 1, Ireland
simon.caton@ncirl.ie

Jens-Erik Mai
University of Copenhagen
Birketinget 6,
2300 Copenhagen, Denmark
jemai@hum.ku.dk

Markus Strohmaier
GESIS
Unter Sachsenhausen 6-8
50667 Cologne, Germany
markus.strohmaier@gesis.org

CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2856500

Abstract
Social media and the resulting tidal wave of available data have changed the ways and methods researchers analyze communities at scale. But the full potential for social scientists (and others) is not yet achieved. Despite the popularity of social media analysis in the past decade, few researchers invest in cross-platform analyses. This is a major oversight as 42% of Online Social Media users have multiple social media accounts. Missing are the models and tools necessary to undertake analysis at scale across multiple platforms. Especially promising in support of cross-platform analysis is the mixed method approach (e.g. qualitative and quantitative methods) in order to better understand how users and society interacts online. The workshop '*Following User Pathways*' brings together a community of researchers and professionals to address methodological, analytical, conceptual, and technological challenges and opportunities of cross-platform, mixed method analysis in social media ecosystems.

Author Keywords
Social Media; Mixed Method Analysis; Research Ethics; Cross Platform; Data Curation

ACM Classification Keywords
H.1.2 User/Machine Systems J.4; Social and Behavioral Sciences.

Mission Statement

This workshop aims to consolidate diverse research practices and methodologies of social media analysis into a more structured and unified vision for user experience, HCI research, and an overarching understanding and observation point of digital studies. Under-addressed by researchers (though not digital advertisers or data mining processes), '*Following User Pathways*' promises a more complete concept of how to braid differing aspects together for the benefit as opposed to exploitation of social media users, as well as to the benefit of (social) researchers.

'*Following User Pathways*' recognizes that only with a mixture of platforms can researchers really disentangle aspects of user experience, engagement and dependencies of and upon social platforms.

Background

A new approach in the area of empirical social research is found in computational social science [7,14], where the interaction of technology, online communities, and individuals' perception within are investigated at a previously unmanaged scale [4,20]. In spite of the prevalence and availability of Application Programming Interfaces (APIs), as well as commercial data mining packages, and an almost endless supply of papers and studies that focus on specific platforms [1,2,5,6,11,15,17], cross platform analyses are lacking in the social media analysis and broader computation social science corpus. This is an oversight, in an age where nearly half of social media users have multiple social media profiles [9]. The current maturity level of social media and social network research is lower than its potential due to this oversight. A level of yet-unknown research bias also exists due to literature's current concentration on single-platform analyses [21]. In order to facilitate more realistic analyses, social models, and theories, researchers need to approach social media as a holistic ecosystem: the scientific community must map user pathways to match users' activities.

First work in the comparison of the same user or phenomenon on different social media platforms, or path mapping, is being attempted [8,16,22], but are still limited in scope. It has, for example, been established that individuals' sentiment valence and conversation styles differ across platforms [8,16], and that unrelated posts on different social media platforms can predict return on advertising investments [22]. But, because users cannot be tracked across platforms the available tools do not match the research need [13]. Missing is a better concept of how to plug all these things together for the benefit as opposed to exploitation of users as well as (social) researchers.

Whilst it is apparent that the technical functionalities exist for capturing multiple streams of data, currently few scientific papers or services undertake this in a way that is easily replicable. Few platforms or packages exist which pull social media data from multiple platforms in order to facilitate mixed method analyses. Nor do many works exist focusing on platforms with smaller market shares. Of particular interest are mapping the path to include alternative platforms like enterprise social networks, professional networks like LinkedIn, or platforms aiming at particular target groups (e.g., researchers on Mendeley or ResearchGate). Visually-based social networks like YouTube, Snapchat, and Pinterest are also of interest considering their growing user bases [12].

Several issues must be considered. Activities performed on platforms often cannot be compared since they are driven by different conceptual frameworks and motivations (e.g. retweeting on Twitter vs. liking on Facebook). For example, differences in motivation for such activities has been neither qualitatively nor theoretically addressed. Likewise unaddressed is expression of self across platforms – i.e., what motivation does a user have behind posting a similar, different or even same message on different platforms? Here, research on the linking of conceptual and analytical models is needed.

The technical challenges also differ significantly from receiving a continuous stream of data (i.e., tweets) vs. Facebook's paginated results. The latter incites large numbers of API calls, which are not limitless. Restricted word counts present an interesting validation challenge considering the overall small *n* for posts outside of Facebook [3]. A similar small-*n* challenge exists for low volume users across platforms. Moreover, the method of data curation is not without its ambivalence. For example Twitter data curation tends to be forward-facing; accessing future Tweets that fulfil a specific set of attributes starting at a given time point [15,19]. Facebook is retrospective; given a Facebook entity (e.g. a person, or page) researchers access current and historical posts, profile, likes etc. From the perspective of analyzing social data, this subtle difference significantly alters the effort and planning needed to curate a data set and the implicit biases associated with the method [10,21].

Also, ethical considerations for research have to be taken into account [11]. A significantly higher volume of data, with a yet-unaddressed level of granularity is afforded by cross-platform analysis. Users may be unaware of the implications of cross-platform analyses. The privacy of social media users in cross-platform approaches should be of paramount concern.

Finally, mixed method analysis is inherently iterative and interdisciplinary. Whilst approaches from computer science and computational social science are becoming more prevalent, the question of research methodology is often a poignant discussion point and challenge that cannot be overlooked; computer and social scientists leverage diverse and often non-overlapping research methodologies. Therefore, cross platform and mixed method analyses need to accommodate a vast array of (interdisciplinary) methodological approaches.

Irrespective of methodology, an important feature of cross-platform and mixed method analysis is the ability to view a community at a variety of resolutions; starting from an individual micro layer, and progressively zooming out via ego-centric networks, social groups, communities, and demographic (sub)groups, up to the macro layer: community [5]. This ability is of significant importance for understanding a community as a whole, as well as following the complete user path [5]. Also, including temporal information in analyses of user activities across platforms (e.g., Twitter is always visited after Facebook) can enhance understanding of how users navigate the information space, process information, and make use of platforms, and, thus, shed light on why particular platforms are used and for what reasons. Key contribution differences are the observation viewpoint and elicitation of points of reference by analyzing multiple platforms [5]. Whilst the scientific value of single platform studies is

significant, their isolated investigations only give us insights into well-grounded research processes rather than assisting in the construction of a general approach.

Goals and Deliverables

This workshop will bring together the diverse community of computational social science and web science researchers who work across platforms and with mixed methods. This is a foundational meeting for the establishment of a unified vision and structured approach for multi-modal and mixed method social media research. We specifically want to address the following points:

1. How can a complete social media path be mapped, and what does a complete representation look like?

This overarching question looks for framing and conceptual modelling of complete use pathways. Especially the qualitative aspects of user motivation and needs, and the quantitative aspect of instantiation design are captured here.

2. Is the value of a complete path higher than the amount of personal data required to map it with respect to data privacy?
3. What are the ethical parameters of path mapping to avoid exploitative conduct?

Necessary to note is that ethical data curation follows the Belmont Principles and/or the guidelines of the Association of Internet researchers [18]. The study design and curation must be reasonable, non-exploitative, and balance data extraction with benefit to society. These questions address the broad ethical issues in internet and cross-platform research.

4. What technical affordances need to be implemented to develop a single framework?

Few instantiations exist that support cross-platform data extraction. Even fewer exist that support mixed-method analyses. This question looks for contributions on integrated and automated cross-platform and mixed-method analyses.

5. What challenges to holistic curation exist?

It is expected that the observation lenses across platforms and with the differing methods capture differing structural, content, and temporal aspects. As such, the data must be reconciled to support a holistic analysis. Theoretical and empirical contributions addressing these challenges are envisioned with this question.

Selected workshop papers will be invited to submit their work for a special issue of *Interacting with Computers* (tbc) or another journal. A secondary goal of the workshop is continuing the dialogue with a further event, to be co-located at either 2016 ACM WebScience or the Association of Internet Researchers annual conference.

Organizers

Margeret Hall is a Senior Researcher at the Karlsruhe Service Research Institute of the Karlsruhe Institute of Technology. Her concentration is in the area of computational social science, particularly the detection

and public sourcing of social indicators from digital communities.

Athanasios Mazarakis is a postdoc of Web Science at CAU Kiel University and at ZBW – German National Library of Economics – Leibniz Information Centre for Economics. He works interdisciplinary in Computer Science, Psychology and Economics. His main expertise is about using non-monetary incentives and gamification to stimulate engagement in social media tools and apps.

Isabella Peters is Professor of Web Science at ZBW – German National Library of Economics – Leibniz Information Centre for Economics and also at CAU Kiel University (joint appointment). Her primary expertise is on how user-generated content (especially social tagging and folksonomies) from social media-platforms can be used for knowledge representation and information retrieval.

Martin Chorley is a Lecturer in the school of Computer Science & Informatics at Cardiff University. His research interests are focused around the interaction of people and systems and include social networks, social media and mobile computing, as well as computational journalism.

Simon Caton is a Lecturer of Data Analytics at the National College of Ireland. He is an active researcher in the computational aspects of analyzing multiple social media platforms simultaneously, as well as the facilitating parallel and distributed computing methods.

Jens-Erik Mai is professor at the University of Copenhagen, Royal School of Library and Information Science in Denmark. Jens-Erik studies basic questions about the nature of information phenomena in contemporary society; he is concerned with state of privacy and surveillance given new digital media, with classification given the pluralistic nature of meaning and society, and with information and its quality given its pragmatic nature.

Markus Strohmaier is Professor of Web-Science at the Faculty of Computer Science at University of Koblenz-Landau, Scientific Director of the Computational Social Science department at GESIS - the Leibniz Institute for the Social Sciences, and a Distinguished Visiting Scholar at Stanford University's Media-X program (since 2010). He has been a visiting scientist/professor at Stanford University during the 2011/12 academic year, at XEROX Parc (2009, 2010-2011) and at RWTH Aachen (2009). He received his PhD from the Faculty of Computer Science at Graz University of Technology in 2004.

References

[1] Allen, S.M., Chorley, M., Colombo, G., et al. Exploiting user interest similarity and social links for micro-blog forwarding in mobile opportunistic networks. *Pervasive and Mobile Computing 11*, (2014), 106–131.

[2] Böcking, B., Hall, M., and Schneider, J. Event Prediction With Learning Algorithms—A Study of Events Surrounding the Egyptian Revolution of 2011 on the Basis of Micro Blog Data. *Policy & Internet 7*, 2 (2015), 159–184.

[3] Braga-Neto, U.M. and Dougherty, E.R. Is cross-validation valid for small-sample microarray classification? *Bioinformatics 20*, 3 (2004), 374–80.

[4] Burrows, R. and Savage, M. After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data & Society 1*, 1

(2014), 2053951714540280.

[5] Caton, S., Hall, M., and Weinhardt, C. How do politicians use Facebook? An applied Social Observatory. *Big Data & Society 2*, 2 (2015), 1–18.

[6] Chorley, M.J., Whitaker, R.M., and Allen, S.M. Personality and location-based social networks. *Computers in Human Behavior 46*, (2015), 45–56.

[7] Cioffi-Revilla, C. Computational social science. *Computational Statistics 2*, 3 (2010), 259–271.

[8] Davenport, S.W., Bergman, S.M., Bergman, J.Z., and Fearrington, M.E. Twitter versus Facebook: Exploring the role of narcissism in the motives and usage of different social media platforms. *Computers in Human Behavior 32*, (2014), 212–220.

[9] Duggan, M., Ellison, N., Lampe, C., Lenhart, A., and Madden, M. *Pew Social Meida Report 2015.* 2014.

[10] González-Bailón, S., Wang, N., Rivero, A., and Borge-Holthoefer, J. Assessing the bias in samples of large online networks. *Social Networks 38*, January (2014), 16–27.

[11] Gosling, S.D. and Mason, W. Internet Research in Psychology. *Annual Review of Psychology 66*, 1 (2015), 877–902.

[12] Hall, M. *Detecting Well-being in Digital Communities: An Interdisciplinary Engineering Approach for its Indicators.* KIT Scientific Press, Karlsruhe, Germany, 2015.

[13] Kivelä, A. and Lyytinen, O. Topic Map Aided Publishing – A Case Study of Assembly Media Archive. *STeP 2004 - The 11th Finnish Artificial Intelligence Conference Proceedings*, (2004).

[14] Lazer, D., Brewer, D., Christakis, N., Fowler, J., and King, G. Life in the network: the coming age of computational social. *Science 323*, 5915 (2009), 721–723.

[15] Lemke, S., Mazarakis, A., and Peters, I. Understanding Scientific Conference Tweets. *Proceedings of the 17th General Online Research Conference (GOR 2015)*, German Society for Online Research (DGOF) (2015), 52–53.

[16] Lin, H. and Qiu, L. Two sites, two voices: Linguistic differences between facebook status updates and tweets. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8024 LNCS*, (2013), 432–440.

[17] Lindner, A., Niemeyer, C., Caton, S., Hall, M., Niemeyer, C., and Caton, S. BeWell: A Sentiment Aggregator for Proactive Community Management. *CHI'15 Extended Abstracts*, (2015).

[18] Markham, A. and Buchanan, E. *Ethical Decision-Making and Internet Research Recommendations from the AoIR Ethics Working Committee.* Association of Internet Researchers, Chicago, IL, 2012.

[19] Mazarakis, A. and Peters, I. Tweets and Scientific Conferences: The Use Case of the Science 2.0 Conference. *Proceedings of the 2nd European Conference on Social Media (ECSM 2015)*, Academic Conferences and Publishing International Limited (2015), 323–331.

[20] Ruppert, E. Rethinking empirical social sciences. *Dialogues in Human Geography 3*, 3 (2013), 268–273.

[21] Ruths, D. and Pfeffer, J. Social media for large studies of behavior. *Science 346*, 6213 (2014), 1063–1064.

[22] Schacht, J., Hall, M., and Chorley, M. Tweet if you will – the real question is, who do you influence? *WebSci 15*, (2015).

Pre-workshop Plans

In order to appeal to the broader CHI audience (thereby integrating diverse methods and less-commonly analyzed platforms), 'Following User Pathways' will recruit from a broad spectrum of disciplines. This is supported by the interdisciplinary program committee, including computer and web science, communications, political science, and psychology. The Program Committee will publicize the workshop via personal outreach, mailing lists and other appropriate venues.

'Following User Pathways' is building an active social media and platform-based presence, expecting approximately 20 papers as a part of the proceedings.

Workshop Structure

Part I. Morning Session – Invited talks, paper presentation and discussion

In this session, participants will present and discuss their work in 20 minute presentations. Question and answer sessions will be a 10-minute panel at the end of each group, generating an interactive review of the current research.

Participants will be placed into conceptual themes by the organizers around the open questions listed in 'Goals and Deliverables.'

Part II. Afternoon Session – Group brainstorming and open discussion

The afternoon session will include brainstorming and discussion of the 'hot topics' that emerged in the discussions in the morning, as well as planning the post-workshop activities. This includes submission information for the Special Issue.

The session will end by collating the workshop outcome to a research agenda, to be posted on the workshop's webpage. All other media associated with the day will be made available on the workshop's websites.

Post-workshop Plans

Workshop papers will be invited to submit their work into a special issue. In order to support the stated goal of a structured approach to user path mapping, we plan to submit a second workshop proposal to ACM WebSci 2016 or the AoIR, pushing the conversation forward.

Call for Participation

Social media and the resulting tidal wave of available data have changed the ways and methods researchers analyze communities at scale. But the full potential for science is not yet achieved. Despite the popularity of social media analysis in the past decade, few researchers invest in cross-platform analyses due to various reasons, e.g. unknown methods and tools, supposed difficult analysis and others.

"**Following user pathways: Using cross platform and mixed methods analysis in social media studies**" co-located with **CHI 2016** brings together a community of researchers and professionals to address methodological, analytical, conceptual, and technological challenges and opportunities of mapping

user across platforms with mixed method analysis in social media ecosystems.

Topics of Interest:

- Motivational analysis across different Social Media platforms
- Multi-dimensional Representations of Person, Event and Society on Social Media
- Discrepancies in Representation of Events Across Platforms
- Barriers to Multi-platform and Mixed Methods analysis
- Ethical, Legal, and Social Implications, especially privacy issues
- Technical and Implementation Aspects of Multi-platform analysis
- Experiences of interdisciplinary multi-platform projects
- Mixed-method Approaches to Social Media Path Mapping
- Addressing Bias in Social Media Studies

Submission:
Interested authors should submit a 3-5 page paper in the CHI extended abstract format to hall@kit.edu. The organizing committee will review submissions and select based on relevance, quality, and diversity of inputs. Selected workshop papers will be invited to submit their work for a special issue of *Interacting with Computers* or another journal. At least one author of each accepted paper needs to register for the workshop and for one or more days of the conference.

Important Dates:
December 21, 2015: Early Bird Submissions
January 10, 2016: Submission deadline
February 7, 2016: Notification of acceptance
May 7-8 (tbd), 2016: Workshop

Organizing committee:
Margaret Hall, Athanasios Mazarakis, Isabella Peters, Martin Chorley, Jens-Erik Mai, Simon Caton, Markus Strohmaier

Questions? Find more information and follow the conversation at
www.facebook.com/followinguserpathways &
http://www.ksri.kit.edu/1516.php