Guidelines for Social Media Mining for Innovation Purposes

Experiences and Recommendations from Literature and Practice

Robert Eckhoff, Mark Markus,
Markus Lassnig, and Sandra Schön
Innovation Lab
Salzburg Research Forschungsgesellschaft
Salzburg, Austria
markus.lassnig@salzburgresearch.at

Abstract— **Social media in general and user-generated content are seen as potential sources for trend detection and innovation. Nevertheless, there are a quite variety of approaches and experiences. Within this contribution, we want to give an overview where we collect and discuss existing recommendations from literature and more. Building on lessons learned from literature and an expert discussion, we present guidelines for social media mining for innovation purposes. These guidelines especially build on experiences with an approach called innovation signals, which combines social media mining technology with an interpretative methodology. We explain guidelines, as "No tool will do your work automatically" or "Narrow your search", and discuss our methodology of guidelines' development as well as the results critically.**

Keywords-innovation; social media monitoring; social media mining; guideline

I. INTRODUCTION

Social Media and its user-generated content is a valuable source for management tools and analysis that try to gain market and consumer insights. Weblogs, social networks or microblogging give insights in customers' wishes, the image of a brand or new trends. Social media monitoring is also important to detect, select and analyze signs or facts for (future) innovation. In our contribution, we will give an overview about different approaches how social media mining is applied for innovation issues and provide insights into current research within the topic. The aim of our contribution is to deliver guidelines for researchers (and also practitioners) when they consider or plan to use social media mining for innovation purposes. Therefore, we will first introduce social media mining for innovation before we will present research question and research design as well as the guidelines we developed. Finally, we will discuss our results.

The developed guidelines should be interesting and helpful for consultants as well as researchers dealing with social media mining for innovation purposes.

II. SOCIAL MEDIA MINING FOR INNOVATION

A. Diverse uses of social media content and the idea of social media mining for innovation purposes

Social media are Web tools and services that allow to communicate, to collaborate, and to share. For example, social networks, discussion forums, Wikis, Weblogs or mailing lists are such applications. Within social media customers, colleagues, experts and others discuss brands, products and services, or related topics and issues. Therefore, social media is not only a way to share and discuss online, but also a good source of information for research and strategic planning. Marketing and PR departments of firms are interested in real-time monitoring of the Web to be able to react in time with a focus on customer satisfaction, customer relationship management, public relations or measurement of actions. Several tools help to monitor interactions within the web and send alerts for special keywords (e.g., the brand's name) [1]. Sometimes active interaction within the social web [2] is called "social media listening". In contrast to the in time monitoring of the social web the usage of social web content for mining purposes uses social media as source to find developments, new topics or interesting discussions. This approach is not only used for innovation purposes, but a quite common idea within innovation research [3]: Social web content is a cheap, non-reactive and authentic; therefore, it is a broad source for diverse methodologies and approaches to get insights or ideas for ongoing trends and future developments.

B. Approaches and Tools

There are several ideas what might be found related to innovation by social media mining [4]: Signs for potential innovations can be diffuse, but "weak signals" might have the potential of future impact on innovation [5]. For example Twitter messages are used as a source to detect weak signals for events [6]. Also, comprehensive overviews for tools that might be used for such weak signals detection [3] or social media monitoring in general [7] are available. The purpose of social media mining might also be the detection of ongoing developments or existing open innovations - These are innovations developed outside the enterprise, for example by customers [8] [9]. Mining for innovation also includes the detection and clarification of new trends, for example the

 ISBN: 978-1-61208-386-5

speed of adaptation of a new term or idea in a community. Some tools that are used to monitor or mine social media for innovation issues are for example, Attensity360, Brandwatch, Netbreeze Navigator, NM Incite - My BuzzMetrics, Radian6 or RapidSentilizer [3].

C. The Process of Social Media Mining for Innovation

Figure 1 gives an overview of the process of social media mining for innovation purposes. It builds on an overview of foresight with technologies [15]. We see it appropriate for a scheme of social media mining, but are skeptical about its linearity.

a) Preparation: This first stage includes all basic assumptions, researchers or customers make: What should be done? What is the goal of research?

b) Selection of tool(s): Another important issue is the usage of tools to support data collection, selection, and analysis.

c) Selection of sources and data: What concrete data should be used for analysis?.

d) Data collection: Data collection includes the sort of collected data, e.g., historical data, current postings, format of data.

e) Data processing: There are many ways for data processing. The selected tools and approaches influence the processing (e.g., from keywords, detection abilities for sentiments).

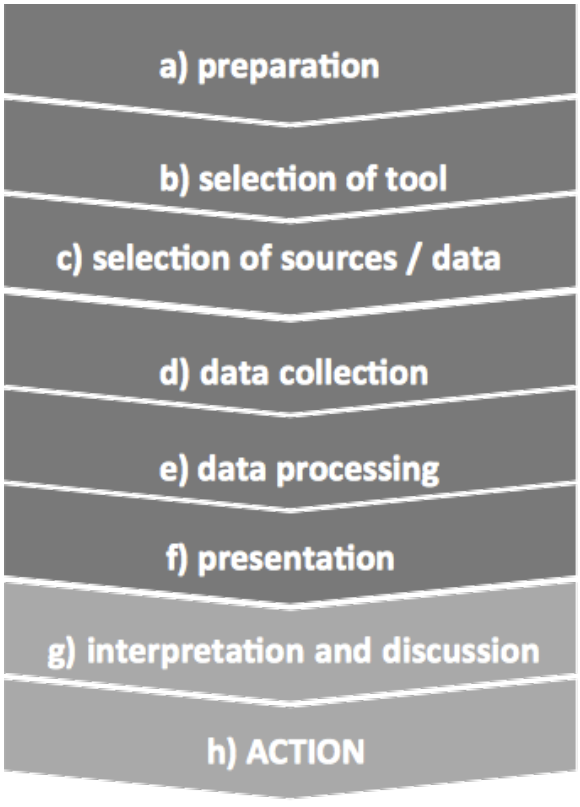

Figure 1. Process of social media mining for innovation purposes, based on [9], figure 3 (own variation)

f) Presentation of data: Finally, the results of the analysis should be presented in an appropriate and versatile way.

The last two phases of the process of innovation through social media mining (g) interpretation and discussion as well as (h) action are also influenced by, but not directly through, social media mining.

D. The Method of "Innovation Signals"

Exemplarily, we will now describe the research design of our project "innovation signals". On the one hand, this is an example for a concrete research design [10] [11]. On the other hand, this is also a description of the background of the involved researchers and might explain their special experiences, even when we sum up and generalize the experiences and future guidelines.

The concept called "Innovation Signals" exploits user-generated content for strategic innovation purposes by combining quantitative and qualitative methods. The Innovation Signals research approach does not rely on technology alone, but unfolds in the development of social media mining technology in unique combination with an interpretative methodology. The process is described as following.

a) Set-up: The set-up of Innovation Signals research mimics the traditional research design of empirical social science. The main goal is to formulate research hypotheses and define conceptual search terms, which contain between 20 and 50 English and German keywords. Then, 40 to 50 publicly accessible social web sources (forums, communities, blogs, newsgroups) are identified and quickly assessed, according to a catalogue of criteria (e.g., quality of contents, length of contributions, intensity of contribution).

b) Detection and monitoring: The social media mining-based technology provides automatic detection of relevant keywords and topics of interest in sources selected before. It first extracts a large amount of user posts (e.g., 200,000 posts) and then, automatically detects emerging keywords, topics and sentiments from compiled discussions and user's publicly available opinions.

The Innovation Signals technology provides answers to the questions in the context of product development and trend detection such as: How do users talk about existing products? What are critical issues? What issues are discussed very intensively? What are emerging topics? How do topics change over time? The technology enables experts to analyze and interpret detected innovation signals in an easy and intuitive way and also to save the most important posts for additional manual analysis and coding.

c) Identification and contextualisation of innovation signals: The automated analysis of textual content enables an efficient information processing, but the machine-processed information still remains ambiguous. In order to

enable effective research, the interactions in the social web must be structured additionally and analysed with social scientific methodology. This means to associate user generated content with relevant statistics, trends and theories to amplify the meaning of the information and to understand the consumers' conversations better and in a broader context.

d) Translation into business opportunities: This phase of the research process utilizes user generated content (in close co-operation with customers/companies) as an additional information source for strategic decision making with regard to the kind of innovation (product, process, business models, strategic innovation fields) to be pursued in order to determine the focus of the product innovation and market strategies and/or to detect new markets and new ideas.

The approach of innovation signals and the technology was developed and used within the project "Innovation Signals – Development of a Social Web Innovation Signals Amplifier System", funded by Austrian Research Promotion Agency. Three bigger and some smaller practical use cases have been delivered – for different branches and industry partners. The skiing industry, electric automobility, and the energy sector have been fields of application.

E. Quality criteria of Future Studies and Social Media Mining

According to our research, there is no comprehensive guideline available for the usage of social media mining for innovation purposes and future developments. Of course, there is a huge list of social media monitoring guides for general aspects as brand issues itself available, e.g., [12], [13]. An approach to develop guidelines is also to adapt quality criteria of future studies [14].

III. Research Question

The guiding question of our research is: What comprehensive guidelines for usage of social media mining tools might be offered to researchers and practitioners in the field of innovation management and future studies?

IV. Research Design

Within our study, we developed a multi-stage approach. First of all, we collected recommendations, lessons learned and comments from literature and the Web from (i) general social media monitoring guides and (ii) from usage of social media mining in research (literature). Short sentences and key advices were written on sheets of paper. Additionally, we made a list of quality criteria for future studies [15]. Building on this, concrete guidelines were to be developed in an expert workshop. The structured discussion included a systematic review of existing experiences as well as collection of our own expert assessments. In a final round, the main guidelines were approved.

V. Results

The development process as well as the guidelines for social media mining for innovation purposes is presented in the following.

A. The development process

The expert workshop was held in September 2014. The moderator and three experts in the field are authors of this text. According to the process of social media mining for innovation from Figure 1 the recommendations and lessons learned were first of all completed with lessons learned from own work (see paragraphs about innovation signals approach).

The key source of the guidelines is the own practical experiences of the experts. They were asked to give advice to beginners, to think about pitfalls and also successes to collect a first set of issues for recommendations. Then, existing recommendations or criteria were used to systematically enrich and broaden the set of possible guidelines. Existing recommendations are sources from literature and the Web [12], [13], existing criteria for tools for special usage [3] as well as general criteria of future studies [14]. Another framework for orientation for the guidelines was the process of social media mining from Figure 1.

After collecting, a phase of re-arranging, combining and also selecting more important issues and topics started. When all experts were satisfied with the first sketch of guidelines, the authors formulated these. According to the formulation of the guidelines we did not find a guide for good guidelines, therefore we orientated our formulation of guidelines on good practice [16]. In a final phase, the recommendations were discussed by all experts and formulated as follows.

B. Guidelines

The following guidelines should support future social media mining, especially for beginners and also potential future customers.

a) No tool will do your work automatically: The expectation that a tool or set of tools might be able to do the work or the majority of work automatically is not realistic – not at all. In contrary, even sophisticated systems, or stand-alone tools are more or less only a small support for comprehensive social media mining; at least by now. Notably, a similar assessment is available for social media monitoring tools: "None of them [the tools] do what you want to do them" [17]. Even more concrete, the experts' experiences with sentiment analysis for innovation purposes or tool-delivered (automatically detected and defined) trends detection are unsatisfying by now.

b) Narrow your search. The more narrowly the search field or topic was defined, the merrier our results and customers' satisfaction was. This is of special importance when researchers are not already familiar with the sector of industry or topic. The data-driven approach is helpful, but only with a clear preference and focus in the beginning. For example, "car mobility" seems to be a good topic for mining

 ISBN: 978-1-61208-386-5

at a first glance. But thinking on the wide range of possible items, it might have been helpful to discuss in the very first meeting topics, which are really of interest, for example, "tuning of electric cars". This opens the opportunity to mine for details within the concrete field of application.

c) The mining is manual social sciences; even better: The mining requires qualitative and quantitative social science. This is not only a consequence of the absence of useful stand-alone tools (see guideline a) but also building on the assessment of the power of the analytical approaches that are not easy to fulfill or overtake by algorithms. Manual work also includes that bottom-up ("data driven") and top-down ("hypothesis driven") approaches should be combined, as this is a typical approach of explorative social science. Re-adjusting working hypotheses, open minds and systematic procedures, e.g., coding and building of categories, are required working competencies for social media mining for innovation purposes. Some approaches directly build on social science experiences, for example, the so-called netnography approach [18]. The need of manual work by researchers is not only owed to missing possibilities for automation, but also the inevitable way to enhance results and realize smarter analyses. For example, the manual collection and selection of appropriate Web forums makes later analysis easier, as there is less " information noise" and unrelated content (eventually with same keywords) included.

d) The process is iterative rather than linear: Whereas typical illustrations of social media mining for innovation purposes are sketched as a linear process (see Figure 1), reality proves that successful social media mining for innovation has to be iterative and repetitive. A re-definition of goals, tuning of selected data and approaches, a combination of bottom-up (data-driven) and top-down (hypothesis driven) approaches, are needed and to be recommended in order to achieve satisfying results. Typically, an explorative process has no steady goals, hypotheses, sources, data, or possibilities to process.

e) Use well-selected topical discussion forums in the Web. The Web is full of opportunities and data where potentially interesting information might be delivered. The collection of data that is not directly about the key issues of the search requires further efforts while analyzing, especially if it contains similar key words and concepts in another context. Data extraction is simpler, if specialized forums are used. Besides this, we also try to choose such specialized forums for another reason. Topical discussion forums, e.g., on privately operated Wordpress-websites, more often not prohibit the analysis of public data (whereas most of the big commercial platforms, e.g., Facebook do).

f) Do not expect outstanding surprises when mining for innovation: According to several years researching for innovations for diverse customers, there had been no big surprises for insiders and experts within a certain domain. Presenting results from social media mining, even for vague "weak signals" to insiders and experts, in no case delivered a real wow effect. The rate of completely new topics or ideas is probably very small, the pure information is seldom perceived (!) as surprising and new.

g) The role of researcher as the customers' consultant is important: A solid know-how of social scientific methodologies seems to be a key factor, but this should be combined with a classical consultant expertise for a successful mining. A constant dialogue with the client, an effective management of user requirements and presentation skills is needed to be successful in form of clients' satisfaction. Being a good consultant also includes being able to transfer issues into the client's language.

h) Give results a meaning. Typically, social media mining delivers a long list of remarkable results, for example, a list of trendy topics or weak signals. Bringing social media mining to a success in form or customer satisfaction as well as potential action the results should get a meaning for the audience. There are several ways to give a good presentation, and not only a long list of themes and topics from your search field:

- It is not the list of topics, but the smart way to organize them in a model. This model might be the process of product development or service (e.g., from first users' contact on).
- Time series might deliver a feeling for a development or trend.
- Numbers are helpful, e.g., how many posts are analyzed and how many of them concern a special topic?
- Exemplary statements (quotes) from the sources should be used to illustrate the numbers.
- Nevertheless, the presentation should not blur information of needs, information about possible solutions or prototypes.

i) Social media is not always a good or analysable source for innovation: Social media is in general a good source for many innovation purposes [e.g., 9]. Nevertheless, it is restricted: typically on languages (English, German, some others) and concerning the users (no offliners). Developing ideas of new mobile games should be done with Japanese data; and future developments of corrida (bullfight) should be done with Spanish content. Additionally, false friends or other meanings of words in another language complicate the analysis and should be considered while selecting sources or data. Whereas "e-bike" in German is a bicycle with an electronic motor (also called Pedelec), "bikes" for example, in English are also motorbikes. "Handy" is the German word for "mobile phones", whereas "handy" is a common adjective in English.

j) Decisions within social media mining are a question of cost-value ratio, not of possibilities. Social

 ISBN: 978-1-61208-386-5

media mining is dealing with an endless source of data, especially if you broaden it from text to pictures and videos. As the mining itself is an explorative work or a spiral work (see d), the constant re-designs of the goal, source, collection and processing regularly needs a weighting up of workload and results and an assessment of the cost-benefit ratio.

C. Additional (concrete) recommendations

The experts also have two additional advices, non-related to the already mentioned guidelines.

- *Look for fitting thesauri!* Of course, thesauri are not available for any topic, but it is always worth to look if one is available. Using thesauri is a smart possibility to enhance results and analysis.
- *Use existing lists for social media mining and monitoring tools' evaluation!* Besides results of such evaluations and customers' feedback on the tools, the criteria used for evaluation might help you to get a clearer picture of your own needs. But always, take in mind guideline (a).

VI. Conclusion and Discussion

To sum up, the experts developed general advices as guidelines that may help to come to more realistic expectations concerning approaches and tools of social media mining for innovation purposes. Additionally, the guidelines give advice how the process might get better results and be more successful. The guidelines themselves are new and original, but in some points they are coincident with existing opinions and research [17]. As existing recommendations and lessons learned were part of the discussion, the presented guidelines can be seen as support for ongoing and future social media mining for innovation purposes.

Nevertheless, the development of guidelines and the guidelines themselves must be discussed critically. One point might be our set of experts. This might be considered as a limiting factor for the validity of the guidelines. The experts are only experienced in English and German speaking contexts and analyses, they have broad experiences with several tools and in several economic sectors, but of course there are still blind spots. As our current work with "innovation signals" focuses on technology-enhanced weak signals detection [19] this also might have influence on the results (compared with other social media mining for innovation approaches).

VII. Next Steps

For future developments and potential adaptions of guidelines for social media mining, a broader base of experiences might be considered. Practically, there is no international network of researchers and practitioners that might be helpful for this special task. But social media mining for innovation purposes seem to be a developing [9], if not boosting approach, so future settings might allow broader developments of guidelines. So, we hope our guidelines provide a valuable first step and a worthy support for beginners and practitioners alike.

Within our workshop and discussion several topics arouse that are still unclear from a research perspective, especially related to guideline (f) "Do not expect outstanding surprises when mining for innovation" [20]. Building on our experiences with social media mining, there seem to be space and possibilities to get better information and insights into innovation signals or trends. But the approaches, by now, are very limited concerning signals for disruptive innovations or other brand-new original issues that might become trendy in the future. Research that deals with historic data to evaluate existing approaches concerning their validity is also an interesting step forward.

Acknowledgment

Thanks to Kathrin Parson and Laurenz Giger for their support within the project.

References

[1] I. Stavrakantonakis, A. Gagiu, A., H. Kasper, I. Toma, I., and A. Thalhammer, A. "An approach for evaluation of social media monitoring tools" in: D. Fensel, H. Kett, and M. Grobelnik (Eds.), Common Value Management, Stuttgart Fraunhofer-Institut für Arbeitswirtschaft und Organisation, 2012, pp. 52-64.

[2] R. W. Helms, E. Booij, and M. R. Spruit, "Reaching out: Involin users in innovation tasks through social media", in: ECIS 2012 Proceedings, paper 193, Online: http://aisel.aisnet.org/ecis2012/193 (2014-09-15).

[3] K. Welz, L. Brecht, J. Kauffeldt, and D. Schallmo, "Weak signals detection: Criteria for social media monitoring tools", Proceedings of the 5th ISPIM Innovation Symposium: "Stimulating Innovation: Challenges for Management, Science & Technology", 09-12 December, 2012, Seoul, Korea.

[4] R. Zafarani, M. Abbasi, and H. Liu, Social Media Mining. An Introduction, 2014, Cambridge University Press.

[5] I. Ansoff, "Managing Strategic Surprise by Response to Weak Signals", California Management Review, 18, 2, 1975, pp. 21-33.

[6] B. Song, „Weak Signal Detection on Twitter Datasets. A non accumulated approach for non- famous events", Master Thesis TU Delft, 2012, Online available: http://www.cs.uml.edu/~hachreka/files/related/Thesis.pdf (2014-07-01)

[7] I. Stavrakantonakis, A. Gagiu, A., H. Kasper, I. Toma, I., and A. Thalhammer, "An approach for evaluation of social media monitoring tools", in: D. Fensel, H. Kett, and M. Grobelnik (Eds.), Common Value Management, Stuttgart Fraunhofer-Institut für Arbeitswirtschaft und Organisation, 2012, pp. 52-64.

[8] H. W. Chesbrough, Open Innovation: The New Imperative for Creating and Profiting from Technology, Harvard Business Press, 2006.

[9] F. Piller, A. Vossen, and C. Ihl, "From Social Media to Social Product Development: The Impactof Social Media on Co-Creation of Innovation", 2012, Online: http://www.alycante.it/wp/From-Social-Media-to-Social-Product-Development.pdf (2014-09-15).

[10] M. Markus, R. Eckhoff, and M. Lassnig, "Innovation Signals in Online-Communitys – ein komplementärer analytischer Ansatz", HMD - Praxis der Wirtschaftsinformatik, 293, 10/2013, pp. 13-21.

[11] M. Lassnig, M. Markus, R. Eckhoff, and K. Wrussnig, "Prospects of technology-enhanced Social Media Analysis for Open Innovation in the Leisure Industries" in: R. Egger, I. Gula, and D. Walcher, Eds., Open Tourism – Open Innovation, Crowdsourcing and Collaborative Consumption challenging the Tourism Industry. Salzburg, 2013.

 ISBN: 978-1-61208-386-5

[12] L. Bush and L. Glazier, „3 guidelines for using social media monitoring during a PR crisis“, Prdaily.com, Post from 2013-09-13, Online: http://www.prdaily.com/Main/Articles/3_guidelines_for_using_social_media_monitoring_dur_15189.aspx (2014-09-15).

[13] Socialmedia hq (without year). “The Definitive Guide to Social Media Monitoring for Business”, Online: http://www.socialmediahq.com/pdf/the-definitive-guide-to-social-media-monitoring-for-business.pdf (2014-09-15).

[14] M. J. Boon, E. Rusman, and M. R. Klink, “Developing a critical view on e-learning reports: Trend watching or trend setting?” International Journal of Training and Development, 2005, 9(3), pp. 1-27.

[15] J. Keller and H. A. von der Gracht, “The influence of information and communication technology (ICT) on future foresight processes — Results from a Delphi survey”, Technological Forecasting and Social Change, Volume 85, June 2014, pp. 81-92.

[16] Deutsche Forschungsgemeinschaft (DFG), Safeguarding Good Scientific Practice, 2014, Wiley. Online available: http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_1310.pdf (2014-09-15).

[17] J. Falls, “Where Social Media Monitoring Services Fail”, published at socialmediaexplorer.com, April 1, 2010, Online: http://www.socialmediaexplorer.com/social-media-monitoring/where-social-media-monitoring-services-fail/ (2014-09-15).

[18] G. Jawecki and J. Fuller, “How to use the innovative potential of online communities? Netnography – an unobtrusive research method to absorb the knowledge and creativity of online communities”, International Journal of Business Process Integration and Management, 3, 4/2008, pp. 248-255.

[19] R. Eckhoff, M. Markus, M. Lassnig, S. Schön, “Detecting Weak Signals with Technologies. Overview of current technology-enhanced approaches for the detection of weak signals”, International Journal of Trends in Economics Management & Technology (IJTEMT), volume III issue V, October 2014, URL: http://www.ijtemt.org/vol3issue5/1_Detecting_Weak_Signals_with_Technologies_Vol_III_Issue_V.php

[20] R. Eckhoff, M. Markus, M. Lassnig, S. Schön, “No outstanding surprises when using Social Media as source for weak signals? First attempt to discuss the impact of social media sources to detect surprising weak signals”, Proceedings of The Ninth International Conference on Digital Society (ICDS) 2015, in Lisbon, Portugal.

 ISBN: 978-1-61208-386-5