# Personal Information Management and Social Networks Re-finding on Twitter

Florian Meier
2nd year PhD candidate
Chair for Information Science
University of Regensburg
Germany
florian.meier@ur.de

David Elsweiler
Supervisor
Chair for Information Science
University of Regensburg
Germany
david.elsweiler@ur.de

## ABSTRACT

The following PhD project argues for analyzing Personal Information Management (PIM) behaviour on social networks and microblogging platforms, such as Twitter, as such services go far beyond their expected usage and have grown a viable source of information to many users. Of the three main information interaction activities that characterise PIM (finding - keeping - re-finding), only the retrieval of information has received researcher attention on such services. However, recent literature suggests that other PIM related activities occur, that should be further investigated. In addition to motivating the project, this position statement outlines the main research questions and approach, details the current status, which sheds light on how people perform PIM activities in the context of social networks and microblogging platforms. Here we detail some of the key findings and describe our tentative plans for the future. It is this plans that I wish to discuss in the doctoral consortium.

## MOTIVATION

The purposes for which people use social networks and microblogging platforms, such as Twitter and Facebook, go far beyond what one would expect. In addition to facilitating connecting and conversing with other friends and contacts, these services have become a valuable source of information in themselves. Many tweets contain useful information such as links to web pages or multimedia content, personal experiences, recommendations, or breaking news; any of which can motivate users to browse, manage and curate their timeline, search a corpus of status updates or query their followers [6, 16, 12]. This PhD project aims to investigate how users behave with social media content, for example, how and why social networking posts are archived and re-accessed, how existing features of social networks support this; and how systems could be improved to meet the needs of their users. The microblogging platform Twitter is a particular system of interest.


*IIiX '14* August 26 - 29 2014, Regensburg, Germany

http://dx.doi.org/10.1145/2637002.2637058 ...$15.00.

Personal Information Management (PIM) as defined by Jones is the "practice and study of the activities a person performs in order to locate or create, store, organise, maintain, modify, retrieve, use and distribute information in each of its many forms as needed to meet life's many goals and to fulfill life's many roles and responsibilities." [9]. PIM has been studied in many contexts, such as paper and electronic documents, email messages and web pages [9]. Some work specifically focuses on the cognitive and psychological processes during PIM [10, 3, 8]. The highly personalised nature of Twitter makes it a particularly interesting focus for PIM. According to Jones there are six senses in which information can be personal and so an object of PIM. Those are: (1) information is controlled or owned by me, (2) information is about me, (3) information is directed towards me, (4) information is sent (posted, provided) by me, (5) information is (already) experienced by me, (6) information is relevant (useful) to me [9]. All these six senses apply to tweets and social status updates, which underlines why investigating PIM behaviour on social media platforms would be important for the PIM community.

PIM consists of three main groups of activities trying to map people's needs and information. These activities are: keeping activities, finding/re-finding activities and meta-level activities such as organising and maintaining personal information collections (PIC). On Twitter, of these activities, only the finding of information has yet received significant attention. Researchers have shown that looking for information in status updates is a common type of use for users of microblogging services. Java et al. [7] were among the first to identify information seeking as a dominant motivation for using Twitter. Later Zhao and Rosson [18] as well as Naaman, Boase and Lai [13] confirmed and extended these findings. Another common behaviour is to query one's followers for help or opinions by asking questions and thus to exploit Twitter as a social Q&A system [12, 14]. Moreover several researchers aimed at the serendipitous nature of encountering valuable or relevant information in microblog content [1], while others analysed Twitter user's deliberate searching behaviour [16].

Via a large-scale query log study Teevan, Ramage and Morris [16] analysed people's motivation for searching Twitter. From the search logs the authors infer how users behave at searching Twitter and how this behaviour differs from more traditional search behaviour such as searching with web search engines. They discovered that users query Twitter mainly for time sensitive and person related information,

resulting in queries that are, compared to web search queries, more common and less changing. They further state that while web search queries are mainly to develop and learn about a topic, Twitter queries are for monitoring content. This behaviour leads to frequently repeated queries, that in web search often indicate re-finding behaviour (finding information that has been seen before)[15]. Teevan, Ramage and Morris [16] argue that re-finding on Twitter is difficult, and state that repeated queries in Twitter are solely an indicator for monitoring behaviour, but not for the need of re-finding information. Elsweiler and Harvey [4] studied the tasks and needs motivating Twitter search via surveying users in a long term diary study. Their results show that 43% of all user reported search tasks had to do with re-finding, standing in contrast with the results of Teevan, Ramage and Morris, were in a pre-study questionnaire only 2 out of 54 participants reported re-finding as a search task.

Although not focus of any in depth study, signs of keeping behaviour can be seen on Twitter. boyd, Golder and Lotan [2] analysed the characteristics and motives behind retweeting and found, that one motivation for doing so was the user's need for saving tweets for future personal access. The evidence suggests that the need for re-finding tweets exists, but Twitter's search facilities do not address all user needs. Moreover we can trace signs of keeping behaviour, which evidently also exist, but have not been investigated any further. Framing this in the context of PIM behaviour several questions arise.

## RESEARCH QUESTIONS

Recent literature shows that PIM related activities are prevalent in microblogging context. However little is known about how people integrate their personal information collection (PIC) of tweets, or the information within these tweets, into their personal space of information (PSI), nor do we know how they manage to do this and what motivates their behaviour. This leads to the following high-level research questions, which represent a spectrum of possible threads of inquiry: (1) Which tasks or needs motivate people to engage in PIM related activities on Twitter? (2) To which extent do PIM related activities occur? (3) What UI features or options does Twitter offer for PIM related activities? (4) Which other methods do users employ to fulfill PIM related activities? (5) How can we offer improved UI features for PIM on Twitter?

## METHODOLOGY

In order to shed light on the stated questions several research methods are to be combined to employ a mixed methods approach, which means that the results of different methods can be triangulated to create a rich picture of user needs and desires and how these may be satisfied. In the early stages of the PhD qualitative and quantitative data gathered via a large-scale questionnaire provide initial insights on how Twitter users engage in PIM related activities. The survey gives a high-level overview on the object of investigation and will mainly address research questions 2 and 3. Further studies depend on initial insights from the survey, which will provide some answers to the high-level research questions as well as unearthing more specific subquestions to be answered.

## PROGRESS MADE SO FAR

A large-scale survey was conducted to get a first high-level overview on how Twitter users engage in PIM related activities. A questionnaire surveyed 606 users (103 from a locally hosted, web-based survey, 503 from the online survey platform Tellwut.com) with respect to three main Twitter related PIM topics and demographic aspects. The topics were: (1) Favouriting Behaviour, (2) Preserving Behaviour, (3) Re-finding Behaviour.

Primary data analysis focused on the favouriting behaviour section [11]. There were two reasons for this: First, the favouriting option is one central function of the microblogging platform, that hasn't received researcher attention to date. Second, favouriting is strongly connected to PIM, as it is the primary feature of the service that allows the user to save tweets in a separate timeline, which enables management and possible re-access of such tweets. Within the favouriting section we surveyed participants, who knew about the favouriting feature (395 or 65.1% of all participants) on how many favourited tweets they have, on the frequency with which they make use of this feature, how often they unfavourite a tweet, how often they visit their favourites list and some others. Results show that the median for the number of favourited tweets participants had was 5, at which in two extreme cases users reported to have 8.000 and 55.660 favourited tweets. 290 participants did not have a tweet marked as favourite at that time. With respect to the frequency of their favouriting behaviour 26.8% stated to never favourite a tweet, 41.7% responded to do so only rarely, whereas 7.8% stated to favourite a tweet at least once a day. The favourites list is visited infrequently as 77.4% responded to never or rarely do so. Participants also stated rarely changing their mind about a favourited tweet as 91.8% never unfavourite a tweet they once favourited.

Moreover, two free form questions queried participants on their motivation for using the favourite button. In order to analyze the respondents comments we applied an approach aligning with Glaserian Grounded Theory [5]. This approach led to a set of categories or codes representing participants favouriting motives. The results of our qualitative analysis show that motives for favouriting are very heterogenous resulting in a taxonomy consisting of 25 different motives and reasons [11]. Despite the diverse nature of reasons, two main favouriting use cases can be identified, these are: favouriting as a response and favouriting for a purpose or as a function. The most dominant purpose favouriting is used for is bookmarking. Within this category we can distinguish between (1) bookmarking for a future use e.g. already with an anticipated need in mind and (2) bookmarking for memory related purposes e.g. for high valued tweets that users collect for keeping. Both practices are strongly connected to keeping and re-finding.

## FUTURE PLAN

First, future steps aim at the completion of survey data analysis. Here, in addition to basic descriptive statistics of keeping and re-finding behaviour, present work focuses on using statistical modeling approaches to understand the relationship between factors such as the amount of followees participants have, and their responses on keeping and re-finding behaviour.

Second the initial findings on favouriting behaviour, pre-

sented in section 4, raise questions regarding the process of favouriting for the purpose of bookmarking that call for further investigation. (1) If people use favouriting for a possible re-access of tweets why is the favourites list visited infrequently? (2) If tweets are saved in the favourites list for a future interaction step, that might be completed after having visited the tweet again, are users going to unfavourite that tweet? The survey's result on participants unfavouriting behaviour would suggest the opposite, as 91.8% state to never unfavourite a tweet. This would overload the favourites list making re-finding a difficult task again. Besides that, frequent unfavouriting of tweets would partially explain, why some people don't have any favourited tweets at all, as they regularly unfavourite tweets, that they don't need any longer (e.g. in form of a spring cleaning, that can be observed in PIM with emails [17]). (3) Favouriting and retweeting have a very special relationship as both features are used for similar motives like the possible re-access of tweets. Further investigations should clarify the picture of how the relationship between both features looks with respect to PIM.

Several methods seem appropriate for further data acquisition and analysis hence complementing and confirming results from the high-level survey. Among these are: naturalistic approaches and lab-based studies. Naturalistic approaches, such as diary studies or log-file analysis, are highly suitable for studying PIM behaviour as they are capable of capturing user's real behaviour in everyday life situations, accounting for the many contexts in which PIM can happen, therefore overcoming the difficulties caused by the personal nature of PIM. Laboratory-based studies, on the contrary, try to simulate this everyday life interaction in the controlled setting of the laboratory, coping with the problem of the availability of personal information collections, but offering the opportunity of conducting highly focused narrow scoped studies. Support by members of the doctoral consortium would be appreciated in order to discuss which research questions to address next and which methods to apply.

# REFERENCES


[1] T. Bogers and L. Björneborn. Micro-serendipity: Meaningful coincidences in everyday life shared on twitter. In *Proceedings of iConference 2013*, pages 196–208. iSchools, February 2013.

[2] d. boyd, S. Golder, and G. Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*, HICSS '10, pages 1–10, Washington, DC, USA, 2010. IEEE Computer Society.

[3] D. Elsweiler, M. Baillie, and I. Ruthven. Exploring memory in email refinding. *ACM Trans. Inf. Syst.*, 26(4):21:1–21:36, 2008.

[4] D. Elsweiler and M. Harvey. Engaging and maintaining a sense of being informed: Understanding the tasks motivating twitter search. *JASIST*, 2014.

[5] B. G. Glaser and A. L. Strauss. *The discovery of grounded theory: strategies for qualitative research.* Aldine, Chicago, 1967.

[6] J. Hurlock and M. L. Wilson. Searching twitter: Separating the tweet from the chaff. In L. A. Adamic, R. A. Baeza-Yates, and S. Counts, editors, *ICWSM*. The AAAI Press, 2011.

[7] A. Java, X. Song, T. Finin, and B. Tseng. Why we twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, WebKDD/SNA-KDD '07, pages 56–65, New York, NY, USA, 2007. ACM.

[8] W. Jones and B. H. Ross. Personal information management. In F. T. Durso, editor, *Handbook of applied Cognition*, pages 471–496. John Wiley and Sons, 2007.

[9] W. P. Jones and J. Teevan. *Personal Information Management*. University of Washington Press, 2007.

[10] M. Lansdale. The psychology of personal information management. *Applied Ergonomics*, 19(1):55–66, 1988.

[11] F. Meier, D. Elsweiler, and M. Wilson. More than liking and bookmarking? towards understanding twitter favouriting behaviour. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*, 2014.

[12] M. R. Morris, J. Teevan, and K. Panovich. What Do People Ask Their Social Networks, and Why? A Survey Study of Status Message Q&A Behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1739–1748, New York, NY, USA, 2010. ACM.

[13] M. Naaman, J. Boase, and C.-H. Lai. Is it really about me? message content in social awareness streams. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, CSCW '10, pages 189–192, New York, NY, USA, 2010. ACM.

[14] S. A. Paul, L. Hong, and E. H. Chi. Is twitter a good place for asking questions? a characterization study. In L. A. Adamic, R. A. Baeza-Yates, and S. Counts, editors, *ICWSM*. The AAAI Press, 2011.

[15] J. Teevan, E. Adar, R. Jones, and M. A. S. Potts. Information re-retrieval: Repeat queries in yahoo's logs. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pages 151–158, New York, NY, USA, 2007. ACM.

[16] J. Teevan, D. Ramage, and M. R. Morris. #twittersearch: A comparison of microblog search and web search. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11, pages 35–44, New York, NY, USA, 2011. ACM.

[17] S. Whittaker, V. Bellotti, and J. Gwizdka. Everything through email. In *Personal Information Management*, pages 167–190. University of Washington Press, Seattle, 2007.

[18] D. Zhao and M. B. Rosson. How and why people twitter: The role that micro-blogging plays in informal communication at work. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, GROUP '09, pages 243–252, New York, NY, USA, 2009. ACM.