

# HCI Research and History: Special Interests Groups on Facebook as Historical Sources

Mechtild Stock<sup>(✉)</sup>

Kerpen, Germany

MechtildStock@gmail.com

**Abstract.** In this article, we discuss whether Social Network Services (SNSs), especially Facebook, are sources for microhistory or history from below. It can be shown that SNSs indeed form a valuable source for historical science. However, due to the huge amount of data in SNSs there is a problem not to be overwhelmed with unmanageable large data sets. Our proposal to structure the data sets is to apply informetrics and statistical methods as means of quantitative history. We work with a case study, i.e. *Kerpener und Ex-Kerpener*. This special interest group on Facebook addresses Kerpen (a small town in Germany) as well as its historical development. We are going to analyze about 2,000 wall posts of the group during 2014. For each post, we investigate its type (text, image, and video), category, topic, number of likes, shares and comments as well as the author's name. The article concludes with a list of decision criteria for historically relevant and credible Facebook wall posts as sources.

**Keywords:** HCI research · Microhistory · History from below · Facebook · Information science · Facebook metrics

## 1 Introduction: Facebook and History from Below

Human-computer interaction on social media is sometimes directed to local or regional cultural information. Prime examples for such interactions are microblogs on Twitter [3] or special interest communities on social networking services (SNSs) as Facebook which are related to contemporary and past local history [7]. In this article, we discuss whether SNSs are sources for history, especially for microhistory [2, 5]. “Microhistory” [1] or “history from below” [9] is history from the perspective of common people. Lay historians and average people act as citizen scholars. SNSs, with their well-documented dialog and content, appear to be excellent sources for constructing an account of history from below. It can be shown that SNSs with their processes of dialogs of “common” people indeed form a valuable source for historical science [8]. Here, historians and archivists are able to locate additional information, which cannot be found through any other sources. Due to the huge amount of data in SNSs, the greatest problem is being overwhelmed with very large datasets [6]. One aim of our study is to divide the amounts of posts and comments into two groups, namely, historically relevant items and less relevant or not credible items (Fig. 1).

We worked with a case study, namely *Kerpener und Ex-Kerpener*. This German Facebook group addresses Kerpen as well as its historical development. We analyzed

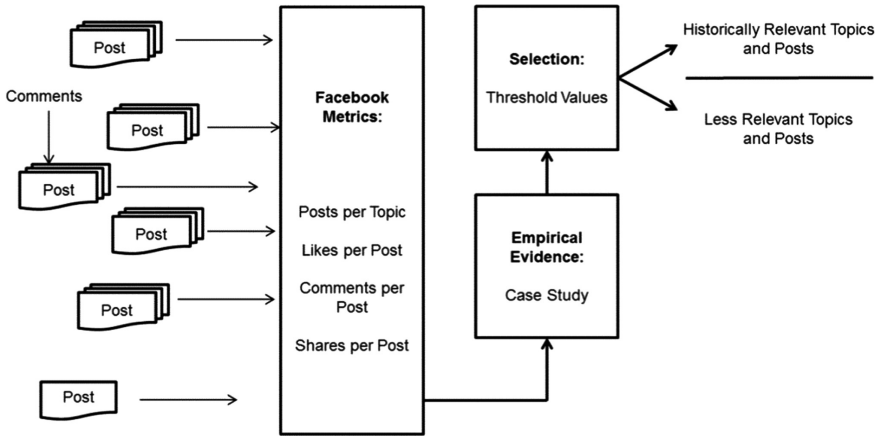


Fig. 1. Research model.

all wall posts of this special interest group during 2014. The article will have practical implications for cultural heritage institutions, e.g. archives [4]. This combination of HCI research, information science and history succeeded as a research approach that can be expanded to include other SNS groups which postings and observations may help produce a fuller historical record of both our time and place.

Facebook is an SNS, an important and popular kind of social media. On Facebook, people are able to create huge amounts of data. A crucial aspect is the credibility of the data. “This kind of system provides first-hand data, but one pressing problem is to distinguish true information from misinformation and rumors. In many cases, social media data is user generated and can be biased, inaccurate, and subjective” [6, p. 113]. Therefore, we need methods to separate historically relevant information from misinformation and rubbish.

To manipulate such “big data” in information services as Facebook, we apply quantitative methods borrowed from information science, more precisely from informetrics. Informetrics originates in scientific communication. It measures the dialog in science in terms of publications and citations. In SNSs, there is dialog as well. The wall posts act as articles; the likes, comments and shares are analogues to citations.

## 2 Methods

Our case study for history from below is the special interest group *Kerpener und Ex-Kerpener*. This Facebook group addresses Kerpen as well as its historical development. Kerpen is a town of 65,000 inhabitants (2012) located in the German Rhineland. The aim of this moderated group is to preserve historical images and videos and to make them accessible. The group *Kerpener und Ex-Kerpener* was founded in 2012 and was able to attract 5,455 members (by the end of 2014).

At the appointed date of January 19, 2015, we downloaded all wall posts of the group *Kerpener und Ex-Kerpener* from 2014 to an offline HTML file (about 42 MB).

We only gathered the visible comments in our offline file. We ignored all those posts lacking in content (e.g. posts without any text, image or video). Our file consists of 1,951 wall posts in total. In consequence of the huge amount of comments (26,319) we decided not to evaluate their content. We constructed a database with the following field scheme: date (month, day), day of the week, author, type of post (text, image, video, text and image, text and video – intellectually coded), shared post (post from external source), number of likes, number of shares, number of comments, kind of image and video (current image, old image, current placard, current video, old video – intellectually coded), content category (intellectually coded), content description (keywords – intellectually coded). First of all, we roughly screened the posts in order to summarize and to determine content categories. Nine categories are the result of our analysis: caution (warnings), curiosity (“what is happening there?”), current impression (“current” means only some months old), news, notice (announcements, tips), old impression, private (all posts of private nature, including recommendations and requests for help), report/criticism (complaints, experiences), request (questions of general interests).

### 3 Results

In 2014, 1,951 posts in *Kerpener und Ex-Kerpener* have been written by 582 different members; hence the share of active members is 10.67%. Only a small segment of members is really active in writing wall posts. The majority of members obviously does not trigger discussions, yet some people response to wall posts by liking, sharing and commenting. Others, indeed, are only “lurkers” and pure consumers. A wall post in the year 2014 has on average 13.17 likes, 0.85 shares and 13.49 comments (Table 1). Who are the active members, regularly contributing posts? We sorted the entire set of posts by authors. We can identify an extremely left-skewed distribution. This means that there are only few highly productive authors and a long tail of authors contributing two posts or one per year. 24 authors produce 50% of all posts. The most productive author alone is responsible for 15.27% of all posts.

We have ranked our wall posts according to numbers of likes, shares and comments. The rankings of posts by likes and shares are – similar to the authors’ distribution – left-skewed distributions. Only few wall posts get many likes and shares. In most cases, the group members like current and old impressions. The top-liked post is an image of a wrong place-name sign at the new motorway A4 (*Bergdorf* instead of *Bergheim*) (Table 2). The interest in Michael Schumacher (a famous former Formula 1 driver) is understandable, since Michael once was a citizen of Kerpen. A moderate highly liked wall post is an impression of winter in Kerpen.

How does the distribution of shares look like (Table 3)? Most of the top wall posts by number of shares are private requests and requests for help; the others are warning notices. One post occupies a high position in the ranking. 722 members share the wall post searching for a hit-and-run driver. Two posts cover the issue of burglary and ask for attention. Six further posts are devoted to dogs and cats, e.g. dog found, dog poisoned, cat disappeared, and – most horrible – cat halved). We summarize that shares are used for current events. The more shares the more historically irrelevant the wall

**Table 1.** *Kerpener und Ex-Kerpener*: Basic figures (2014) (N = 1,951)

Basic figures	Kerpener und Ex-Kerpener
Number of posts	1,951
Number of members	5,455
Number of unique post authors	582
Share of active members	10.67%
Number of top post authors	24
Number of likes	25,686
Likes per post (average)	13.17
Number of shares	1,658
Shares per post (average)	0.85
Number of comments	26,319
Comments per post (average)	13.49
Number of images and videos	1,236
— Number of current images	679
— Number of old images	427
— Number of current placards	111
— Number of videos	19

**Table 2.** Top posts by number of likes (N = 1,951)

Rank	Likes	Kind of post	Description
1	640	Current impression	Wrong place-name sign ( <i>Bergdorf</i> instead of <i>Bergheim</i> )
2	338	News	Michael Schumacher awake
3	275	Old impression	Winter in Kerpen (2010)

**Table 3.** Top posts by number of shares (N = 1,951)

Rank	Shares	Kind of post	Description
1	722	Request for help/private	Wanted: hit-and-run driver
2	175	Caution	Burglary/tramp's sign
3	161	Request for help/private	Found dog

posts seem to be. But there are exceptions: The posts about burglary are not necessarily historically irrelevant, however.

Shares and likes do not need much cognitive effort; they are just one click, a touch of a button. Contrariwise, to produce a comment (maybe including images and videos besides pure texts) requires elaborate cognitive work. The top posts by number of comments include requests, notices, news, current impressions and warnings (the mentioned post about burglary) (Table 4). A hot topic in Kerpen is an empty apartment tower which has been set on fire several times. Questions of general interest (e.g., What do you associate with Kerpen?) trigger high numbers of comments. The often commented post about the kiosk around the corner describes a shop with an upholstered

**Table 4.** Top posts by number of comments (N = 1,951)

Rank	Comments	Kind of post	Description
1	359	Notice	Fire: Apartment tower Maastrichterstraße
2	353	Request	What do you associate with Kerpen?
3	299	Current impression	Kiosk around the corner

sofa on the sidewalk in front of the house. Within the top commented posts you can identify historically relevant posts (such as the problematic apartment tower) as well as gossip and tittle-tattle (e.g. the kiosk around the corner).

Our results clearly exhibit highly significant differences between multimedia and text-only posts in terms of the average numbers of likes and comments, but no statistically noticeable difference with regard to the average number of shares. Multimedia posts received on average 17.77 likes per item in contrast to only 4.73 likes per textual post. This is nearly four times the amount in favor of multimedia posts. In contrast, text-only posts gain 19.13 comments on average, while multimedia posts only get 10.41 comments on average. This is just about twice as much, but now in favor of textual posts. Obviously, multimedia posts often provoke many likes (meaning “This image pleases me,” and therewith everything has been said) and only few comments. Text-only posts lead to the opposite user behavior. Such posts are moderately high liked, but provoke a lot of comments.

Outstanding categories by the average number of likes are old und current impressions. Both categories receive more than 23 likes per post. Obviously, people do like old as well as new images of their hometown. In contrast to the high number of likes, both impressions’ categories only get moderately high numbers of comments (about 11 and 12 comments per post) and actually no shares. Only very few likes per post go to the categories private and request, but both categories include large numbers of comments (private 15 and request 28 comments per post). Private posts and requests of general interests call for answers (comments) and not for likes. There is just one category getting lots of shares: caution. Wall posts in this category are devoted to current burglaries and to warnings concerning dangerous situations for cats and dogs. Here, fast information dissemination is important, which can be reached by immediate sharing of the caution-posts. Additionally, these posts have high numbers of comments (29 on average) and moderately numbers of likes (8 on average). All other categories contain only small numbers of shares or no shares at all.

Top topics are defined by the absolute number of wall posts, sorted by keywords. We identified three topics as “Top Topics.” Do these posts provide historically relevant information in addition to news reported in local newspapers? The top topics show moderately high numbers of likes and comments as well as nearly no shares. 30 wall posts are on the topic thunderstorm. In the local press (*Kölner Stadtanzeiger*; June 10, 2014), we only find one article. In the wall posts, people report on subjective feelings, concrete impressions of vested interest, and offers of help. 25 wall posts are devoted to the topic Maastricht Street. This subject is frequently reported in the local press, too. The *Kölner Stadtanzeiger* covers, among others, arson attacks on the unoccupied

building (June 2, 2014), actions by the city administration of Kerpen (June 6, 2014), and security problems of the residents (June 26, 2014). In contrast, the wall posts concentrate on offense, on personal experience and hints to handle the problem. One wall post on the topic Highway A4 reports the wrong place-name sign Bergdorf instead of Bergheim (our top-liked post). This post was even the source of an article in *Kölner Stadtanzeiger* (September 16, 2014).

## 4 Discussion

Decision criteria for historically relevant and credible Facebook posts as sources are:

- high number of posts per topic in a given time interval (e.g., a year) (but one has to know which topics are relevant – our approach was to index all posts through categories and keywords),
- there is a high probability of historical relevance for the categories old impression, current impression and news; additionally perhaps selected posts from the notice and report/criticism categories,
- historically relevant multimedia posts (images, videos) have a moderate high number of comments and a high number of likes,
- some wall posts of historical relevance exhibit a high number of comments,
- high numbers of shares seem to be indicators for rapid requests and warnings which are seldom historically relevant.

To sum up: Facebook is an important source which complements other historical sources. On Facebook you can find information you will hardly find elsewhere: first-hand impressions, images and comments of “common people.” Since we only worked with one case study, it is necessary for further scientific investigations to analyze other Facebook groups which are interested in historical aspects of their environment. Additionally, the application of Facebook metrics should be broadened and calibrated. All in all, the new combination of HCI research, information science and historical science proved to be successful and is expandable.

## References

1. Brewer, J.: Microhistory and the histories of everyday life. *Cultural Soc. Hist.: J. Soc. Hist. Soc.* **7**(1), 87–109 (2010)
2. Davalos, S., Merchant, A., Rose, G.M., Lessley, B.J., Teredesai, A.M.: ‘The good old days’: an examination of nostalgia in Facebook posts. *Int. J. Hum. Comput. Stud.* **83**, 83–93 (2015)
3. Malone, M.: Tweeting history. An inquiry into aspects of social media in the Egyptian revolution. In: Hall, B.L., Clovers, D.E., Crowther, J., Scandrett, E. (eds.) *Learning and Education for a Better World*, pp. 169–182. Sense, Rotterdam (2012)
4. McCown, F., Nelson, M.L.: What happens when Facebook is gone? In: *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, pp. 251–254. ACM, New York (2009)

5. Nack, F.: Social media in history. In: Proceedings of the 2012 International Workshop on Socially-aware Multimedia, pp. 51–56. ACM, New York (2012)
6. Pinheiro, A., Cappelli, C., Maciel, C.: Providing tools to enable information audit in social networks. In: Stephanidis, C. (ed.) HCI 2015. CCIS, vol. 529, pp. 113–117. Springer, Cham (2015). doi:[10.1007/978-3-319-21383-5\\_19](https://doi.org/10.1007/978-3-319-21383-5_19)
7. Silberman, N., Purser, M.: Collective memory as affirmation: people-centered cultural heritage in a digital age. In: Giaccardi, E. (ed.) Heritage and Social Media Understanding. Heritage in a Participatory Culture, pp. 13–29. Routledge, London (2012)
8. Stock, M.: Facebook: a source for microhistory? In: Knautz, K., Baran, K.S. (eds.) Facets of Facebook. Use and Users, pp. 210–240. De Gruyter Saur, Berlin (2016)
9. Thompson, E.P.: History from below. *Times Literary Suppl.* 7, 279–280 (1966)