Research Article

Eelco Herder*, Daniel Roßner, and Claus Atzenbeck

Reflecting on Social Media Behavior by Structuring and Exploring Posts and Comments

https://doi.org/10.1515/icom-2020-0019

Abstract: Social networks use several user interaction techniques for enabling and soliciting user responses, such as posts, likes and comments. Some of these triggers may lead to posts or comments that a user may regret at a later stage. In this article, we investigate how users may be supported in reflecting upon their past activities, making use of an exploratory spatial hypertext tool. We discuss how we transform raw Facebook data dumps into a graph-based structure and reflect upon design decisions. First results provide insights in users motivations for using such a tool and confirm that the approach helps them in discovering past activities that they perceive as outdated or even embarrassing.

Keywords: Social media, posts, comments, user behavior, visual exploration, hypertext

1 Introduction

Social networks and social media are widely spread. According to the *Digital 2020* report,[1] there are 2.5 billion active Facebook users, which makes Facebook the world's most-used social platform, followed by YouTube with 2.0 billion users.

Online communication and social networking follow different patterns than face-to-face conversations or phone calls. The absence of non-verbal and visual feedback and other factors is compensated for by netiquette (the rules of etiquette that apply when communicating over the Internet) and platform elements that shape the conversation.

Platforms such as Facebook continuously apply smaller or bigger changes to their user interfaces and recommender algorithms; users then adapt their behavior (or not) in response, which then might be taken into account for further platform changes. A notorious example of a recent change was Facebook's introduction of different emotions (love, haha, wow, sad, angry) as refinements of the simple 'like' button [39], which enabled users to dislike a post. Similarly, the 'care' response – introduced during the COVID-19 pandemic[2] – has become widely used as well, for various purposes.

Social networks like Facebook can guide user responses (likes, comments, follow-up posts) as well as user mood [28] by changing the relative proportions of different content (among others, posts or images from friends, posts from various pages, suggested posts or advertisements). While the exact mechanisms used are largely black-box [12], Facebook acknowledges that its algorithms respond to user activity (where any reaction to a post, comment, photo, video or advertisement is considered a sign of interest – be it positive or negative).

Arguably, platforms like Facebook have an interest in maintaining their users' attention and increasing user activity and social feedback – the more time spent within the platform, the more opportunities for serving advertisements and other types of sponsored suggestions. It has been shown that users' opinion and behavior is influenced by these dynamics [28].

Several studies have investigated how users may regret certain posts or comments that were posted in a particular state of arousal and//or that received non-anticipated negative responses [41]. These studies are very helpful in shaping the netiquette among social media users and for reflecting on one's own social behavior. An aspect that has not received much attention until now is how users may reflect on how their own (clicking, liking, posting and commenting) behavior reinforces negative or addictive aspects of a social platform like Facebook.

1 See https://datareportal.com/reports/digital-2020-global-digital-overview, data updated to 25 January 2020.

2 https://www.socialmediatoday.com/news/facebook-unveils-new-care-reactions-to-help-express-responses-to-covid-19/576318/

***Corresponding author: Eelco Herder,** Radboud Universiteit Nijmegen, Nijmegen, the Netherlands, e-mail: eelcoherder@acm.org
Daniel Roßner, Claus Atzenbeck, Institute of Information Sciences, Hof University, Hof, Germany, e-mails: daniel.rossner@iisys.de, claus.atzenbeck@iisys.de

Examples of behavior that users typically regret include spending too much time on checking for new content and then realising that they only get ‘more of the same’ [7], being aroused by intentionally polarising headlines or topics, and responding to ‘clickbait’ [29].

In this article, we investigate how users may be supported in reflecting upon their past activities, and to separate purposeful activities from actions that they may want to reduce. Doing so is expected to have two mutually reinforcing benefits: it may reduce unwanted responses triggered by, for example, sponsored posts; furthermore, the reduced responses will most likely be interpreted as reduced interest and taken into account in future selections of material to show.

More specifically, we discuss the development and preliminary evaluation of our system *Mother* as a hypertext tool for exploring social media data [19]. The goal is to increase the users’ awareness of what data and, more importantly, which associations between information units the platform provider stores, and which insights it may algorithmically deduct from ongoing, reoccurring or co-occurring terms and topics as well as from temporal relations.

The remainder of this article is structured as follows. In Section 2, we discuss relevant related work on social media behavior, personal online data and triggers for possibly regretful actions. We then present Mother’s spatial hypertext tool in Section 3 as an exploratory UI for such data. Section 4 describes the steps from a Facebook dump to its visual exploration in greater detail. Section 5 includes an exploratory evaluation of Facebook posts and a reflection on Facebook commenting behavior. Finally, Section 6 concludes this paper, raises open research questions, suggests future work.

2 Problem Statement and Background

In this section, we discuss how social media platforms shape user activity by analyzing and exploiting patterns in user behavior. We start our argument with the observation that social media platforms have various (monetary) reasons for providing triggers to keep users engaged. In the subsequent subsection, we summarize a wide body of research on which triggers and communication strategies lead to increased user engagement and responses. Another line of research that we discuss shows that users may regret having posted particular posts and comments, either or not triggered by (commercial) social media campaigns.

For the reasons above, it is important that users will be able to review and reflect upon their actions. We believe that the European General Data Protection Regulation (GDPR) provides opportunities for developing tools to support these activities, which will be the focus of the final part of this section.

2.1 Monetization of Social Media Activity

Arguably, user activity in social media platforms is strictly voluntary. However, the traces of such social interaction and communication of social media users create significant opportunities for the platform provider. For example, Facebook can analyze people’s posts, behavior, or habits, trace users’ click histories on public websites or identify patterns on various other user data. Information about its users enables social media platforms to issue ads specific to the individual user.

The extent of a company’s benefits achieved by exploiting user activity within their platform(s) becomes clear when looking at Facebook’s annual financial report,[3] which states that 98 % of its 2019 revenue is based on advertising, a total of 69.5 billion USD. This is an increase of 27 % compared to 2018 or 74 % compared to 2017.

Personal data is used by web platforms for a variety of purposes, varying from personalization and recommendation to monetization, for example via advertisements and nudges to continue visiting the platform [18]. Both real and perceived discrepancies between the use of personal data for the benefit of the end user on the one hand, and the use of the same data for monetization has led to several privacy concerns. A particular concern is the interpretation of data into, among others, user interest profiles, beliefs and demographics, consumer behavior or even health status [1].

There is a conflict of interest between collecting information about users (possibly by combining various sources such as Facebook and WhatsApp, both platforms owned by Facebook, Inc.) and the users’ right of privacy. The same is true for third-party companies crawling social media data for the same purpose [11]. Unexpected results in personalized advertisements have been noticed, which suggests that information of various sources have

3 FORM 10-K for Facebook, Inc., *Annual Report Pursuant to Section 13 or 15(d) of the Securities Exchange Act of 1934 for the fiscal year ended December 31, 2019,* http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/45290cc0-656d-4a88-a2f3-147c8de86506.pdf

been used by the social media providers [20]. Moreover, as explained in the next subsection, there is a body of research on triggers for user activity that reinforces the effectiveness of such advertisements.

2.2 Triggers for Non-Planned and Possibly Regretful Activities

A platform such as Facebook enables users to communicate with one another, but it also allows companies, organizations, celebrities and other parties to reach out to users with similar communication means. Recent studies have confirmed that in such mediated communication contexts users strongly respond to and mimic language features and other verbal, visual and non-verbal cues [31].

In marketing and communication studies, these effects are actively being investigated with the aim to identify successful triggers and measure them using 'digital engagement metrics' [43]. These effects are successfully employed in marketing strategies. For instance, it is found that visual features lead to *likes*, rational and interactive features to *comments* [25].

Behavioral research on user responses in terms of heart rate and skin conductance level just before pressing a like button inspired strategies 'to enhance cognitive and emotional engagement' as 'pathways to virality' [2].

It is also known that certain topics and communities are more vulnerable to unwanted, toxic comments and discussions than other topics or communities [4]; this line of research mainly appears to focus on the direct triggers for toxic comments.

Triggers as discussed above may lead to non-planned responses and behavioral patterns that users may regret afterwards, as will be discussed below.

2.3 Regrets About Social Media Behavior

In a much-cited study on Facebook regrets [41], participants explained several reasons for regretting: posting sensitive content, among others about alcohol and illegal drug use, about religion and politics, or about personal and family issues; posting negative or offensive comments, or revealing lies and secrets. Reasons and motivations for these posts and comments include thinking that they are funny, or for venting frustrations, or good intentions that did not work out as intended. Most posts became regrets after unforeseen consequences or responses.

In a follow-up study [40], a number of privacy nudges have been designed and evaluated, including notices that 'anyone on the internet can see your post' and a 10-second timer during which users can still cancel their post. Once the nudges were understood and considered usable enough, users started to appreciate them.

A recent literature study [17] concluded that most people have regretful experiences on social media, most of them related to privacy, and that such experiences may lead users to stop using social media. Particularly regretful experiences arising from the respondent's actions (as discussed above) increase the likelihood of deactivation.

Social media regrets are not just limited to single unfortunate actions, but also to the time spent on social media on itself. Factors such as the fear of missing out may contribute to social media addiction [9]. Increasingly, it is argued and shown that social media companies design their platforms in a way that renders them addictive, as this is precisely what the attention-economy business model is all about [8].

2.4 Personal Data, the GDPR and Interpretability

Increasingly, governments introduce regulations that protect citizens against the company's market power or lobbyists, including practices as discussed above. The European Union introduced the General Data Protection Regulation (EU) 2016/679 (GDPR) [16]. This opens, among others, the right to EU citizens to request their data from platform providers.

The introduction of the General Data Protection Regulation, the GDPR [16], in Europe has led to several restrictions in which data can be used by industry and researchers alike and provides end-users with means for requesting transparency. Following the principles of responsibility, explainability, accuracy, auditability and fairness, several initiatives for responsible (HCI) research have been proposed [38]. A further opportunity that the GDPR offers to the scientific community is the result of Article 20, the "Right to data portability",[4] which states that:

> "The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format […]"

Recently, researchers investigated user expectations and practices regarding the GDPR Right to Data Portability by asking users to request their data from the German loyalty

4 https://gdpr-info.eu/art-20-gdpr/

program Payback. It turned out to be "unexpectedly simple and uncomplicated" to request the data, but participants believed that the data did not "paint the complete picture". Particularly, Payback did not provide any derived data, such as profiling or classification [3].

Even though users may receive their data in human readable format, it is still difficult to explore it in a meaningful manner and to reach a high awareness of the personal data that is stored and computed by the platform provider's algorithms. In most cases, the *connections* between data snippets are more of interest than the data itself. Thus, a tool is required that can be used for exploring the various connections and associations between data.

Analysis and visualization of social media activities and connections is paramount in social network analysis [21], but these methods are typically targeted at whole user populations, in order to identify clusters of users, topical clusters or temporal patterns. Visualization of a user' individual patterns for reflection is far less common. The area of Personal Visual Analytics aims to find and create design spaces that enable exploration, create awareness and support reflection [22]. In this article, we present Mother as an example of such a tool.

3 Mother as a Tool for Visual Data Exploration

Our approach for visualizing user activity is rooted in spatial hypertext. In this section, we introduce the concepts of hypertext, navigational hypertext and spatial hypertext, followed by an overview of Mother, the tool we used and further developed for personal visualization of social media activities.

3.1 Spatial Hypertext: Context and Visualization

Dumps of personal data or posts from social media platforms may be voluminous. Therefore, they generally need to be teared apart into smaller coherent informational units and associated in a semantically meaningful way. The result is a weighted undirected graph in which the edges represent the associations between information units. We will describe the full process in Section 4.

A naive way for displaying this data would be graph visualization, for example, using frameworks like Graphviz.[5] However, there may be too much information with many relevant associations in between to be displayed. Considering the full graph would result in a visualization that is hard to read and, thus, of low relevance for the users who wants to explore their own data. For this purpose, an iterative process is required that also can be found in *visual analytics*: "The visual analytics process combines automatic and visual analysis methods with a tight coupling through human interaction in order to gain knowledge from data" [24]. The underlying mantra reads: "Analyze first, Show the Important, Zoom, filter and analyze further, Details on demand" [23].

There are various ways to enable such an iterative approach of finding relevant information. For example, filters can be applied to the given graph that allow users to remove irrelevant information or change the relevance of certain data. Such filters can be seen as a context or a lens through which the graph is seen.

An important aspect of exploring data is the representation of *associations* that come from the user's mind. In this perspective, filters are more abstract than direct associations between informational units. A user may have certain relations in his or her mind, however, filters do not allow to associate those directly. They are applied to a given data set and, thus, are "outside" the resulting information space (that contains Facebook data in our case). As such, we consider filters as an abstract context or "layer".

Our approach is a tool that enables users to associate Facebook data and to create a context in which further data is presented at the same time. In fact, the result of the structuring process *is* the context in which relevant further data gets presented. The underlying concept of users associating information has its origin in hypertext.

From its very beginning [13, 30, 15], hypertext always has been considered a medium for representing human associations, a medium for humans to express their (interconnected) ideas.

The most common structure type is based on nodes and links, as it is widely known from the World Wide Web: nodes (i. e., websites for the WWW) are interconnected using links (i. e., URIs pointing to other websites). This type is also called "navigational hypertext", as it allows users to navigate nodes by activating links.

What does navigational hypertext mean in the context of exploring Facebook data? With the ability to use links, users would be able to associate data and, thus, create a context of data or information that belong together. Suggestions can be computed based on this context. The draw-

5 https://www.graphviz.org

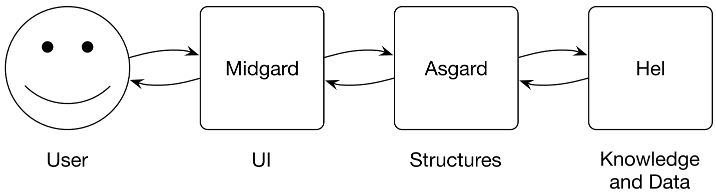

Figure 1: Architecture of the Mother system.

back is that each link must be explicitly created and possibly modified. This is unfortunate for evolving structures, as those exist in the case of explorational browsing Facebook content.

Besides navigational hypertext, there have been other types discussed in the past, including taxonomic hypertext [32] or argumentation supporting structures [14], which target specific tasks at hand. A special type is *spatial hypertext* [27], which follows a *desk-on-table* metaphor: similar to physical paper notes on a desk, objects can be moved on a 2D canvas. Associations between nodes are encoded by their position, arrangement, distance, size, color, orientation, or other visual cues. The associations are implicit: they appear by interpreting the space. In order to make the system "understand" the user-created structure, so-called parsers compute associations based on spatial distance ("spatial parser"), visual appearance ("visual parser"), temporal user interaction ("temporal parser"), or content similarities ("content parser") [35].

The big advantage of using spatial hypertext is that structures can be created and modified at ease and at a low cognitive load [37]. As such, spatial hypertext helps in exploring unknown knowledge by supporting creating, modifying, or destroying contexts of information to which the system reacts to. Spatial hypertext also helps dealing with the evolving nature of the user's Facebook data context: "One of the strengths of spatial hypertext is its ability to support the expression of evolving interpretations." [36]

The focus on the user-created context as the target of a user's structuring action makes spatial hypertext an appropriate medium for the task at hand. It supports the evolving nature of the structure and reduces cognitive overhead during the structuring process. Furthermore, using specialized parsers the spatial structure can be used as a basis for suggesting additional relevant Facebook data.

3.2 Mother

Our system *Mother* is a component-based open hypermedia system (CB-OHS) [5]. As such, it provides an infrastructure for multiple structure services (in particular the spatial structure service, including its various parsers), knowledge bases, or front-end applications running on various devices. Mother consists of three basic layers [6] – see Figure 1:

1. *Hel,* which includes all *knowledge-based components*
2. *Asgard,* which hosts Mother's *structure components*
3. *Midgard,* which includes all *user interfaces* or components that have similar functionalities or purposes

Mother provides, among others, GUI applications for desktop computers (written in JavaFX) and for Android. It currently uses various structure/annotation components, such as a spatial structure service, a metadata service, or a link service. Components in Mother can be easily addressed or combined. For example, the linking services makes use of the metadata service in order to comment or label link endpoints [34].

Mother's spatial structure service includes parsers that are capable of interpreting the user created structure. The result is a weighted graph in which edges represent relationships between informational units. Mother uses this structure to query so-called *suggestion nodes* from the machine's knowledge bases. Such suggestions are displayed around the user generated context in a semantically meaningful way. This opens the possibility for an interactive workflow, in which the user fills and modifies the knowledge space step by step to which the machine responds with relevant additional information.

The user may chose to pick any of the system's suggested nodes. This would transform the node to be a user node: it becomes an element of the user-generated context. As such it is considered for upcoming parser runs. In another perspective, creating the context by adding selected items suggested by the system from its knowledge base lets the user browse an (unknown) information space.

With respect to the application domain of exploring social media data, Mother provides:

1. A graph-based structure with information units taken from the Facebook dump and connected with weighted edges (a detailed description of the Facebook dump is provided in Section 4.1)
2. A spatial hypertext UI that allows users to create contexts as a result of their exploration of data

3. The parsers that compute a weighted graph from the implicit information given by the user, which is used for querying the knowledge graph

In the following section, we will discuss design considerations and decisions for instrumentalizing Mother as a spatial hypertext tool that allows users to explore and reflect upon their social media behavior.

4 From Raw Data to Visual Exploration

The GDPR obliges platforms such as Facebook to offer their users a copy of their personal data. In this section, we explain what this data looks like and how knowledge – in the form of topics and connections between these topics – can be extracted from these logs. In order not to overload users, several design principles and decisions are needed to decide which temporal and topical relations to take into account.

4.1 Description of the Facebook Data Dump

Following the GDPR regulations, Facebook allows its users to obtain a copy of their personal data as a simple download,[6] either in (human-readable) HTML format or in (machine-readable) JSON format. In both cases, the user receives an archive with files that contain, among others, the user's own posts, comments, likes and reactions to posts of others (both friends and Facebook pages), search history, lists of friends, subscribed groups and pages, and interaction with advertisements.[7]

Upon first inspection, it becomes apparent that the provided user data is strictly limited to the data provided by this specific user: users do *not* receive other users' comments or likes on their posts, nor do they receive the content of their friends' posts that they liked or commented on. As the social interactions between posts, likes and comments are an important ingredient of Facebook's algorithm,[8] this implies that the data dump cannot be used for better understanding the inner workings of Facebook.

Still, the textual contents of the posts as well as post frequency statistics would provide rich material for users to obtain insight in and to reflect on their Facebook usage, including the reporting of life events, work-related announcements, discussions with friends, shared silly pictures or memes, and interaction with advertisements. However, particularly for active Facebook users, the lengthy, chronologically ordered lists of posts is not directly useful, as it does not allow users to recognize overarching themes and their content-related, associative and temporal connections.

4.2 Knowledge Extraction

As a first step towards a visualization, we created a script to process the JSON Facebook post data of a user into a graph-based format, with posts, keywords, months and years as vertices, connected by edges with various weights. The keywords are extracted from the Facebook posts, converted into lowercase, lemmatized, stopwords removed and only keywords that appear in at least 5 posts are stored, in order to keep the number within limits. Edges between posts, years, months and keyword were created and weighted based on tf-idf (term frequency–inverse document frequency) and/or co-occurrence.

These vertices and edges served as a basis for experimenting with several configurations of word and post visualizations and their connections.

4.3 Exploratory User Interface

The principle of user interaction with Mother is that one selects a single entity (in our case, a Facebook post), which is then displayed along with *related entities* (other posts and keywords) as recommendations that can be followed in order to create a narrative. An example is shown in Figure 2.

As explained in Section 3, the visualization of related entities is inspired by the concept of spatial hypertext, where associations and thus context is encoded by visual means. Proximity, color, size and shape are examples for such clues. Our implementation focuses on proximity, as decoding positions of elements is a very accurate perceptual task for humans [26]. Recommended entities can be selected (using click or touch) and subsequently will be permanently added to the 2D workspace.

All selected entities form the visual context of the application, which is used to lay out recommendations as described and to search for them, with help of spatial parsing

6 https://www.facebook.com/settings?tab=your_facebook_information

7 For a full overview, see https://www.facebook.com/help/930396167085762?helpref=uf_permalink

8 See, e. g., https://blog.hootsuite.com/facebook-algorithm/

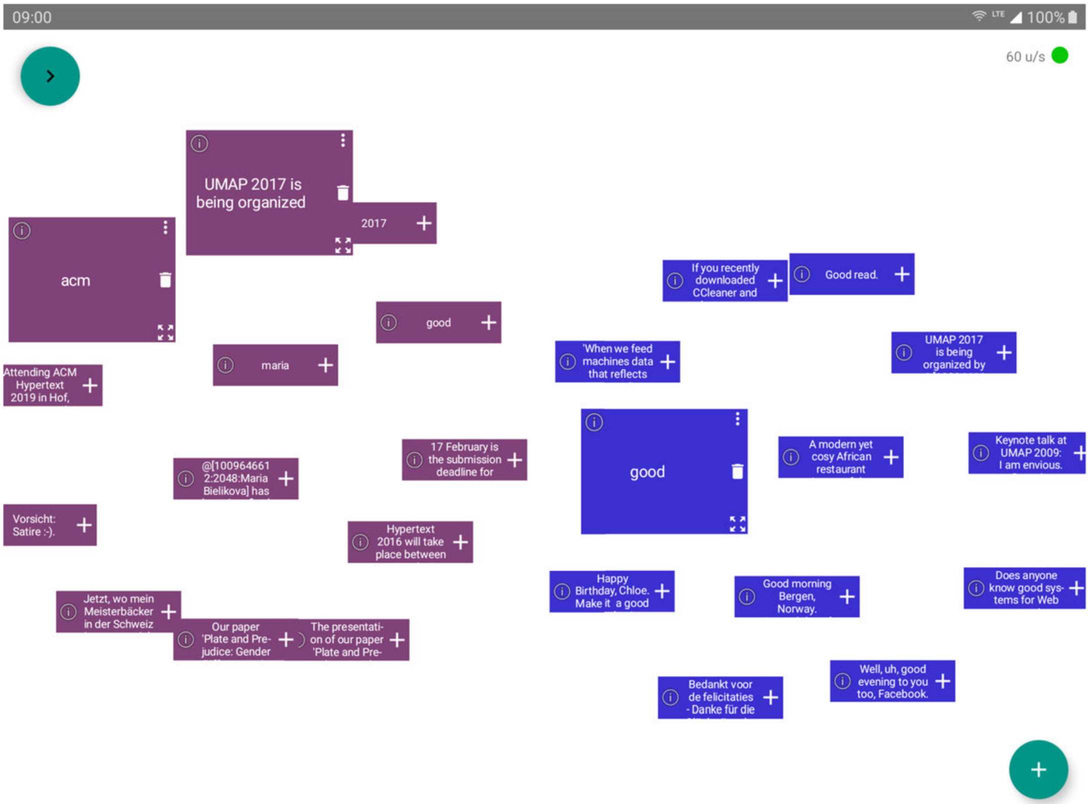

Figure 2: Screenshot of the prototype. One post and two keywords are added. Related entities surround the selected posts and words by means of attraction and repulsion, and two spatial clusters are created based on content similarity.

(cf. Section 3). The (re-)positioning of recommendations happens while the user is interacting with the system.

Interaction, in the form of moving, resizing and adding/deleting selected nodes, is an important part of the visualization:

1. Temporal user interaction is analyzed by a temporal parser, which refines the result of spatial parsing; thus the system gets a better understanding of the space and comes up with better matching recommendations.
2. Recommendations adapt to a changed context by changing their position or by appearing/disappearing. These changes give the user a further feedback channel to the otherwise restricted two dimensions of the work space.

The movement of the recommendations is handled by means of a physics metaphor, where the distance of entities is controlled by attraction and repulsion among them – as explained in more detail in [33]. The coloring, as depicted in Figure 2, is controlled by the application and shows the user how the space is interpreted by the spatial parser. While it calculates an interval scale, the result can be broken down to a nominal label: either two entities are *related* or *not related*. Entities that are labeled as related share the same color, the matching recommendations are colored accordingly.

The first developed application area of Mother concerned the movie domain, where movies are connected with one another through actors, genres and other entities [5]. These tight relations allowed users to discover and explore their own areas of interest. Similarly, our first visualization of the Facebook domain – making use of the authors' own Facebook profiles – recommended and displayed the posts that were content-wise closest to the selected post. Content-wise this approach made sense, but resulted in relations that were too obvious for the user (e. g., birthday wishes were related to other birthday wishes). After all, there is a difference between exploring an unknown domain, and *introspection*, the examination or observation of one's own mental and emotional processes; in this process, the most meaningful connections are the ones that are still meaningful, but not entirely obvious.

For this reason, we decided to only recommend posts with a content similarity of $0.6 \pm x$, with the remaining weight calculated by a sum of temporal similarity (post in the same hour of day, day of week, month), manually tuned and evaluated by the authors in several sessions. Furthermore, we also added the keywords as vertices and related them to the posts. Given the large differences in posting behavior, even between the authors, in terms of frequency and content, it was concluded that no optimal *a priori* values could be found. Instead, we opted for a configuration that led to the first author's observation that he typically posted his reflections on a day in the early evening, along with some typical themes of these reflections; in addition, some randomness was added in order to prevent users to get locked in a small number of favorite themes.

5 Exploratory Evaluation

The development and iterative refinement of the Facebook post visualization already provided several insights in the type of relations that one would consider meaningful, interesting and relevant for introspection. As we are designing a solution for a range of foreseen user needs or wishes, good design science practice [42] is to have several iterations of design/development and (preliminary) evaluation.

In this exploratory evaluation, we are interested in finding out in which type of posts and terms users were interested and for which reasons: Would users typically try to confirm their most common patterns or themes, or would they try to discover new, surprising relations in order to better understand past events or past behavior? Would this behavior mainly be motivated by introspection and/or would users also aim to investigate possible privacy threats?

We finish the section with a comparison of user posting behavior with commenting behavior, observing that user comments are typically more spontaneous and prompted by various triggers than user posts.

5.1 Methodology

Given the qualitative research questions, we chose an exploratory, scenario-based study setup with convenience sampling. Three participants, from the age of 27 to 35, were recruited and asked to provide us with their Facebook posts, which were used for the study and finally deleted, in order to prevent privacy and security issues. The evaluation itself involved the exploration of three themes (i. e., keywords given by the participant, representing hobbies, work-related announcements or life events). We invited them to select a post, and then to further explore related posts or keywords. No explicit time limit was set for the test.

After the evaluation, we asked the participants several open questions regarding their issues with Facebook data in general, to what extent the visualization would help to obtain answers with respect to these issues, which meaningful or surprising relations they discovered, and whether they would have any other question, wishes or ideas for the Facebook visualization.

5.2 Results

Our first participant had a history of 1096 Facebook posts, posted between 2009 and now, with an average of 9 posts per month. Most of these were short announcements of one or more photos (e. g., "wonderful small things"), as reflected by the low average number of 16.3 words per post. Consequently, the recommended related posts were typically based on an overlap of one or two often used phrases. Consequently, she could not relate to many of the displayed relations. She acknowledged that this was probably because most of her posts were visual and not so much textual. As points for improvement, she mentioned the inclusion of related places, persons, photos and links. Furthermore, she noticed – similar as the first author – the many birthday wishes in her log.

The second participant had a Facebook log of 275 posts between late 2011 and late 2019, with 78 active months out of the 106 months in this period. The average post length was 25.6 words. In his posts, he mainly reported about activities of the youth organization that he is involved in. As these activities follow regular patterns and have regular relations, it was very easy for the participant to recognize the thematic clusters. However, there were also several posts that he could not remember having written and he also could not reconstruct when and why the post was written, the context remained entirely unclear.

The third participant found the visualization "exciting" and interacted with it for a long period. This participant had a history of 155 posts between July 2011 and now, with an average of 2.6 posts in each active month. With an average of 33.9 words per post, her posts were relatively long. This participant was interested to find out how her personal interests and writing style developed over time. Sometimes this led to interesting observations and explorations; she recognized a forgotten event in which she sold her study books. Some other relations remained unclear,

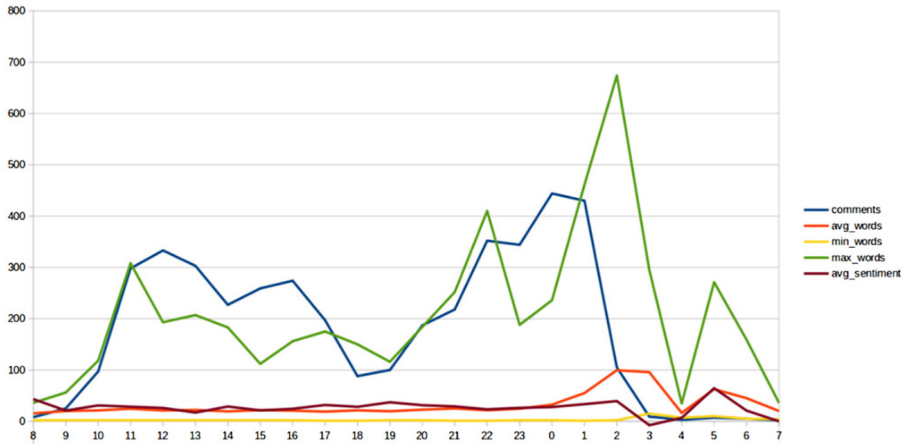

Figure 3: Aggregated daily statistics of Facebook comments, representing a period of about ten years. The x-axis represents the posting time, from 8 am to 7 am the next day.

which led to some frustration and the wish to be able to decide herself on the factors and weights for the post and word relations.

Apart from exploring one's post history, the tool also prompted participants to think about privacy-related questions, such as: what happens with a Facebook profile after one's death and where exactly are the personal data stored? Monthly statistics revealed that each participant's post behavior in terms of post frequency and average post length has remained stable and similar in the past decade.

5.3 Exploration of User Commenting Behavior

The initial design and evaluation of Mother as a spatial hypertext tool for reflecting on Facebook activities was focused on Facebook posts. The observations during our exploratory evaluation confirmed the observations in Section 4.3 about introspection activities. The tool helped participants to be reminded of earlier events, interests and occasions that they did not have in active memory anymore. Still, participants did recognize most posts – even relatively old ones – and could reconstruct the context in which they were posted.

The literature discussed in Section 2.2 suggests that comments are often spontaneous responses triggered by external factors. However, research on Facebook regrets appears to be limited to user responses to these comments, without analysing patterns in commenting behavior. A plausible reason for this is that commenting behavior is considered too sensitive for sharing with researchers.

As a first step toward understanding the characteristics of commenting behavior, the first and second author informally inspected and discussed their own comment history, visualized similarly as the posts – see Figure 2. Most comments could either be classified as brief empathic responses to friends' posts ("Same for me", "Funny", "Get well soon") or as longer stories (sometimes rants) invoked by a particular friend's or page post or a discussion in a group.

Often, the tone of the comments was observed to be more humorous – or meant to be humorous – and even more informal than the Facebook posts. We recognized that our comments might even better reflect who we are and what triggers us than Facebook posts (which are usually deliberately shared with a particular audience). At the same time, most comments were not considered relevant anymore and perhaps in hindsight even a bit embarrassing.

These observations are in line with the visualization of daily commenting behavior – aggregated from a period of more than ten years – in Figure 3. The increasing number of comments in the course of the evening, as well as the average length of comments, reflects relatively spontaneous behavior that arguably does not necessarily need to be stored, accessed or used for eternity.

6 Discussion and Conclusion

In this article, we presented the adaptation of a spatial hypertext tool, Mother, in order to allow users to explore their Facebook post and comment history. The post data can be downloaded by the users themselves, thanks to the data portability requirements of the European GDPR. Even though users can directly scroll through the posts in chronological order, it does not allow for recognizing overarching themes or relations. Mother aims to fill this gap by providing navigation via recommended related posts, based on a combination of content similarity and temporal relations.

An important lesson learned during the design of the system and exploration of the authors' own profiles is that there is a difference between exploring an unknown domain (e. g., movies) and introspective exploration of one's own activities and posts: for unknown domains, close semantic relations (such as actor X plays in movie Y) are meaningful and useful, but when exploring one's own activities, these relations turn out to be to obvious to be useful. These observations are in line with classic hypertext literature, in which it was assumed and observed that users would create (spatial) *cognitive maps* of information structures, subjective representations of the environments that one would discover, learn and traverse in a similar way as one would do when moving to, for example, a new, unknown city [10]. Despite large differences in Facebook use in terms of post frequency, post topics and post types (e. g., short vs long, text-based or photo-based), a combination of semantic closeness (excluding the closest relations) and temporal similarity (including seemingly less obvious relations, such as hour of day) delivered a rich domain that generally led to thematically meaningful clusters.

The exploratory spatial hypertext for user posts and comments has been evaluated by analyzing and discussing the authors' own data, and three participants participated in a qualitative scenario-based study. The results provide several directions for further research. First, a larger-scale study on user interactions with and observations about their own posts is expected to provide further insight in how people remember, recognize and connect life events. Further, in line with studies on Facebook 'regrets', it seems that the lifetime of certain spontaneous actions (such as liking, commenting and sharing) ideally needs to be reduced – which eventually might lead to stricter and well-founded regulations on the processing of personal behavioral data.

Finally, we believe that regular exploratory, introspective and reflective studies like the one presented in this paper are very instrumental in further shaping and reshaping our social media landscape, particularly in terms of triggers that users are expected to respond to. A particular challenge of such tools for personal visual analytics is that they need to fit in personal routines and environments [22], as otherwise they remain largely unused.

References

[1] Esma Aïmeur and Alexis Tremblay. Me, myself and i: Are looking for a balance between personalization and privacy. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, pages 115–119, 2018.

[2] Saleem Alhabash, Nasser Almutairi, Chen Lou, and Wonkyung Kim. Pathways to virality: Psychophysiological responses preceding likes, shares, comments, and status updates on facebook. *Media Psychology*, 22(2):196–216, 2019.

[3] Fatemeh Alizadeh, Timo Jakobi, Jens Boldt, and Gunnar Stevens. Gdpr-reality check on the right to access data: Claiming and investigating personally identifiable data from companies. In *Proceedings of Mensch und Computer 2019*, pages 811–814. 2019.

[4] Hind Almerekhi, Haewoon Kwak, Joni Salminen, and Bernard J Jansen. Are these comments triggering? predicting triggers of toxicity in online discussions. In *Proceedings of The Web Conference 2020*, pages 3033–3040, 2020.

[5] Claus Atzenbeck, Daniel Roßner, and Manolis Tzagarakis. Mother – an integrated approach to hypertext domains. In *Proceedings of the 29th ACM Conference on Hypertext and Social Media*, pages 145–149. ACM Press, 2018.

[6] Claus Atzenbeck, Thomas Schedel, Manolis Tzagarakis, Daniel Roßner, and Lucas Mages. Revisiting hypertext infrastructure. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media*, pages 35–44. ACM Press, 2017.

[7] Nancy K Baym, Kelly B Wagman, and Christopher J Persaud. Mindfully scrolling: Rethinking facebook after time deactivated. *Social Media+Society*, 6(2):2056305120919105, 2020.

[8] Vikram R Bhargava and Manuel Velasquez. Ethics of the attention economy: The problem of social media addiction. *Business Ethics Quarterly*, pages 1–39, 2020.

[9] David Blackwell, Carrie Leaman, Rose Tramposch, Ciera Osborne, and Miriam Liss. Extraversion, neuroticism, attachment style and fear of missing out as predictors of social media use and addiction. *Personality and Individual Differences*, 116:69–72, 2017.

[10] Patricia M Boechler. How spatial is hyperspace? interacting with hypertext documents: cognitive processes and concepts. *CyberPsychology & Behavior*, 4(1):23–46, 2001.

[11] Joseph Bonneau, Jonathan Anderson, and George Danezis. Prying data out of a social network. In *Proceedings of the 2009 International Conference on Advances in Social Network Analysis and Mining*, pages 249–254, 2009.

[12] Taina Bucher. The algorithmic imaginary: exploring the ordinary affects of facebook algorithms. *Information, communication & society*, 20(1):30–44, 2017.

[13] Vannevar Bush. As we may think. *The Atlantic Monthly*, 176(1):101–108, 7 1945.

[14] Jeff Conklin and Michael L. Begeman. gIBIS: a hypertext tool for team design deliberation. In *Proceedings of the ACM Conference on Hypertext*, pages 247–251. ACM Press, 1987.

[15] Douglas C. Engelbart. Augmenting human intellect: A conceptual framework. Summary Report AFOSR-3233, Standford Research Institute, 10 1962.

[16] European Union. Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing

of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation). *Official Journal of the European Union*, 59(L 119):1–149, 5 2016.
[17] Shion Guha, Eric PS Baumer, and Geri K Gay. Regrets, i've had a few: When regretful experiences do (and don't) compel users to leave facebook. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*, pages 166–177, 2018.
[18] Eelco Herder. The need for identifying ways to monetize personalization and recommendation. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, pages 291–294, 2019.
[19] Eelco Herder, Claus Atzenbeck, and Daniel Roßner. Hypertext as a tool for exploring personal data on social media. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media*, 2020.
[20] Eelco Herder and Boping Zhang. Unexpected and unpredictable: Factors that make personalized advertisements creepy. In *Proceedings of the 23rd International Workshop on Personalization and Recommendation on the Web and Beyond, ABIS'19*, pages 1–6. ACM, 2019.
[21] Itai Himelboim, Marc A Smith, Lee Rainie, Ben Shneiderman, and Camila Espina. Classifying twitter topic-networks using social network analysis. *Social media+society*, 3(1):2056305117691545, 2017.
[22] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):420–433, 2014.
[23] Daniel Keim, Gennady Andrienko, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. Visual analytics: Definition, process, and challenges. In Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North, editors, *Information Visualization, volume 4950 of* Lecture Notes in Computer Science, pages 154–175. Springer Berlin Heidelberg, 2008.
[24] Daniel Keim and Leishi Zhang. Solving problems with visual analytics: Challenges and applications. In *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, pages 1:1–1:4. ACM, 2011.
[25] Cheonsoo Kim and Sung-Un Yang. Like, comment, and share on facebook: How each behavior differs from the other. *Public Relations Review*, 43(2):441–449, 2017.
[26] Jock Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2):110–141, 4 1986.
[27] Catherine C. Marshall, Frank G. Halasz, Russell A. Rogers, and William C. Janssen. Aquanet: a hypertext tool to hold your knowledge in place. In *Proceedings of the 3rd ACM Conference on Hypertext*, pages 261–275. ACM, 1991.
[28] Emmanuel Mogaji. This advert makes me cry: Disclosure of emotional response to advertisement on facebook. *Cogent Business & Management*, 3(1):1177906, 2016.
[29] Kevin Munger, Mario Luca, Jonathan Nagler, and Joshua Tucker. The (null) effects of clickbait headlines on polarization, trust, and learning. *Public opinion quarterly*, 2020.
[30] Theodor Holm Nelson. Complex information processing: a file structure for the complex, the changing and the indeterminate. In *Proceedings of the ACM 20th National Conference*, pages 84–100. ACM Press, 1965.
[31] Jahna Otterbacher, Chee Siang Ang, Marina Litvak, and David Atkins. Show me you care: Trait empathy, linguistic style, and mimicry on facebook. *ACM Transactions on Internet Technology (TOIT)*, 17(1):1–22, 2017.
[32] H. Van Dyke Parunak. Don't link me in: Set based hypermedia for taxonomic reasoning. In *Proceedings of the 3rd Annual ACM Conference on Hypertext*, pages 233–242. ACM Press, 1991.
[33] Daniel Roßner, Claus Atzenbeck, and Tom Gross. Visualization of the relevance: Using physics simulations for encoding context. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media, HT'19*, pages 67–76, 2019. Association for Computing Machinery, New York, NY, USA.
[34] Daniel Roßner, Claus Atzenbeck, and Daniel Urban. Weblinks: Augmenting web browsers with enhanced link services. In *Proceedings of the 3rd Workshop on Human Factors in Hypertext (HUMAN'20)*, 2020. In print.
[35] Thomas Schedel and Claus Atzenbeck. Spatio-temporal parsing in spatial hypermedia. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media*, pages 149–157. ACM, 2016.
[36] Frank M. Shipman, Haowei Hsieh, Preetam Maloor, and J. Michael Moore. The Visual Knowledge Builder: a second generation spatial hypertext. In *Proceedings of the 12th ACM Conference on Hypertext and Hypermedia*, pages 113–122. ACM Press, 2001.
[37] Frank M. Shipman, J. Michael Moore, Preetam Maloor, Haowei Hsieh, and Raghu Akkapeddi. Semantics happen: Knowledge building in spatial hypertext. In *Proceedings of the 13th Conference on Hypertext and Hypermedia*, pages 25–34. ACM Press, 2002.
[38] Eva Thelisson, Kshitij Sharma, Hanan Salam, and Virginia Dignum. The general data protection regulation: An opportunity for the hci community? In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2018.
[39] Sarah Turnbull and Simon Jenkins. Why facebook reactions are good news for evaluating social media campaigns. *Journal of Direct, Data and Digital Marketing Practice*, 17(3):156–158, 2016.
[40] Yang Wang, Pedro Giovanni Leon, Xiaoxuan Chen, and Saranga Komanduri. From facebook regrets to facebook privacy nudges. *Ohio St. LJ*, 74:1307, 2013.
[41] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro Giovanni Leon, and Lorrie Faith Cranor. "I regretted the minute i pressed share" a qualitative study of regrets on facebook. In *Proceedings of the seventh symposium on usable privacy and security*, pages 1–16, 2011.
[42] Roel Wieringa. Design science as nested problem solving. In *Proceedings of the 4th international conference on design science research in information systems and technology*, pages 1–12, 2009.
[43] Gunwoo Yoon, Cong Li, Yi Ji, Michael North, Cheng Hong, and Jiangmeng Liu. Attracting comments: digital engagement metrics on facebook and financial performance. *Journal of Advertising*, 47(1):24–37, 2018.

Bionotes

Eelco Herder
Radboud Universiteit Nijmegen, Nijmegen, the Netherlands
eelcoherder@acm.org

Eelco Herder is an assistant professor at the Digital Security Group at Radboud Universiteit Nijmegen, the Netherlands. His research focuses on the fine balance between the benefits of personalization and perceived and actual risks associated with privacy matters.

Daniel Roßner
Institute of Information Sciences, Hof University, Hof, Germany
daniel.rossner@iisys.de

Daniel Roßner is a research fellow in Visual Analytics at the iisys Institute of Information Systems at Hof University, Germany.

Claus Atzenbeck
Institute of Information Sciences, Hof University, Hof, Germany
claus.atzenbeck@iisys.de

Claus Atzenbeck is a professor at the iisys Institute of Information Systems at Hof University, Germany, where he leads the Visual Analytics Research Group.