

# ‘Datafied’ Reading: Framing Behavioral Data and Algorithmic News Recommendations

Jesse Haapoja

Department of Computer  
Science  
Aalto University  
Espoo, Finland  
jesse.haapoja@aalto.fi

Airi Lampinen

Department of Computer and  
Systems Sciences  
Stockholm University  
Stockholm, Sweden  
airi@dsv.su.se

## ABSTRACT

There are increasing concerns about how people discover news online and how algorithmic systems affect those discoveries. We investigate how individuals made sense of behavioral data and algorithmic recommendations in the context of a system that transformed their online reading activities into a new data source. We apply Goffman’s frame analysis to a qualitative study of Scoopinion, a collaborative news recommender system that used tracked reading time to recommend articles from whitelisted websites. Based upon ten user interviews and one designer interview, we describe 1) the process through which reading was framed as a ‘*datafied*’ activity and 2) how behavioral data was interpreted as socially meaningful and communicative, even in the absence of overtly social system features, producing what we term ‘*implicit sociality*’. We conclude with a discussion of how our findings about Scoopinion and its users speak to similar issues with more popular and more complex algorithmic systems.

## Author Keywords

Behavioral Data; Datafication; Algorithmic System; Frame analysis; Recommender systems; Online Journalism

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

Using information technology to transform social behavior into quantified data is referred to as *datafication* [8]. This phenomenon has been argued to have led to the digital quantification of many aspects of social life, such as friendships as performed on Facebook, expressions of tastes for personalized marketing and media recommender systems [8], and health [35, 21]. Algorithms play an increasingly

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

*NordiCHI’18*, September 29–October 3, 2018, Oslo, Norway

© 2018 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6437-9/18/09...\$15.00

<https://doi.org/10.1145/3240167.3240194>

important role in our everyday practices [44], and more and more of our daily behavior is tracked so that it can serve as a data source. While the term “datafication” has been scarcely used in the field of Human–Computer Interaction (HCI), tracked behavioral data, including their potentials and troubles, have long been a key concern for our community.

Receiving news is one practice that is transforming due to datafication and algorithmic systems that rely on behavioral data. There are growing concerns about how people discover news, and how algorithmic systems play into these discoveries. We approach this theme by documenting the story of Scoopinion, a collaborative news recommender system that relied on tracked reading data (in particular how long a user spent on an article) to recommend ‘*stories worth reading*’ from whitelisted websites. Scoopinion is now defunct, having shut down in 2016 after failing to become financially sustainable. We turn to it as an illustrative, early example of an algorithmic system that abandoned click-based recommendation to resist clickbait (Oxford English Dictionary defines clickbait as “(on the Internet) content whose main purpose is to attract attention and encourage visitors to click on a link to a particular web page.”) and promote longer-form articles. Approaches that focus on time spent on content instead of mere clicks have since been adopted also by, for instance, Facebook [5], and Youtube [45].

Research on algorithmic systems has, so far, had a major focus on widely-used systems, such as Facebook, Uber, Twitter and Google. By focusing only on the giants of contemporary platform economies, our community risks failing to explore issues that do not appear in their context and missing the opportunity to consider what we might learn from the struggles of alternative systems as they try to survive over time. Here, our focus on a less prominent and less complex system – and the visibility into its logics that access to its lead designer provided – made it possible for us to study reasoning behind design choices regarding the proverbial “black box” and document an account of *why* the service was designed as it was.

Our qualitative study illustrates how both those using and those developing this particular algorithmic system made sense of how it turned online reading activities into behavioral data as well as how they drew on wider cultural

meanings to form shared understandings of the system and its mission to counteract clickbait culture. We complement user interviews with an interview with Scoopinion's lead designer. This was conducted after the service closed so as to reflect on the story of the system after the fact. We use Erving Goffman's *frame analysis* [18], a micro-sociological approach for studying how individuals make sense of experiences in their daily life and how everyday experience is organized by individuals, as a theoretical approach. Our findings depict how participants came to understand reading as an activity that can be transformed into data that meaningfully captures something of the motivations for reading. In other words, we illustrate how reading got to be seen as a 'datafiable' activity. Secondly, we describe how reading, once datafied, became a way of implicitly engaging with other Scoopinion users, even as there were no overtly social features in the system, such as user profiles or tools for 'liking' or commenting content.

Our study contributes to the literature on algorithmic systems within HCI in two ways. First, we advance understanding of datafication by offering an empirical example of how the cultural practice of reading was 'datafied' in the context of Scoopinion, covering the points of view of both the system's lead designer and its early users. Second, we contribute with the notion of *implicit sociality* that relates to how behavioral data was interpreted as socially meaningful and communicative, even in the absence of overt social cues in the system. Instead of showing explicit traces of other users' behavior and thus enabling social navigation [7], the framing of datafied reading and the implicit sociality related to it allow interpreting aggregated, algorithmic recommendations similarly as explicit information about other's tastes.

## BACKGROUND

We now introduce research on *algorithmic systems* and *datafication*, both from the emerging domain of critical algorithm and data studies and from HCI scholarship. We, then, discuss our theoretical background, especially Erving Goffman's concept of *frames*, and how it has been used in the study of socio-technical systems.

### Algorithmic systems

In recent years, algorithms have gone from being a basic concept in computer science to a term that gets actively interrogated and debated in the social sciences and humanities, and even in broader public debates. The HCI community has taken actively part in these conversations about the roles and implications of algorithms, as evidenced, for instance, by a panel on the power algorithms have over humans at the CHI conference in 2016 [26] and a workshop on algorithms and work at the CSCW conference in 2016 [42]. Moreover, recent work within HCI has looked at questions such as algorithmic management of Uber drivers and their sense-making of the system [25], and Facebook's capabilities to manipulate emotions by tweaking the algorithms that produce its Newsfeed [19, 23].

Our everyday life is increasingly mediated by *algorithmic systems* [44]. We use the term algorithmic systems, rather than just algorithms, since the former includes other meaningful parts of the system, too, such as human actors and data. As Dourish [10] points out, algorithms need to be understood as relational beings: they are hard to understand in separation, without considering the data that they use. In social scientific research, significant effort has been placed on making sense of and critiquing the promise of objectivity and the power that algorithms have [15, 22]. Kitchin [22] calls for studies of algorithmic systems that analyze how they act in the world and how they are designed, in order to understand the nature of algorithms and the work they do in contemporary life. Neyland and Möllers [31] remind us that the assumed social power of algorithms only makes sense in light of how such power is constituted through associations between rules, people, processes and relationships.

Prior research on users' experiences and views of algorithmic systems has focused, for example, on how Facebook's algorithms are understood [12, 13, 33], how users may modify their behavior in response to what they know (or believe) about a service's algorithms [12] and on trust in algorithmic management [24]. Although useful for many purposes, algorithmic systems do, at times, make predictions about our interests that may seem faulty or feel emotionally intrusive [4]. Bucher [4] gives an example of this in documenting how a Facebook user was reminded of his recently deceased daughter in his 'Year in Review'.

Moreover, algorithmic systems need to be understood in the context of the economic obligations they may have [15]. For example, if a service or an app is to "stay alive", it typically needs to find a business model to cover its costs or turn a profit – it may not be enough that those adopting a system enjoy it or find it beneficial. Those who create algorithmic systems have to make choices regarding their functioning but in doing so, they are rarely free of (economic) constraints.

Research on algorithmic systems can be seen as a continuation to, or at least in parallel with, the vast body of work on *recommender systems* where such systems have been developed and studied for decades [34, 37]. This genre of algorithmic systems is focused on the creation of tailored recommendations with the help of varied data sources, typically collected from prior interactions with the system [3]. Our interest here is not on recommender systems per se, even though Scoopinion belongs to this category.

### Datafication

*Datafication* refers to the process of using information technology to transform social behavior into quantified data [8, 29]. Giving credit on the term's origin for Mayer-Schoenberger and Cukier [29], Van Dijck describes datafication as "*the transformation of social action into online quantified data, thus allowing for real-time tracking and predictive analysis*" [8, p. 198]. Further, she argues that datafication is gradually becoming normalized both in science and wider society, with researchers hoping to gain

new insights on data sets collected from online behavior and with consumers exchanging their data for services.

According to scholars in the emerging field of critical data studies, *datafication* has led to the digital quantification of many aspects of social life, such as friendships, interests and expressions of tastes [8] as well as even health where research has often focused on the possibilities to collect data from users with the aim of feeding it back to them in order to change their behavior in some desired way [34, 20, 21].

While the term datafication has been less commonly used within HCI, it seems obvious that the phenomenon itself is a central concern for our community. This is evidenced, for example, in research on services that rely on mobile sensing [e.g. 2] and in studies on different health applications that track, quantify, and visualize behavior [e.g. 20, 21]. HCI researchers have recently taken on further related topics, including but not limited to the design of data collection and aggregation [14], human-data interaction (which refers to how people engage with data that is deemed personal) [6], and the role of data in so called “smart cities” [30].

### **Goffman’s conceptualization of frames**

Theoretically and methodologically, we build on Erving Goffman’s work, in particular his conceptualization of *frame analysis*. Manning [27] has characterized Goffman as a theorist of credibility: frame analysis can be seen as a method of understanding the circumstances that make experiences feel real. Credibility encompasses two themes central to Goffman’s work: trust and deception. Manning [27] notes that much of Goffman’s writing shows that the methods for fostering trust are the same as those used for deception.

Goffman [18, p. 8-10] conceptualized *frames* as a way to organize meanings individuals relate to any given situation, providing answers to the question “*what is it that’s going on here?*” In other words, people observe and interpret activities by situating them in frames, and without these frames, they do not understand what they observe. Goffman [18, p. 85] specifies that frames are not something that are either ‘in someone’s head’ or ‘out there to be perceived’ but that they incorporate both participants’ responses and the world they are responding to. In brief, frames are relational rather than purely cognitive. In Goffman’s conceptualization, frames can be divided into two primary classes: *natural* and *social frameworks*. Natural frameworks refer to events such as rain, in which no willful agent is present. In this paper, we focus on social frameworks. These refer to activities that involve intentional actors, such as a game of chess [18].

Framed activities can be *keyed*. A key, in this context, refers to a “*set of conventions by which a given activity, one already meaningful in terms of some primary framework, is transformed into something patterned on this activity but seen by the participants to be something quite else*” [18, p. 41]. As a classic example, play fighting is a keying of fighting – one that is quite different from fighting “for real”.

Keying, then, changes the meaning of activity, typically by adding a new shade of meaning so that the established frame is shifted and due to that another direction emerges.

Keyings can be rekeyed, too, adding another layer to the activity and thus creating a nested and processual structure of experience [18, p. 82]. To illustrate a nested structure, Goffman gives the example of a play within which characters are rehearsing a fictional play. We use the idea of experience as *layered* and *processual* to guide our analysis. To connect the notion of frames to our concerns about how users and designers make sense of socio-technical systems, it is important to note that any given situation can be framed in multiple ways [18]. As a result, for service designers, there can exist preferred user understandings that align with the goals of service stakeholders [15], as well as alternatives that do not match their goals or even actively contest them. Decisions made regarding the aims of service design create, more or less intentionally, a certain definition of the service and the situation of its usage.

The notion of framings resembles Eglash’s [11] definition of reinterpretation of technology where the semantic association of technology differs from the intended. While literature on *appropriation* [9, 11, 36] sheds light on the different meanings technology may have for different people, frame analysis can help us understand how the development of socio-technical systems is embedded in the larger system of shared meanings in the surrounding society. For instance, understandings of being friends on Facebook build on established meaning of friendship, although it does not map on to it 1:1.

We are not the first to apply the concept of *frames* in the study of information systems. For instance, Orlikowski and Gash [32] used the concept to study how different members of an organization understood a new groupware system implemented in the company. Moreover, frame analysis has often been used with the aim of identifying and naming different frames [28, 32]. As a departure from this, often cognitively oriented approach, we focus on the *processual nature of framing* by examining micro-level processes of organizing experiences instead of naming different frames.

To be clear, our aim with this paper is not to contribute to the literature on frame analysis, but to build on it as an analytical approach that guides our study of an algorithmic system and its relationship to more widely shared meanings (such as understandings of reading).

### **CASE TECHNOLOGY: SCOOPINION**

Scoopinion was a personalized news recommender system that focused primarily on promoting longer-form magazine articles. Our description builds on our observations of the system, informal discussions with the Scoopinion team, and a formal interview with Scoopinion’s lead designer.

Scoopinion recommended feature stories from online magazines and news sites based on behavioral information collected by tracking users’ reading habits via a browser

plug-in. The plug-in monitored and collected information about user behavior on whitelisted sites of online newspapers and magazines. Tracked behaviors included time spent on the site, scrolling patterns, reading speed, and read-through rates. Scoopinon used the aggregated reading behavior data from its users to recommend stories that it deemed engaging and immersive based on behavioral metrics.

Scoopinion compared the reading habits of its users with each other and recommended material from others sharing similar reading habits. That is, if other people who shared similar habits as a particular user had read a certain article, the algorithm would offer it to that user, based on the assumption that the user might be interested in articles that others with similar reading patterns chose to read “well”. Recommendations were delivered via the system’s website (Figure 1.) and via e-mail digests. The website included both personalized content targeted for the particular user and non-personalized content, such as a list of the most popular articles among all users.

Scoopinion got its start when the team behind it participated in an annual Finnish media innovation competition called Uutisraivaaja<sup>1</sup>. The competition rewards its winners with funding for innovative (technology) projects in the field of media. The team was driven by critical views on internet media and an appreciation of journalism that contributes to the society by providing high quality information for citizens on what is happening in the world. The service has since shut down as the team was unable to make the service profitable.

At the time of our user interviews, Scoopinon was focused on anonymous crowd filtering of long-form stories. There was no explicitly social functionality, such as user profiles or messaging features. Rather, Scoopinon simply tracked users’ reading activities on whitelisted sites and then served recommendations based on the resulting behavioral data. The interface included neither explicit traces of other people, nor any means of contacting other users.

## MATERIAL AND METHODS

We adopted qualitative methods in order to gain insight into 1) how Scoopinon users understood the system and interpreted the recommendations they received and 2) how the system was understood and promoted by the team

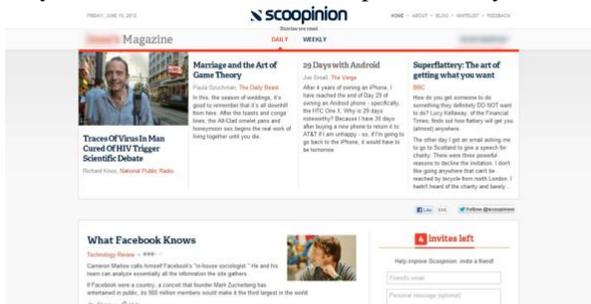


Figure 1. Picture of Scoopinon’s personalized interface

developing it. Our research materials consist of ten semi-structured individual interviews with Scoopinon users and one designer interview. After analyzing the user interviews initially with grounded theory methods [16], we went deeper with the help of frame analysis [18], considering them in conjunction with the designer interview.

## User Interviews

We employed *purposive sampling*, that is, non-random sampling of individuals, groups, or settings where the processes and experiences of interest are most likely to be observed [40]. The prerequisite for participation was that participants had used Scoopinon at least for a year at the time of being interviewed. The founders of Scoopinon facilitated recruitment by contacting potential participants to check their willingness to participate in a research project related to the service, conducted by the authors. The first author then received the contact details of those willing to participate in order to set up an interview.

The first author interviewed ten Scoopinon users in July and August 2012. We decided to revisit these interviews now that the topic of algorithmic systems and datafication has become newly topical and now that it is possible to reflect on the story of Scoopinon on the whole. The participants, all of whom lived in the Helsinki metropolitan area and were native Finnish speakers, had joined Scoopinon soon after it first launched in 2011, and all of them had some kind of a personal connection to the Scoopinon team. Three participants were women, the other seven were men. Participants’ ages ranged from 25 to 34 years. All had either completed a higher education degree or were students in higher education. All participants reported being daily users of diverse information technologies.

Seven participants had the Scoopinon browser plug-in installed at the time of the study. Of the other three, one had never installed it, another had uninstalled it due to technical issues, and the third one reported having installed it in a browser he used infrequently. This had an effect on the data Scoopinon could collect from these three participants and how recommendations could be made for them: the browser plug-in was necessary for tracking reading behavior and the system used the resulting data to make recommendations. Were the browser plug-in not installed, Scoopinon only collected data on the links that the user opened through its website or in e-mail digests, resulting in a constrained set of information about the user and his/her reading activities.

Our interview protocol was semi-structured. Participants were asked questions about their media consumption, their experiences of using Scoopinon, and their views of the current system and the changes that the service had gone through while they had been using it. To ensure informed consent, participants were informed that the interviews are for a study, that their personal information will be kept

<sup>1</sup> <http://www.uutisraivaaja.fi/en/>

confidential, that only members of the research group will have access to the interview transcripts, and that they could withdraw from the study at any point or choose not to answer particular questions. No compensation was offered. The interviews were conducted in Finnish, audio recorded and transcribed verbatim by the first author. The length of the interviews varied between 45 minutes and an hour. The excerpts in this paper are translations from the original Finnish transcripts. To protect the anonymity of the participants, participants (PX) are identified by numbers (in place of “X”).

### **Designer Interview**

The first author conducted an in-depth semi-structured interview with the person in charge of Scoopinion’s design. This interview took place in April 2017, after the service was shut down. The aim was to augment insights from the user interviews with an account from the design perspective. We wanted to explore how the people who build the system had intended it to be understood by users and gain a contrasting perspective to the views of the interviewed Scoopinion users. An ethical issue we needed to tackle was that the anonymity of the designer could not be guaranteed if Scoopinion were to be discussed with its real name in this paper. We discussed the issue with the designer and, upon reflecting on the matter, he preferred to forego anonymity so that the paper could document Scoopinion’s story. To avoid confusion, we acknowledge the designer by name in the Acknowledgements section.

### **Analysis Process**

Initially, the first author analyzed the user interview transcripts with a grounded theory approach [16, 41]. In line with Strauss and Corbin [41], this part of the research process was driven by a broadly defined topic of interest: experiences related to online news recommendations. We formed the more precise research questions and conducted the literature review only after this initial analysis. The analysis, first, involved going systematically through the interview transcripts and coding them line by line. After this the material was re-coded by comparing and combining prior codes with each other. Findings in this study mainly derive from data classified as “Interaction with the (Scoopinion) algorithm”, “Perceptions about the articles received through online social filtering” and “Sociality in Scoopinion”. Later, we approached these descriptive classes with Goffman’s idea of frames [18] in order to analyze more deeply the micro-level processes of organizing experiences.

After analyzing the user interviews, the findings were discussed among the authors and the resulting insights were used to guide the designer interview. This, in turn, was analyzed in conjunction with revisiting the original analysis in light of what was learned from the designer interview. This was done to balance and contrast user and designer views on the system and the data it collected. This led us to investigate how Scoopinion users understood the process of turning

reading data into recommendations, and how they came to experience an implicit form of sociality.

### **Limitations**

Our research materials are limited in that by choosing to target long-term Scoopinion users and, thus, having to rely on the founders to facilitate recruitment, we ended up with interviewees who had personal connections to the team. Our participants should not be taken to represent the Scoopinion user base as a whole. As the service was in an early stage at the time of the study, long-term users at that point had mostly been recruited to use the system by the Scoopinion team.

As another issue, one might consider it a weakness that the designer interview was not conducted at the same time as the user interviews, but rather several years later, only after the service had already been shut down. The user perspectives should be understood as accounts of experiencing the system in use, whereas the designer interview addresses both insights from our initial analysis of user interviews and the team’s experiences of building the service after the fact. This had the benefit of providing us with a more reflective designer account than what would have been possible while the service was still actively in development.

### **FRAMING BEHAVIOR-BASED RECOMMENDATIONS**

Our findings are organized in two main sections: First, we consider the process through which reading was framed as a ‘datafiable’ activity. Second, we discuss how reading was transformed from a relatively solitary activity to one that connects people implicitly – something that can be considered socially meaningful and communicative despite the lack of overtly social system features.

#### **Transforming reading into data**

We will, first, discuss how the Scoopinion team considered reading time as a measure of an article’s quality, its potential to match a reader’s interests, and how “well” it had been read by users. This was an attempted fix to problems rising from the perceived tendency of internet journalism to favor quick, short stories with catchy titles. The team considered this click-based culture detrimental for people’s ability to become more deeply informed citizens. Second, we turn to users’ experiences, illustrating that tracking reading data was a key concept for their understanding of Scoopinion.

#### *Framing reading as a measurable behavior*

Let us, first, consider the role of the Scoopinion team in establishing a frame of reading as a datafiable activity. In communicating with its user base, the Scoopinion team purposefully framed reading as a meaningfully measurable behavior. To use Goffman’s terms [18], reading as an activity was *keyed* by the team and the service they built: The tracked reading behavior is a keying of the act of reading, that is, a type of documentation of it. Behavioral data can be understood as a keying of the measured behavior as it is fundamentally understandable only through the activity that it represents. This closely resembles the notion of *datafication* [e.g. 8, 29] – the process of turning social behavior into numbers. Furthermore, reading as a datafiable

activity was, then, framed by the Scoopinion team as a more precise and meaningful way to represent reading, and as such preferable to click-based approaches to quantifying reading.

Based on the designer interview, the goal of the Scoopinion team was not to be able to claim that Scoopinion's plug-in would capture reading in some pure way 'as it is', but rather, to achieve a 'good enough' state where the system tracked quantifiable aspects of reading that were deemed relevant and somehow representative of the act of reading: "All the way to the end the data collection methodology was at the prototype-phase because it did what it was supposed to do, and if it had been a lot better, it would not have made the brand any better... so I did not put a lot effort on the tracking being super precise" (The designer)

It is worth noting that the Scoopinion team had implemented measures to counter situations where the tracking might produce misleading data and to block attempts to "game" the recommendation algorithm. We interpret this as a precaution to pre-empt cases where users purposefully look for weaknesses in the system to outsmart the system (and its designers) by behaving in strategic ways. The designer stated that the team had implemented timers that would, for instance, recognize situations where a page had been open for too long and where a user was going through an article unnaturally fast. The team had to consider how to write software that would capture their understanding of authentic reading as closely as possible, translating the cultural practice of reading into a form that could be tracked and measured with an algorithm. As Seaver [38] has pointed out, this kind of interpretive, cultural work is needed to create algorithmically recognizable versions of events.

#### *Framing reading time as a proxy for good quality*

The designer characterized reading time as an imperfect, yet "better than clicks" proxy for three things. These were approximately how many words of the article were read, the quality of the article, and the match between the article and the reader's interest. The team wanted the system to be "biased" towards lengthier articles as their vision was to promote articles with deeper analysis, typically longer journalistic works.

The designer considered tracking reading as a way to overcome some of the troubles with internet journalism. He explained how the team, early on, chose to create a system that would try to counter the trend of click-driven journalism: "We positioned ourselves against click-basedness in the communication [of the brand of Scoopinion]". We argue that, here, the team made use of broader cultural understandings related to reading in order to direct Scoopinion users to form an understanding of the system as a positive being in the world, a welcome alternative to systems that did not question the click-based economy of the internet.

#### *User understandings of datafication*

We now turn to consider the case from the point of view of Scoopinion's users. Overall, our interviewees did not contest the frames that the Scoopinion team promoted through their design and communications: First, they largely agreed with the idea that reading can be meaningfully captured and transformed into data. Second, they seemed to believe that the captured behavioral data represented plausibly more than just what was measured – in this case, longer reading time got interpreted as a sign of higher quality.

However, participants provided examples of how documenting reading based on time spent on articles might not capture all that they felt it was supposed to capture: "Reading longer [from Scoopinion's viewpoint] does mean something, but it might only mean that a browser window has been left open" (P4). Here, leaving a browser tab open was brought up as a possible source of erroneous data, since the system could interpret such moments to mean the user had spent plenty of time on an article even if they had actually just left a page open, idle and forgotten. As described above, the Scoopinion team had measures in place to deal with these concerns, but the users were not aware of them.

In terms of the broader story Scoopinion was telling about the value of tracking time spent on content rather than mere clicks, this framing effort seemed to have been effective: our interviewees often differentiated Scoopinion's model of tracking online reading behavior from click-based approaches to measuring the popularity of online content. In undermining metrics that are based on how many times a piece of content had been accessed, participants echoed the rhetoric that Scoopinion used to promote itself. When asked how he would explain what Scoopinion is to someone who is not familiar with the service, a participant answered in this way: "I would probably explain how it shifts from the number of clicks [an article has] generated to what people actually read, so it is a service that recommends good-quality articles from all over the world to the user" (P5).

Another interviewee gave a similar response, characterizing Scoopinion as a more trustworthy recommender since it accounts for the time readers have spent on different articles: "Maybe you could describe it like this; it chooses recommendations based on what others read and how much they [articles] have been read, but also for how long they have been read. It is not enough to just take a look at the headline, so it is a little bit more trustworthy". (P8) Here, reading is portrayed as an intentional behavior: choosing to read a particular piece is seen to involve having judged the material interesting enough to spend time on it. Thus, another layer is added to the keying of reading as a documentation: The documentation of reading behavior gets to be understood as an implicit recommendation.

Finally, for the Scoopinion users we interviewed, the experience of receiving recommendations from the service was not only about the recommendations themselves. Rather, it related closely to the act of reading and the understanding

that reading is an activity that can be datafied meaningfully, as long as the measures used are not mere clicks. In our interpretation, in describing Scoopinion as “*a little bit more trustworthy*”, the participant claimed that the more detailed behavioral data Scoopinion relied on lead to recommendations that more reliably represented people’s evaluations of the articles they had read online. Thus, the interest others had shown in an article by reading it closely was taken as an indicator of the article’s quality.

### **Communicative cues in the absence of social features**

At the time of the user interviews, Scoopinion provided users neither with explicit traces of other people, nor with any means of contacting other users via the service. Yet, the behavioral data that Scoopinion relied on was interpreted as socially meaningful and communicative.

#### *Framing behavioral data as authentic information*

Scoopinion was understood to give a more authentic representation of how people assessed articles than systems relying on click-based measurements of popularity. One participant stated that “[t]he articles [that Scoopinion recommends] have really been read, and not like the “boob news” that has received most clicks ... You can really trust that it offers articles that really are worth reading” (P9). Prior research has dealt with how authenticity is constructed in the scope of self-presentation on Last.fm [39], a service where user profiles about music-listening rely on behavioral tracking, rather than self-reported preferences. In contrast, we face here a question about the authenticity of a different activity – reading – and how its qualities are understood to be measurable and interpretable from behavioral data.

One way to understand notions of authenticity in this context is that Scoopinion explicated behavioral cues that online reading *gives off*. Goffman [17, p. 2] distinguished between the information *given* and that what is *given off*. Given information refers to what is actively communicated: for instance, introducing oneself. Information that is given off refers to that part of interaction where there seems to be less deliberation on what sort of impression gets made. One can verbally claim not to be nervous, but one’s body may hint otherwise through a trembling voice or shaky hands. Scoopinion did not ask individuals to communicate their interests and preferences explicitly. Instead, it tracked their online reading behavior and made predictions with the help of the resulting data based on the cues about the users’ preferences that it contained.

The articles Scoopinion recommended were understood to be interesting on the whole as opposed to just having a catchy headline. One participant contrasted Scoopinion’s recommendations to links that circulated on Facebook, describing the latter in a derogative way: “[They] are largely completely irrelevant news, something along the lines of “A boy caught a fish, it was big”. It’s feel like my life is wasted if I click them open” (P9). The focus on reading data made this distinction possible by revealing whether people engaged with an article for a prolonged time. Also, unlike in

the Last.fm case, the interviewed Scoopinion users expressed little interest in modifying their reading habits after installing the tracker. It was possible to request the system to “forget an article”, but users stated that they had not used the feature or had, at most, just tried it once. Since the reading data was not shared with other users in a way that could be linked back to any individual, self-presentation was not at stake here.

While reading patterns were sometimes understood to be revealing of a person’s character and the tracked behavioral data was seen as a source of authentic information, some interviewees challenged this frame by pointing out that automatic tracking of reading does not capture the context of reading: An article may be written in a way that makes it slow to read, so spending a long time reading something is not necessarily a sign of high-quality writing. For example, a participant questioned the logic of time spent on reading as a sign of quality in this way: “I think there is a difference between seeing what someone reads and if someone recommends something. What she or he has just read isn’t necessarily a good article” (P10). Another interviewee pointed out that technical problems on websites could make it look like something had been read for a prolonged time: “People have looked at that site for a longer period of time, and it [Scoopinions recommender system] isn’t based on click amounts or other trivial information, but it could be that the site has just been hard to use and someone has been on the site for a long time because of that, or the site could have spent a long time loading” (P8). As a final example, we may read not because we want to, but because we are (or feel) required to do so: A participant mentioned such a practice, explaining that sometimes “a topic does not exactly interest me, but it could be something that we could do something with [at my workplace] and so I might mark it up for reading it later” (P2).

#### *Framing recommendations as a communal accomplishment*

Overall, Scoopinion was understood to help users discover and select meaningful reading materials. Key reasons for using Scoopinion were gaining access to new materials, receiving help in ‘filtering’ worthwhile reads from the wealth of stories published online and ‘killing time’, for instance when participants were on a bus and needed something to occupy their time. Moreover, Scoopinion users knew that their reading activities were, in turn, transformed into data and that, by allowing it, they were participating in creating Scoopinion recommendations. Here, information *given off* by one’s solitary reading activities became a form of (voluntary) participation in datafication and the production of news recommendations that it enabled.

The designer explained how he sometimes felt that the people using the Scoopinion browser plug-in and thus producing reading data to the service were important not only for the data they produced per se, but also because their reading data warranted the service’s claims about shifting away from a click-based recommendation model. In this way, users made Scoopinion as a brand a better

recommender: without the users' participation, there could be no claims of using reading data from the community to produce individualized recommendations.

The designer stated that he believed that users got more out of the experience of using Scoopinion when they were both receiving personalized recommendations and helping others get better recommendations by using the browser extension. The designer discussed this idea as an example of the commons, making a comparison to the Foursquare application where checking in at different venues benefits both the individual and other users. However, as mentioned before, all users did not use the browser plug-in. This implies that many users got a "good enough" experience without the system having as much data on their preferences as would have been possible. This has similarities to many online services that rely on user contributions, one notable example being Wikipedia, where those who just benefit from the collective good far outnumber those who also actively contribute to it.

When asked how he would describe Scoopinion's strengths, a Scoopinion user explained that recommendations are in a sense a service that users provide one another, with the help of the system: "*The filtering way of looking at it, there's a huge amount of information and news-streams, that someone does a little something for you, filters and recommends good articles to you*" (P6). Here, instead of the data, understanding of reading itself was transformed: reading as an activity was no longer solitary, as the tracking added a social dimension to it. This transformation turned reading from a relatively unconnected behavior to something potentially consequential: communication to the system about material one enjoys and wants more of, as well as, implicitly, recommends to others. What might have been just time killed was, thus, turned into a contribution to the community – something of a gift.

Participants considered that articles were being recommended because they had been read by others and, second, that because of the presumably authentic nature of the data on which the recommendations were based, the service recommended high-quality articles. While Scoopinion's user interface did not show how many people had read any given article, participants discussed recommended articles as pieces of content that had interested multiple individuals. This formed the basis for experiences of engaging with the service. In explaining when to visit Scoopinion's website, one participant stated the following: "*If it seems like I can't find anything interesting from anywhere, I might go to Scoopinion to see what has interested other people more from this kind of long journalism*" (P2).

Another participant's description of the service was that "*you can discover [from Scoopinion] articles that people have considered worth reading or what they have delved into*" (P3). For this participant, the key way to frame the discovery of articles was tied to the recommendations being based in

others' implicit evaluations. Instead of considering Scoopinion recommendations in terms of the style or the topics of articles that tend to get selected, the participant highlighted how the system measures users' evaluations of what is interesting and makes recommendations based on that. It is noteworthy that the implicit technique of tracking reading behavior was an important aspect of how users made sense of their own participation in the system. Implicit evaluations, that is, acts of reading on whitelisted sites, were understood as endorsements.

Participants highlighted the importance of others' actual reading behavior as the basis for recommendations. They viewed recommendations they received through Scoopinion as articles that had been read and implicitly endorsed by others. Individual reading choices were less relevant than the emergent opinion of the community on the whole: "*It doesn't matter who reads it, instead that it's read and the article rises [in popularity] because of that, so we let the audience decide about that*" (P6). Aggregation of data from multiple readers played a key role in transforming reading as a personal activity into the communal accomplishment of recommendation. Thus, interaction with the system was understood as going further than clicking buttons on the interface. One's reading behavior itself was a form of interaction, where one allowed others to benefit from one's behavioral data, while receiving personalized news recommendations in return. Thus, the recommendations could be thought of as a communal accomplishment, not just a service produced by the recommendation system.

#### **Why do users' understandings of datafication matter?**

We are interacting with algorithmic systems more and more in our daily lives. As users of such systems, we typically have only limited possibilities to scrutinize the algorithms driving them. This was the case also with the Scoopinion users we interviewed. As one participant put it: "*I have no clue how well it [Scoopinion] chooses articles for me*" (P1). However, the users were able to make sense of the system through their understandings about the data it was said to collect. Overall, we interpret the tone of the interviews to imply that the datafication process was more important for the interviewed users than the recommendation algorithm per se. Datafication 'brought humans into the system' by capturing what was considered to be authentic information about Scoopinion users. While it was hard for interviewees to articulate what about the recommendation algorithm was valuable for them, they more readily discussed the value of the tracked behavioral data and their willingness to provide it for the system: "*When that information is used for my gain, it is okay to snoop it*" (P9).

Overall, designer and user perspectives tended to be in sync, to a point that when discussing findings of the user interviews the designer stated that based on what the first author was telling him, their "*branding had succeeded*". Scoopinion had emphasized in their marketing that they

measure reading time rather than mere clicks, and this point was central to how participants described the service, too.

## DISCUSSION

In this paper, we have focused on the story of Scoopinion, a now-defunct algorithmic system that was built to recommend news. While its popularity pales in comparison to widely used and studied systems such as Facebook, Uber, Twitter and Google, studying it gave us the opportunity to gain a holistic account of an algorithmic system, from the point of view of its lead designer and its early users.

Our findings illustrate how both those using and those developing this particular algorithmic system made sense of it and drew from wider cultural meanings to form shared understandings of the system and its mission to counteract clickbait culture. We now discuss the implications of our study for HCI, by considering, in turn, how *datafication* of human behavior is dependent on cultural understandings we share with others as well as how systems that collect and use behavioral data in an opaque way rely on trust and how *implicit sociality* between users can help establish it.

### The social origins of behavioral data

Scoopinion was based on a particular view that click-based approaches to journalism are detrimental. Accordingly, the service was built to counter clickbait and, more broadly, to challenge click-driven journalistic practices. The Scoopinion team was an early mover in challenging at the time dominant click-based practices. They did so by adopting a different way of measuring reading: they replaced clicks with time spent on content as the key measurement that drove algorithmic recommendation. The preference for time-based measurements seemed to have been internalized by the users we interviewed, too.

Participants were prone to understand the collected data to represent others' intentional acts of reading. This was key to the value they felt they could gain from the system. Others' implicit endorsement was taken to confirm that by reading recommended articles, 'time killed' was not 'time wasted'. This helps to understand why they came to view using Scoopinion as a social activity even in a situation where no explicit interaction with other users took place.

Gillespie [15] has pointed out that notions of relevancy can always be contested. While participants considered the way Scoopinion transformed their reading activity into behavioral data broadly preferable to the tracking of clicks, this is not to say that there could not have been better ways for determining the relevancy and quality of online news content. For example, even as measuring reading time might capture interest in an effective way, it might not be as successful in capturing the perceived importance of any given article.

What remains to be noted is that to function, a system like Scoopinion requires a cultural context where both reading online and measuring the time spent doing so make sense. Here, the ideas Goffman presented in Frame Analysis can

help us understand the situation [18]. For there to be a possibility to disagree on how something should be measured to capture it in an appropriate manner, there needs to be a shared cultural context where measuring itself is considered a relevant activity. Users were able to discuss why measuring time was preferable to measuring clicks since they had a culturally shared idea of what reading is. This datafication can be understood, in Goffman's [18] terms, as a form of keying – the data made sense for users only when it was considered through the activity of reading even though, once datafied, the established frame of reading shifted to encompass new meanings related to measurability and implicit endorsement. As such, in datafying human behavior, we need to account for the social origins of the resulting behavioral data. Before a practice can be datafied in a way that is meaningful to those whose behavioral data it builds on, there needs to be some shared understanding of the practice in question.

Based on our study, one significant avenue for future research is how different ways of datafying behavior can affect both the design and user experience of algorithmic systems. Next to solving the many practical and ethical difficulties related to gathering behavioral data, it is important to reflect on what (and by whom) is considered relevant or significant about the behavior that systems are supposed to measure. Compared to larger systems like Facebook, Scoopinion actively marketed their data collection practices and what they could offer to users with them. This dynamic is more complicated when it comes to larger and more complex systems where those whose behavior is datafied are prone to be the product rather than the user. While we should not be naïve about the incentives that drive datafication, being clear on what data is collected, minimizing data collection by focusing only on what is most relevant, and explaining how the data is used and what it is thought to represent could help users understand their own role in regards to different systems as well as to evaluate its implications for them personally and for the broader society.

### The role of trust and transparency in algorithmic recommendations

Much of Scoopinion's underlying functionality was hidden from its users, but this of course did not prevent them from reflecting on the system. Also, even if users had been able to peer inside the proverbial black box, it would have required specialized skills to understand the code and data that they would have found inside it. Based on our designer interview, the Scoopinion team wanted people to understand the operating logics of the system, but revealing the technicalities was not, in their judgment, the best way of accomplishing this. We consider it a strategic decision on the part of the team to emphasize behavioral data as the key aspect of the system, rather than trying to clarify in any meaningful detail how the system processed this input.

Ananny and Crawford [1] have pointed out that opening *algorithmic black boxes* may privilege seeing over

understanding. Yet, leaving parts of the recommendation process invisible necessarily limits users' potential to critique, question, or alter it [43]. This exemplifies the central role that trust plays in allowing shared frames to work. Another aspect where trust became important in our study was that Scoopinion users had no way of perceiving other users in the system. Yet, understanding recommendations from the system as based in behavioral data and as a communal accomplishment of the user base relied heavily on the users' trust that there were other users reading and contributing with their data on Scoopinion. In other words, users had to trust both that there were others using the system and that they constituted a group whose implicit endorsements were worth relying on.

Services that filter content typically make it impossible for their users to see what was filtered out of the slice of reality that the service offers to them [43]. This is both the promise and the problem of these types of algorithmic systems: On one hand, they shield us from the endless amount of online information by filtering it, but on the other hand, we might end up missing information that is important without ever knowing what choices were made for us. This calls us to consider what the meaningful limits of transparency are in any given encounter with an algorithmic system. In many cases, users of algorithmic systems are left to believe in the sincerity of developers and companies, with little means to audit the claims they are presented with. This reliance on trust enables misuse, such as false claims about systems. As Manning highlights by characterizing Goffman as a theorist of credibility [27], trust and deception are closely related. While the relationship between Scoopinion and those of its users we interviewed appeared to be generally trusting and non-contentious, this is not the case with many other systems, as evidenced, for instance, in Facebook controversies over the past decade.

One topic for future research, then, concerns situations where trust towards an algorithmic system has been lost or where it has always been lacking. Studying these more contentious cases could further illuminate how users come to trust (or doubt) services, not only in terms of security or privacy, but also in how they become convinced that services offer them what they promise and that the promised service is, in fact, relevant and ethically acceptable for them. That said, we argue for the importance of studying algorithmic systems that are less prominent and less popular than giants like Google and Facebook (and as such, less centrally sites for heated moral battles) in part because by virtue of being less complex, they are potentially more understandable not only to their users but also for researchers who wish to examine and critique them. This can help in separating out issues that have to do with their algorithmic nature from critiques that concern, more precisely, business models. In addition, we wish that our study can serve as a reminder that algorithmic services – even the current giants – are not eternal or inevitable. Thus, we need to continue making efforts to

generate knowledge that goes beyond specific systems, however culturally dominant they may seem at times.

## CONCLUSION

Our results can be summarized into two main findings: First, participants made sense of the news recommender system Scoopinion's recommendations in relation to frames, first, of online reading as a *'datafiable' activity*, elements of which can be meaningfully captured based on the time invested in it, and second, of behavior tracking as a credible way to infer preferences about reading materials. These user understandings were tied to Scoopinion team's efforts to promote the idea that reading could be meaningfully tracked and translated into a quantified form.

Second, with these frames in place, Scoopinion users approached the system as an opportunity to tap into the outcomes of other people's reading activity and, in turn, participate by providing their own data as material for algorithmic recommendations. Here, reading – often a solitary act – was turned into something that connected people implicitly and was considered socially meaningful even though no direct social interaction was taking place. Our work sheds light on *implicit sociality* as one possible reason for why people welcome algorithmic systems to their daily lives. The social connection that Scoopinion users experienced was largely accomplished through datafication.

Examining how users understand datafication and how algorithmic systems use their data, as well as designing with these lessons in mind, is increasingly important for HCI. Designers of algorithmic systems should carefully consider what data is necessary for their system, what this data is thought to represent, and how these positions are communicated to users. As we have shown, these explanations can have an important effect on users' experiences of engaging with algorithmic systems.

## ACKNOWLEDGMENTS

We thank our anonymous interviewees as well as the Scoopinion team, in particular Ville Sundberg who was interviewed for the study. We also thank Kari Vesala, Barry Brown, and the anonymous reviewers for their generous and insightful comments on the manuscript. This work was supported by funding from the KONE foundation project Algorithmic Systems, Power, and Interaction, Academy of Finland grant 295950, and the Swedish Foundation for Strategic Research (RIT15-0046).

## REFERENCES

1. Mike Ananny and Kate Crawford. 2016. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 1-17.
2. Paul M. Aoki, R. J. Honicky, Alan Mainwaring, Chris Myers, Eric Paulos, Sushmita Subramanian, and Allison Woodruff. 2009. A Vehicle for Research: Using Street Sweepers to Explore the Landscape of

- Environmental Community Action. *ACM Transactions on Computer-Human Interaction*: 10.
3. Shlomo Berkovsky, Tsvi Kuflik, and Francesco Ricci. 2008. Mediation of user models for enhanced personalization in recommender systems. *User Modeling and User-Adapted Interaction*, 18, 3: 245-286.
  4. Taina Bucher. 2017. The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 20, 1: 30-44.
  5. Robyn Caplan and Danah Boyd. 2018. Isomorphism through algorithms: Institutional dependencies in the case of Facebook. *Big Data & Society* 5, 1.
  6. Andy Crabtree and Richard Mortier. 2015. Human Data Interaction: Historical Lessons from Social Studies and CSCW. In *Proceedings of the 2015 14th European Conference on Computer-Supported Cooperative Work, ECSCW'15*, 3–21.
  7. Andreas Dieberger, Paul Dourish, Kristina Höök, Paul Resnick, and Alan Wexelblat. 2000. Social navigation: techniques for building more usable systems. *interactions* 7, 6 (November 2000), 36–45.
  8. Jose van Dijck. 2014. Datafication , dataism and dataveillance : Big Data between scientific paradigm and ideology. *Surveillance & Society* 12, 2: 7487.
  9. Paul Dourish. 2003. The appropriation of interactive technologies: Some lessons from placeless documents *Journal of Computer Supported Cooperative Work* 12, 4: 465–490.
  10. Paul Dourish, 2016. Algorithms and their others: Algorithmic culture in context. *Big Data & Society*, 3, 2.
  11. Ron Eglash. 2004. Appropriating technology: An introduction. In *Appropriating technology: Vernacular science and social power*, Ron Eglash, Jennifer L. Croissant, Giovanna Di Chiro, and Rayvon Fouché. U of Minnesota Press, US, VII-XXI
  12. Motahhare Eslami, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. 2016. First I “like” it, then I hide it. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)* 2371-2382.
  13. Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. “I always assumed that I wasn’t really that close to [her].” *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems – (CHI '15)*, 153-162.
  14. Melanie Feinberg. 2017. A Design Perspective on Data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2952–2963.
  15. Tarleton Gillespie. 2014. The relevance of algorithms. In *Media technologies: Essays on communication, materiality, and society*, Tarleton Gillespie, Pablo J. Boczkowski, and Kristen A. Foot. MIT Press, Cambridge, US, 167-193
  16. Barney Glaser and Anselm Strauss. 1967. *The Discovery of Grounded Theory*. Aldine.
  17. Erving Goffman. 1959. *The Presentation of Self in Everyday Life*. Anchor Books
  18. Erving Goffman. 1986. *Frame analysis: An essay on the organization of experience*. Harvard University Press.
  19. Jeffrey T. Hancock. 2015. The Facebook Study: A Personal Account of Data Science, Ethics and Change. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, 1–1.
  20. Simon Kristoffer Johansen and Anne Marie Kanstrup. 2016. Expanding the Locus of Control: Design of a Mobile Quantified Self-Tracking Application for Whiplash Patients. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, 59:1--59:10.
  21. Elizabeth Kazianas, Mark S. Ackerman, Silvia Lindtner, and Joyce M. Lee. 2017. Caring through Data: Attending to the Social and Emotional Experiences of Health Datafication. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*, 2260–2272.
  22. Rob Kitchin. 2017. Thinking critically about and researching algorithms. *Information, Communication & Society*, 20, 1: 14-29.
  23. Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111, 24: 8788–8790.
  24. Min Kyung Lee. 2018. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society* 5, 1.
  25. Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with Machines. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 1603–1612
  26. Caitlin Lustig, Katie Pine, Bonnie Nardi, Lilly Irani, Min Kyung Lee, Dawn Nafus, and Christian Sandvig. 2016. Algorithmic Authority: The Ethics, Politics, and Economics of Algorithms that Interpret, Decide, and Manage. *Proceedings of the 2016 CHI Conference*

- Extended Abstracts on Human Factors in Computing Systems - CHI EA '16*: 1057–1062.
27. Philip Manning. 2000. Credibility, Agency, and the Interaction Order. *Symbolic Interaction* 23, 233.
  28. Igor Mayer, Harald Warmelink, and Qiqi Zhou. 2016. A frame-reflective discourse analysis of serious games. *British Journal of Educational Technology*, 47, 2: 342–357
  29. Viktor Mayer-Schonberger and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. John Murray Publishers, London.
  30. Donald McMillan, Arvid Engström, Airi Lampinen, and Barry Brown. 2016. Data and the City. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*: 2933–2944.
  31. Daniel Neyland and Norma Möllers. 2017. Algorithmic IF ... THEN rules and the conditions and consequences of power. *Information, Communication & Society*, 20, 1: 45-62.
  32. Wanda J. Orlikowski and Debra C. Gash. 1994. Technological frames: making sense of information technology in organizations. *ACM Transactions on Information Systems* 12, 2: 174-207. <https://doi.org/10.1145/196734.196745>
  33. Emilee Rader and Rebecca Gray. 2015. Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)* 173-182.
  34. Paul Resnick and Hal R. Varian. 1997. Recommender systems. *Communications of the ACM* 40, 3: 56-58.
  35. Minna Ruckenstein and Natasha Dow Schüll. 2017. The Datafication of Health. *Annual Review of Anthropology* 46, 1
  36. Antti Salovaara. 2008. Inventing New Uses for Tools: A Cognitive Foundation for Studies on Appropriation. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments* 4, 2: 209–228.
  37. J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. 2007. Collaborative Filtering Recommender Systems. In *The Adaptive Web*, Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl. Springer, Berlin, Germany, 291-324.
  38. Nick Seaver. 2013. Knowing algorithms. Media in Transition 8. Retrieved April 25, 2017 from <http://nickseaver.net/papers/seaverMIT8.pdf>
  39. Suvi Silfverberg, Lassi a Liikkanen, and Airi Lampinen. 2010. "I'll Press Play, But I Won't Listen": Profile Work in a Music-focused Social Network Service. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 1-10.
  40. David Silverman. 2006. *Interpreting qualitative data: Methods for analyzing talk, text and interaction*. Sage.
  41. Anselm Strauss and Juliet Corbin. 1990. *Basics of qualitative research. Grounded theory procedures and techniques*. Sage.
  42. Susann Wagenknecht, Min Lee, Caitlin Lustig, Jacki O'Neill, and Himanshu Zade. 2016. Algorithms at Work: Empirical Diversity, Analytic Vocabularies, Design Implications. *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*: 536–543.
  43. Michele Willson. 2013. The politics of social filtering. *Convergence: The International Journal of Research into New Media Technologies*, 20, 2: 218-232.
  44. Michele Willson. 2017. Algorithms (and the) everyday. *Information, Communication & Society*, 20, 1: 137-150.
  45. Youtube. 2012. Why we now focus on watch time. Retrieved April, 12, 2018 from <https://youtube-creators.googleblog.com/2012/08/youtube-now-why-we-focus-on-watch-time.html>