

**An agnotological analysis of APIs: Or, disconnectivity and the ideological
limits of our knowledge of social media**

John, N. A., & Nissenbaum, A. (2018). An agnotological analysis of APIs: or, disconnectivity and the ideological limits of our knowledge of social media. *The Information Society*, 1-12.

Nicholas A. John and Asaf Nissenbaum

Department of Communication and Journalism, The Hebrew University of Jerusalem,
Jerusalem, Israel

EMAIL: n.john@huji.ac.il

TWITTER: @nicholasajohn

*This is an Accepted Manuscript of an article published by Taylor & Francis in The
Information Society on 29 December, 2018, available online:*

<http://www.tandfonline.com/10.1080/01972243.2018.1542647>

Abstract

Adopting an agnotological perspective, this article extends the critical literature on APIs (application programming interfaces) by systematically showing that social media APIs are largely blind to acts of disconnectivity such as unfriending and unliking. We do this through analysis of the traces of social media usage that are not accessible through APIs as gleaned from the technical documentation published for developers by 12 major SNSs. Our findings make two main contributions. First, we show for the first time that APIs offer virtually no access to data about disconnectivity. Second, we show that APIs offer a very limited historical perspective, particularly regarding disconnectivity. However, for types of users that might spend money on advertising, far more historical and disconnectivity-oriented information is accessible through the API. This has practical consequences for research, and contributes to an agnotology of social media that sheds critical light on the advertiser-friendly atmosphere of connectivity that social media try to create.

An agnotological analysis of APIs: Or, disconnectivity and the ideological limits of our knowledge of social media

Agnotology: A missing term to describe the cultural production of ignorance (and its study)

Robert N. Proctor (2008), in *Agnotology: The making and unmaking of ignorance*

Introduction

This article contributes to the critical literature on application programming interfaces (APIs) by systematically showing for the first time that social media APIs leave us largely ignorant as to acts of disconnectivity such as unfriending, unliking, and so on, arguing that this has significant epistemological consequences. We do this through a close analysis of the digital traces of social media usage that are and are not accessible through platforms' APIs as gleaned from the documentation by them for the APIs of 12 major social network sites (SNSs). That is, our empirical materials are the extremely dry and technical manuals for using APIs, which detail the kinds and extent of data made available through them. These are published for developers by social media platforms. We read through them in search of methods for collecting social media data about disconnectivity. Our study makes two main empirical contributions. First, we show that APIs offer virtually no access to data about disconnectivity, unless such information might encourage you to spend more money on advertising on the platform. Thus, for instance, Facebook Page admins can see how many users hid their posts or reported them as spam, but individual users cannot. Furthermore, and relatedly, we show that these APIs offer a very limited historical perspective, particularly in regard to disconnectivity. Thus, Facebook's API lets you see when you liked something, but

not when you unliked it. Taken together, these findings have important impacts for the epistemology, or rather, *agnostology*, of social media and information technology research.

Consider this #1: When you make a friend on Facebook, or gain a follower on Twitter, you receive a notification, but when you are unfriended, or when someone unfollows you, you do not. Or this: When you like something on Facebook, it is recorded on your Activity Log. However, when you unlike something, this action does not appear on your Activity Log (the URL for which is [facebook.com/\[username\]/allactivity](https://facebook.com/[username]/allactivity)). Or this: Facebook has a page called “A World of Friends,” located at facebook.com/peace. On this page, Facebook presents daily updates of the number of friendships formed “just yesterday” between people from the opposing sides of three national conflicts: Israel/Palestine, Pakistan/India and Ukraine/Russia. There is no page called “A World of Enemies” that shows the number of friendships broken in the last 24 hours.

Facebook would rather we did not concern ourselves with negativity. The company told the creator of the app and browser extension, Who Deleted Me, that he was creating a negative user experience.¹ Facebook, and other SNSs, let us unfriend, unfollow, mute and block other users, but they do this in order to keep us on the site. As Light and Cassidy (2014) have shown, for SNSs disconnectivity is the lesser of two evils: better you unfriend someone than leave the service because of them.

Consider this #2: Unless you have been keeping your own log, you have no way of knowing how many Facebook friends you had a year ago. Only Facebook does. Or

this: Even if you can remember some of the people you unfollowed (on Twitter, Instagram, etc.), you have no way of knowing when you unfollowed them.

While for many people SNSs function as a kind of personal archive (e.g. Zhao et al. 2013), the platforms themselves are actually selective about what we can remember through them in a way that is strongly biased towards positivity. The culture of social media is one of connectivity (van Dijck 2013) and sharing (John 2016).

While this culture of positivity is clearly expressed through the user experience and interfaces of social media platforms, in this article we show that it also infiltrates a set of technological mechanisms used by researchers of all kinds to study social media, namely, social media APIs. By shaping the way data are collected from SNSs, any biases in APIs—and here we are interested in epistemological ones—will filter up into theory as well. Bodle (2011) defines APIs as “software tools that enable interoperability or the sharing between websites and online services” (321), while Bucher (2013) adds that, “Essentially APIs are interfaces that facilitate the controlled access to the functionality and data contained by a software service or program.” Originally developed to enable discrete blocks of software to “talk” to one another, they are nowadays an essential part of the infrastructure of social media. Third-party developers can use SNSs’ APIs to post status updates or tweets from an app on a user’s behalf, and they can also access information about the user (and sometimes their friends). Indeed, Rieder et al. (2015) remind us that “social media companies provide such interfaces to third party developers to foster the growth of an application ecosystem” (2), and that “APIs provide seemingly robust and convenient access to vast data pools for research purposes” only as a “side effect” (2). But even as a side effect, APIs are “an exceptionally useful point of access to social media data for researchers” (Bruns and Burgess 2016, 19).

The fact that APIs were developed in commercial contexts for commercial ends has not been lost on critical students of APIs, and explains much of what scholars find problematic about using APIs for research. Indeed, as Burgess and Bruns (2015), have noted, “increasingly the tools, techniques, and methodologies of corporate data mining are becoming less distinct from academic ones” (99). For instance, Lomborg and Bechmann (2014) observe that the “usefulness of APIs for researchers is very much dependent on the developers and commercial providers of the service” (260) with the result that if a company feels it can commercialize data, it may simply remove access to them via the API. For this reason, critical approaches to APIs focus on the power relations that they both embody and enact. Puschmann and Burgess (2013), for example, view APIs as “material instantiations of regulatory instruments used by the platform provider” (45), while Bucher (2013) argues that “web APIs constitute important sites for the study of power relations,” adding that “APIs need to be understood as historical projections of power within specific organizational and epistemic structures.”

In the remainder of this article we present the theoretical background to our interest in social media APIs and review the critical literature about APIs. We then describe our methods of data collection and analysis before laying out our findings. In presenting our findings we differentiate between information APIs make accessible about the logged in user and her activities on the one hand and about other users and their activities on the other. We also differentiate between individual and institutional users. We establish that there is clearly a bias towards connectivity and the present in social media APIs, with the crucial caveat that institutional or commercial users—who might spend money on advertising—have access to both negative feedback and historical data. The implications for social media research are

then discussed. In particular, we argue that because certain kinds of data (about unfriending, blocking, etc.) are systematically unavailable to API-based studies, researchers may find themselves unwittingly reinforcing the culture of connectivity (van Dijck 2013) promoted by the platforms themselves, thereby creating a skewed view of social life on the web.

Literature

The most relevant literature to this research is that concerning APIs and data collection from social media. However, given that we are driven by questions about what we do not and cannot know, we shall start with a brief exegesis concerning *agnostology*.

Agnotology is defined by Robert Proctor, the historian of science who coined the word, as “A missing term to describe the cultural production of ignorance (and its study)” (Proctor 2008, 1). Proctor urges us “to think about the conscious, unconscious, and structural production of ignorance, its diverse causes and conformations” (3), and suggests a tripartite classification of types of ignorance: “ignorance as *native state* (or resource), ignorance as *lost realm* (or selective choice), and ignorance as a deliberately engineered and *strategic ploy* (or active construct)” (3, italics in original). The first kind of ignorance (as “native state”) is that of the newborn baby, or of humanity from eras past, and is a lack of knowledge we seek to mitigate over time. For Proctor, this view of ignorance is closely associated with the project of modernity in two ways: first, it is the “gaping hole” that science seeks to fill; but second, and ironically, it is also a resource we need to keep drawing on, for without ignorance our scientific endeavors are rendered somewhat pointless. As Proctor puts it, “ignorance is needed to keep the wheels of science turning” (5). The second (ignorance as “lost realm”) entails questions of “political geography,” such as “Who knows not? And why not?” (6). Here Proctor discusses research left unfunded, often because other research is given

precedence. We take particular encouragement here from Proctor's work on agate, about which much less is known than other, more valuable rocks and minerals, such as diamonds or coal. Important for our concern here is the observation that there is not much money to be made from understanding agate better. That is, if knowledge has a political economy, then so does ignorance. Proctor's third and final kind of ignorance (the result of a "strategic ploy") is actively produced by entities that wish to foster ignorance or doubt, such as the security industry, the tobacco industry or climate change deniers. When ignorance is strategically produced (by the tobacco industry, by Facebook...), then it is the case that "They know, and may or may not want you to know they know, but you are not to be privy to the secret" (8). While it may be politically pertinent as to whether or not they want us to know that they know, the ignorance thus produced is the same ignorance.

Proctor himself readily admits to the arbitrary nature of such distinctions, and indeed a question that we seek to answer here is what kind of ignorance social media APIs produce. One thing we can say is that we are not dealing with false knowledge; this is not a situation of misinformation or lies, but rather the plain absence of knowledge. On the one hand, APIs were not developed as tools for research, and we might see their (hypothesized) blindness to disconnectivity as the result of selective choice. On the other hand, though, and as we shall see, it seems possible that at least part of the ignorance that APIs fail to dissipate is the product of strategic decisions that serve the companies' obligations to their shareholders and funders and their desire to attract advertisers. As Proctor notes, "Ignorance can be an actively engineered part of a deliberate plan" (9). Does this adequately explain the knowledge we cannot glean through the APIs under study here? Are SNS APIs neglecting disconnectivity based on their vision of social networking or purposeful commercial interests? Our analysis

will bring us closer to an answer. Before that, though, we offer a review of the critical literature on APIs and show how this article addresses as yet unasked questions.

The first thing we note is that “None of the publicly accessible APIs offers researchers the same degree of access that one of the company’s own engineers might enjoy” (Driscoll and Walker 2014, 1749). Indeed, Mary Gray (2017) argues that researchers are ultimately dependent on tech companies for data, and have to find a way to collaborate while serving the public interest and avoiding bias. Therefore, the backdrop to this entire discussion is our knowledge that social media platforms know everything about everything that is done on them while researchers only know that which social network sites deem knowable. The reason for this is, of course, that the SNS’s database is the source of its income. Therefore, as Skeggs and Yuill (2016) note, “An API will only ever give partial access to data that is carefully controlled by its providers to enable sufficient openness to encourage innovative spin-offs by outside developers, which can be harvested for future internal product lines, and, at the same time, ensure sufficient closure around that data so as to protect the provider’s competitive interests’ (2016, 1360). An analysis of APIs is one way of gauging what social media platforms have made knowable. As the major interface in the asymmetrical relationship between researchers and social network platforms, APIs have been critiqued in a number of ways. However, none of the critiques issued to date address the almost complete absence of the ability to receive data about disconnectivity through social media platforms’ APIs, which is the focus of this study.

A major group of critiques of API-based big data research has focused on *access* to social media data and the financial resources that collecting them demands. Burgess and Bruns (2015) have described the evolution of Twitter’s control over access to its data. When it first

launched, Twitter had a very liberal approach to APIs, which, it has been argued, is the main factor behind its success (Bucher 2013). Over time, though, it placed restrictions on what could be done with its APIs. Of particular relevance to researchers were the reduction of the number of permitted calls, a restriction to the last 3,200 tweets per user or search term, and the narrowing of the hose to 1% of tweets being tweeted at that time. In parallel, Twitter established partnerships with data brokers who received licenses to resell Twitter data (Burgess and Bruns 2015). A consequence of this, Burgess and Bruns argue, has been the expansion of “hashtag studies” and the need for most researchers to make do with “low hanging fruit,” as the costs of large-scale Twitter studies have soared beyond the reach of the vast majority of researchers (see also Bruns 2013). This is a clear indication of the contours of the political economy of social media research, and of the impact research tools have on the topics studied.

However, even if Twitter reverted to its former model of allowing APIs access to the full Twitter stream, and even if there were no financial costs to collecting data from any SNS, researchers would still be faced with the problem of the *opacity* of APIs and the ability of companies to change them at will, which might undermine their validity and usefulness as research tools. As Ruths and Pfeffer note, “researchers are left in the dark about when and how social media providers change the sampling/filtering of their data streams” (Ruths and Pfeffer 2014, 1063). Rieder (2016) describes this in his documentation of changes to Facebook’s API, which ultimately broke a research tool painstakingly developed by Rieder and used by scholars the world over. Likewise, Driscoll and Walker (2014) describe Facebook and Twitter as “unstable” platforms that are “privately owned and operationally opaque” (1746).

While the critiques mentioned so far concern the data made available through APIs, Baym (2013) makes an intriguing point about the pertinence of data that we do not have. She points to the use of samples of record store staff as the basis for measuring audience size in the music industry, a method of data production that entirely missed the rising popularity of hip-hop. Ultimately, argues Baym, “Claims based on analyses of social media data must be closely scrutinized *with an eye toward what they omit*” (emphasis added). boyd and Crawford (2012) also discuss data we do not have when they observe that “Twitter and Facebook are examples of Big Data sources that offer very poor archiving and search functions” (666). As a result, they argue, “researchers are much more likely to focus on something in the present or immediate past [...] because of the sheer difficulty or impossibility of accessing older data” (666). However, there are important differences between Baym’s and boyd and Crawford’s arguments and ours. To start, Baym is referring to data that were not collected because the data collectors were bad at collecting data, but make no mistake that SNSs are logging (and almost certainly studying) data about disconnectivity that they are not making available through their APIs. Baym’s example says: you should have been collecting *this* data, and not just *that* data. We say: they *are* collecting this data and that data, but not making all of it available for scrutiny outside the company. boyd and Crawford, meanwhile, point to a bias to the here and now in data-based social media research with which we have no quibble. However, we say that even if APIs offered a much more historical perspective (beyond a user’s last 3,200 tweets, for instance), they would still remain radically blind to certain (negative) actions taken only a millisecond ago.

We have presented four main types of critique of the use of APIs for research to date: data are expensive; APIs are opaque; APIs are restrictive; and not all of the data even exist. These critiques are crucial interventions into our understanding of social media API-based research,

but they lack an important aspect, namely, attention to the unavailability of data about disconnectivity through them. These are not data that companies have not thought of collecting—and as we shall show, some of them *are* available to some users in some circumstances—but rather digital traces that are not available through APIs. This is not to say that research into disconnectivity is impossible (there are other research methods than collecting data through APIs); nor is to argue that all knowledge resides within the databases of social media companies. It is, however, to see social media APIs as cultural artefacts that reflect and reproduce the epistemological priorities of some of the most powerful companies in the world today. As a result, APIs shape the way we observe, think about, and study the social networking world that these companies provide us.

Methodology

Similarly to Bucher and Helmond (2017), we see APIs as an interface between platforms or services and users, and hence as a site for observing the different kinds of relationships afforded by the APIs between SNSs and users. To this end, we subjected the documentation² of the APIs of 12 SNSs to qualitative analysis in order to understand what information they let users glean (or not). The SNSs were selected by triangulating sources: Alexa (for number of visitors), Wikipedia (for numbers of visitors and users), and Web of Science (for number of mentions in academic research). This process established which SNSs are the largest and most significant platforms, including for researchers. The 12 sites with the highest average ranking according to these three sources were included in the study. In descending order, they were: Facebook, Google+, Twitter, LinkedIn, Sina Weibo, LiveJournal, Habbo, Foursquare, Flickr, Pinterest, Instagram, Tumblr. However, of these Google+, Pinterest, LiveJournal and Habbo had only a cursory API or no functional API at all.

The first stage of the research was to systematically read through all of the technical documentation³ for the APIs of the platforms included in the study while asking: What does this API let us know about disconnective activities carried out on the site? Because we were not expecting to find many (if any) ways of gathering data about disconnectivity through social media APIs, we paid special attention to the negative spaces created by the positive functionality of the APIs we analyzed. That is, if an API enables one to see a kind of positive feedback (such as liking or friending), we asked whether it enabled knowledge about an equivalent negative action, such as retracting a like or unfriending. Also, if an API provides options for looking back in time, we looked to see whether historical information about the equivalent negative action could also be gathered. More generally, where APIs allow the study of tie formation, we looked for the ability to collect data about tie dissolution. In doing so, our analysis followed each of the eight substantive APIs to examine closely three categories of user vis-à-vis the observer: strangers (users with whom the observer has no tie), ties (friends, followers, etc.) and self (i.e., one's own information). For each of these, relevant API endpoints were listed to outline possibilities of attaining information, which were divided into data on connectivity or disconnectivity (where available). This first reading focused on the default options APIs account for, disregarding heightened privacy settings that might be available to users (such as making friend lists or account details private).

In the second stage of analysis, having established that commercial users (who might pay for advertising) are afforded wider access to metrics about disconnections and negative feedback (see below), we reread the documentation, this time carefully examining the options made available to different categories of user.

Findings

The findings show an overwhelming bias towards connectivity in social media APIs. While APIs allow users to extract lists of current friendships, followings, group memberships, and so on, they include hardly any information about blockings and disconnections, along with an inability to track changes in connections over time. However, this bias is far less pronounced when it comes to commercial users (or at least, users the SNS considers more likely to spend money on advertising). For instance, the owners of Pages and Apps on Facebook *can* get information about unlikes and various other kinds of negative feedback. Furthermore, businesses are afforded a diachronic view that individual users (and researchers) are denied. Thus, businesses on LinkedIn and venues on Foursquare, for instance, are able to track engagement over time and discern downturns.

In the rest of this section we present our findings in three groups. We start with what APIs do not enable users to know about other people and other people's content. We then look at what APIs cannot tell users about themselves and their own data before finally examining what APIs let commercial or institutional users (owners of Pages, Apps, locations, etc.) know about themselves and their customers. The findings are grouped in this way for the sake of convenience. To be sure, there is a logic—we move from the most restrictive to the most liberal in terms of data accessibility—but these are not categories that one finds in the API documentation itself.

Other people and other people's content

We start by looking at the information accessible through APIs about other people and other people's content. At the most basic level of connectivity, most platforms' APIs provide data about how many contacts (Facebook friends, LinkedIn connections, etc.) users have and who

they are. Moreover, on Twitter, it is possible to use the API to see the relationship between any two users (suggesting that, in theory at least, the API could be used to reconstruct the entire social graph of Twitter; Kwak, Chun, and Moon 2011 did something like this for the Korean Twittersphere). The Twitter API also allows one to see a given user's lists and the lists that they are listed in (unless any of those lists are private, in which case one would need to have permission to see them). The Flickr API allows the same for group memberships. The Facebook API gives access to a list of pages liked by a given user, including when the user liked each of them. On the Foursquare API, looking into users reveals their lists of visited venues, check-ins, and "mayorships" of venues.

APIs generally provide information about positive actions made in relation to content. Thus, for instance, the Instagram API allows one to see who has liked a photo, and which other photos share a tag with it, while the Twitter API will show who tweeted a given tweet, and also who has retweeted a given tweet. The Tumblr API will similarly show how many likes and reblogs a given post has had. The Facebook API also enables one to learn how many likes an object (e.g. a photo) has.

APIs do provide data that might be considered related to disconnectivity, but only on specific and limited issues. For example, the Twitter API will show whether a user's account is protected, and we have already mentioned that private lists are not viewable through the API (or indeed through any interface). It will also show whether a certain user or tweet is withheld from certain countries, or, for content, whether it has been withheld following a copyright claim.

But what knowledge about other people or their content is inaccessible through platforms' APIs? At the most basic level, APIs do not tell us whom other users used to be friends with in the past (or whom they followed, or who followed them) or what they used to like (but no longer like). Even numerical accounts of such changes in a user's connections are not available. Also, one cannot see individuals other users have blocked or muted. Similarly, with a focus on content, platforms' APIs shed no light on negative feedback, such as users unliking photos (that is, revoking a previous like). On Foursquare, the only platform on which an item, namely venues, can be disliked, this action is hidden from everyone but the user who performed it, while everyone's liked venues remain visible through the API. It might be reasonably argued at this point that APIs should not give access to such information about other people for reasons of privacy. In terms of agnotology, this would be a strategic ploy but one that seems widely acceptable. Indeed, one can think of cases where it would be grossly unethical to expose a user's former ties (see, for instance, Haimson et al. 2016, on the management of Facebook ties by transgender people). More generally, a right to be forgotten should not be trampled over by data-hungry researchers, and our critique here should not be read as calling for unlimited access to all data about other people. However, as we shall now see, information about people or content that we ourselves no longer like is equally inaccessible through social media APIs.

Me and my content

As just mentioned, basic considerations of privacy can help us understand why APIs might not (and should not) provide lists of users other people have blocked. However, privacy has nothing to do with why APIs do not provide lists of one's own former friends to users. Similarly, privacy considerations cannot account for the lack of historical data about content I used to like (but no longer like) available through APIs. And indeed, not a single API makes

it possible to see former ties on social network sites: not former Facebook friends, or former Twitter or Instagram followers/ees, and so on. Twitter offers users analytics about their own account,⁴ including a graph showing number of followers over time. However, it does not show who these followers are. Moreover, the data behind this graph are not accessible through the Twitter API. Users may find ways to compensate and follow their activity closely enough to notice changes to the social environment (Bevan, Pfyl, and Barclay 2012), but this remains a subversion of the platforms' design, as demonstrated by the APIs, and not a reliable source of information. Put bluntly, social media APIs render us blind to the historical composition of our own networks, to say nothing of the lack of information available about people with whom we used to share a tie.

As to the information that is available, anything that APIs let one know regarding other people and other people's content can also be known regarding oneself and content one has posted. Thus, for instance, if an API lets me see how many followers another user has, it will also let me see how many followers I have. However, one can also gather information about oneself that one cannot gather about other people. For instance, one can use the Twitter, Tumblr or Facebook API to receive a list of users one has blocked. The Flickr API can list specific forms of connection within one's list of contacts, such as those marked on the site as family members or friends. There are also various ways of obtaining information about content we have created or interacted with on SNSs. For example, the Instagram and Flickr APIs will show how many comments one's post or picture got; the Tumblr API will show how many likes a blog post that one wrote received; the Twitter API can show how many times a tweet was retweeted, and by whom. Also, through the Tumblr API, and unusually for the APIs we studied, it is possible for a user to see how many and which posts the user liked before or after a certain date. Thus, the Tumblr API enables a diachronic perspective that is

absent from the other APIs—at least when it comes to their use by individual (as opposed to commercial) users. Another unusual feature is found in the Foursquare API, which lets a user get a recommendation for nearby venues she is likely to visit (and subsequently check in to or like). The Facebook API also lets users know when they liked certain content—the API provides the timestamp for likes (and indeed, the Activity Log shows this too), but does not allow for the definition of a timeframe as part of the call to the API.

However, it is not possible through APIs to glean data about what one used to like but no longer likes, such as unfavorited tweets, or unliked Facebook pages or Foursquare/Swarm locations, to say nothing of content one has deleted. As mentioned above, one can use the Foursquare API to see locations one has disliked, but this is because Foursquare has a “dislike” button; this is not the same as seeing locations from which one has removed a previously awarded a “like.”

There might be a case to be made for preventing users from easily seeing who unfriended them, but it is hard to think of one for preventing users from researching whom they themselves unfriended (or disconnected from in other platform-specific ways) in the past.⁵ Likewise, while we may be grateful for the ignorance embedded in APIs when it comes to other people’s access to information about things we used to like, our inability to see back into our own past through APIs appears to be a reflection of their bias against disconnectivity and negativity and of a focus on positive aspects of the present. This is hard to categorize when viewed as a production of ignorance. It may certainly emanate from the strategic needs of SNSs, as a way to curate and control user experience, keeping it positive by making negative dynamics forgettable or unknowable. On the other hand, SNSs could simply be creatures of the here and now, for which past states simply do not justify attention. In this

state, ignorance about disconnectivity creates a lost realm, a topic not worthy of devoting attention resources from users or server resources from platforms.

By way of a brief summary, when an individual user uses a social platform's API to obtain information, they will not be able to get negative information, and certainly not negative information about the past. So far, we have discussed individual users of social media platforms and the information they can obtain about others and themselves through APIs. We turn now to discuss business users—or users who might potentially spend money on the platform—and the information APIs offer to them.

Commercial users and their crowds

Many of the limitations on individual users' uses of APIs are lifted when it comes to commercial or public entities and their representations on SNSs, e.g. Facebook Pages or Apps, Foursquare Venues, or firms on LinkedIn.⁶ Although not all SNSs have the option of creating such a business-oriented entity, those that do allow for a far wider breadth of data to be accessed through the API, including information on both connectivity and disconnectivity over time. The lack of these data in other parts of the API is therefore either an ideological choice or a financial one (or both). In either case, access to information about the negative and historical elements of SNS use is generally not dependent on negating privacy concerns, but on having a commercial relationship with these sites.

The best examples for this come from Facebook and its "Insights" interface, which is also available through the API. It allows the owner of a Page or an App on the site to track its popularity and success among its Facebook audience, including its ups and downs over time, as well as specific negative responses. The admin of a Facebook Page can receive a multitude

of breakdowns and reports on interactions with it, including: likes, posts on the Page, any click on any of the Page's content, visits to the Page, and the reach of its content (i.e., how many users saw it on the News Feed). These and other metrics supply both a count of the interactions and the number of users performing it, as well as various demographic data. Additionally, most of these numbers are available in different resolutions of day, week, or month. Apps receive information in the same scope and various resolutions, albeit with slight differences, such as targeted details about usage of specific parts of the application (e.g. dialogs or notifications). While this is only a partial account of the data made available (Facebook's full explanation is over 7,000 words long), it makes clear that a Page or an App can know in great detail if and how its social network status has changed for better or worse, and which kinds of people have been responsible (from aggregate demographic data given by Facebook). These data are not endless, but Facebook guarantees access to two years' worth of records, after which it is subject to removal.

Moreover, a few specific points within the "Insights" interface are dedicated to negative feedback. Under engagement reports for a Page or an App, administrators receive a count of negative actions and the number of unique users who performed them, as well as a count of different types of action—hiding some or all of their content, unliking, reporting content as spam and so on. These are available retroactively for different timeframes for a Page or App, or over the entire lifetime of a specific post.

Altogether, it is clear that commercial users of Facebook can learn a lot about their standing within the network and any negative or disconnective patterns associated with their SNS activity. This pattern is repeated, on a smaller scale, in other business-oriented SNSs, namely, Foursquare and LinkedIn. Foursquare, a location-based network, allows businesses in which

users “check-in” to retroactively view the number of check-ins, unique or new visitors, and sharing to Facebook or Twitter. They can also receive the names of the ten most frequent and most recent visitors within a defined timeframe. In LinkedIn, the operators of company pages can receive data on the number of followers at different points in time, as well as impressions and engagements with posts. Business users of both networks can therefore track connections and disconnections to their pages by simple retroactive comparison. Another option for all companies allows them to break down their followers into different sectors based on various details. Furthermore, LinkedIn’s API used to state that companies recruiting employees through LinkedIn gain access to their applicants’ list of followed companies, industries and jobs, as well as the suggestions on who to follow given to that user by LinkedIn.⁷ These options yet again reveal to commercial users a far wider scope of the connections made within the network, both for their own company’s profile and for those they come in contact with.

When put together, the different segments of APIs directed at commercial users paint a clear picture. Many of the details, seminal to the experience of SNS users yet missing from APIs and interfaces, are easily available to those who are potentially willing to spend money on advertisement or promotion. For these users, the logic that directed the structure of the rest of the API is suspended, in favor of affording a torrent of data aimed at improving their standing in the network or reaching wider audiences.

A final indicative example comes from comparing Facebook’s Pages, meant for use by businesses and organizations, to the site’s Groups platform, meant for communities or collectives. While both are designated as platforms for crowds of users to interact, Group administrators receive none of the privileged information given to Page owners. Unlike

commercial users, they are not deemed worthy enough to know how their Facebook presence has fared in the past few months. This contrast makes the differentiation abundantly clear: connectivity and disconnectivity data, whether synchronic or diachronic, are available to the SNSs themselves but given out in very limited ways to users, except those who may spend money on these platforms. It does not matter if the data are requested for one's own account or for managing a popular communal page, only business ties are exempt from the SNSs' manufactured ignorance concerning disconnectivity or historical patterns of connectivity.

Discussion

Our analysis of the APIs of 12 major SNSs produced two main and interrelated findings, and a crucial caveat, which take us beyond existing critical appraisals of APIs. First, social media APIs do not enable individual users to obtain knowledge about negative actions on social media platforms. Users cannot know whom they have unfriended on Facebook or unfollowed on Twitter, to say nothing of knowing who has unfriended them.⁸ Similarly, it is impossible to know what negative responses one's social media activities have inspired among other people, such as causing them to hide a post or mute a conversation. While this lack of information may mean different things on different platforms and their specific cultures, the pattern of excluding disconnectivity data from APIs is indicative of an overarching logic. This is significant for API-based academic research in that one way of studying social media use is to ask participants to install an app that accesses the API from the perspective of the participant, feeding data back to the researcher (for instance Skeggs and Yuill 2016). By extension, of course, APIs will not provide such information about third parties. Second, APIs are quite blind to history (as noted by boyd and Crawford 2012), particularly disconnective history. As we saw, there are APIs that allow time frames to be defined, showing us when we responded positively to a piece of content or another user. However,

APIs give us no information about when we stopped liking that content or user. You can know what you currently like or to whom you are connected, but you cannot know what you used to like or whom you used to be connected to. Yet, when the user is a potential paying customer of the SNS, its API may very well make data available about negative feedback, and it might offer a (relatively) deeper historical perspective.

Our findings establish quite clearly that the functionality of the APIs of the social media platforms analyzed here both reflects and reproduces the companies' commercially-driven understanding of the user experience of different categories of user—the individual and the commercial or institutional. For the individual, social network sites envisage a positive experience (hence the absence of a “dislike” button, which, *inter alia*, would have the potential to open up new forms of bullying). Notifying users when they have been unfriended, or allowing developers to develop onsite apps that notify users, would not be conducive to the atmosphere of positivity and connection that Facebook (and others) wish to produce. Communal users that are not commercially oriented are met with the same approach. Only for the institutional user is the user experience based around analytics, insights, and metrics for measuring engagement. The objective, from the SNS's point of view, is encouraging the spending of money on advertising. What this also makes clear is that the lack of certain analytic abilities in APIs that face individual users is not due to technical limitations, but rather is a choice made by the SNS.

These original findings are significant because if what we can know is a function of how we can know, this study has implications for our ability to research what we do with SNSs, and, to the extent that they are a proxy for broader social processes, for those processes as well. In this we echo the calls made by Puschmann and Burgess (2013) and Rieder (2016) for SNSs

to make more data accessible through their APIs. The danger we see with API-based research on its current trajectory is that the theory researchers develop might be shaped by the kinds of empirical materials the tools at their disposal are able to give, or that scholars may be “tailoring [their] research questions to data availability” (Gaffney and Puschmann 2013, 57). The available data, we argue, are directed towards positivity and connectivity. While we acknowledge that research programs can always be analyzed from a political economy perspective—especially when it comes to matters of funding—we wish here to issue a reminder that the agenda of Facebook and marketers is not necessarily that of researchers.

In other words, researchers should avoid being partners with marketers and the SNSs industry in promoting—even unwittingly—a culture of connectivity that primarily serves the interests of commercial organizations. Moreover, such partnerships are liable to further a perception of APIs as neutral or natural products of social media platforms, rather than the manifestation of business strategies and interests that they actually are. Of course, social scientists have always been interested in connectivity, and we are certainly not suggesting that they should cease to be. Equally, though, any theory of social life that ignored phenomena such as breaking up, quitting a workplace, changes in taste, or dropping out would at best be partial, and any methodology that *a priori* rendered such phenomena invisible would surely be discounted. However, unfriending, unfollowing, unliking, leaving a group, muting a conversation, and more, are all meaningful social actions to which SNSs’ APIs are blind. We are not suggesting that APIs exclusively hold the key to knowledge about life online, but it is important to acknowledge that the advantages they can offer over other methods of data collection only accrue to a narrow subset of social and online phenomena. Moreover, if realistically researchers must collaborate with tech companies, as Gray (2017) suggests, this state of affairs might be a good indication as to the pitfalls of such a union—when a wealth of

data is made available, there might reside wide gaps that go unnoticed. In other words, APIs demonstrate how tech companies are unwilling to share data that does not suit their general strategy, while sharing enough of everything else to obscure the creation of ignorance.

In the hypothetical situation of such data being made available by social media platforms, SNS studies would be enriched in several ways. The ability to trace and measure social polarization, for instance, would be greatly advanced by the availability of data on the deletion of connections and friendships on SNSs. Were broader data sets to be made available through APIs, they may uncover which kinds of events lead to different rates of disconnectivity and among whom (did Clinton supporters unfriend Trump supporters more than the latter unfriended the former?). Another process this data may shed light on is the dissipation of trends or the collective rejection of specific social entities, as reflected by wide disconnection from topics, groups, or institutions in their SNS incarnations (e.g. disconnections from Facebook groups supporting a certain ideology or cause). Furthermore, having such data available may contribute to the experience of regular SNS users: availing them of a view of negative interactions or feedback, currently available only to commercial users, would give users some insight into the implications of their actions online and may allow for more thoughtful behavior. That said, it is also crucial to consider some of the potentially negative implications of this data being available. Beyond the obvious issues of privacy, it may add to the ongoing and extremely troublesome use of social media in manipulating public opinion to benefit certain commercial and political interests. Even if user privacy is diligently preserved, the availability of more data may be exploited by those who seek to influence and divert public discourse.

As they stand, though, the biases of SNS APIs have practical implications for social science researchers of all hues. Specifically, because negative data are unattainable through APIs, researchers must seek alternative means of attaining them, such as surveys. Indeed, currently the only way we know of to measure unfriending, for instance, is through surveys, such as those conducted by Sibona and Walczak (Sibona and Walczak 2011), John and Dvir-Gvirsman (2015), and Zhu, Skoric, and Shen (2016). However, these have a number of disadvantages: researchers have to somehow ensure that their sample is representative of the Facebook population they are studying (which is difficult because Facebook does not publish precise enough demographic information); surveys are expensive; and they may be perceived as less reliable because they appeal to respondents' memories. Moreover, researchers run the risk of their survey-based data becoming utterly redundant should the platform eventually decide to publish the data. Given all this, researchers might decide that these directions are less worthwhile to pursue. In other words, it is not only that the API does not allow for certain lines of research to be pursued, but the limitations of the API might actually discourage researchers from pursuing those lines of inquiry in other ways, either because these other ways might be seen by their peers as less than adequate, and because there remains a chance that they will be gazumped by the platform itself.

Returning to the agnotology of APIs, we are now in a position to characterize the kind of ignorance produced by the biases described above. We take it as given that the SNS itself knows everything that takes place, while the API allows access to a predetermined subset of the data produced by users' actions on the site.⁹ It is clear that the API's functions and the scope of the data it has access to are defined by employees of SNSs. In this sense, the limitations placed on our knowledge—as researchers, but also as users of these services—fall into Proctor's third type of ignorance: it would seem to be the result of a strategic ploy on the

part of SNSs, and the subsequent ignorance is an intended consequence aimed at improving (or at least, not harming) the companies' profits. SNSs allow us to disconnect so that we remain within the network (Light and Cassidy 2014); however, they do not allow us to know anything about disconnectivity for precisely the same reason. At the same time, though, an understanding of the origins of APIs as enabling blocks of software to "talk" to each other, or as providing access to a database, would direct us away from too conspiratorial a view and suggest ignorance of the second type discussed by Proctor, namely, ignorance as a "lost realm." This view allows us to see the production of ignorance presented above as systematic, and as rooted in a cultural milieu, though without accusing SNSs of the kind of skullduggery that Proctor exposes the tobacco industry as engaged in.

The question as to which of these forms of the production of ignorance is predominant is in itself unknowable, as it requires insight into the inner workings of the SNSs that publish them. However, it is also a moot question, since APIs are the obligatory gateway for any application relying on an integration of SNS data or experience. Since they must use the APIs, the same patterns of ignorance are carried into the wider ecosystem of external interfaces available in applications, quashing any possibility of challenging the dominant form of connective culture.

Obscuring negative social dynamics may benefit the advertiser-friendly atmosphere SNSs try to create, but it comes at the price of misrepresenting social realities. Moreover, the biases exhibited by APIs for the benefit commercial users further this misrepresentation. Regular users are left blind to much of the negative interactions that affect their participation and representation within SNSs, and unable to manage this cardinal part of current the social sphere, while commercial users can do so easily. The result is both distortion and prejudice in

our view of the social scape—imagine if stock markets only reported on prices rising, except for certain, well-connected shareholders. The deeper implications of various conflicts may therefore be hidden behind this façade of connectivity and positivity, and this should be a concern for society at large but even more so to the academic endeavors utilizing SNSs’ APIs for social science.

Notes

¹ See <http://www.businessinsider.com/how-to-know-if-someone-unfriends-you-on-facebook-2015-7> (accessed 21 August 2016).

² As of end 2015.

³ If the reader wishes to get a taste of this terribly dry material, documentation for part of Facebook’s API can be found at <https://developers.facebook.com/docs/graph-api/using-graph-api>.

⁴ Accessible at [https://analytics.twitter.com/user/\[username\]/home](https://analytics.twitter.com/user/[username]/home).

⁵ We note that the information Facebook provides when a user requests their Facebook data does include a list of users whom they have unfriended.

⁶ We know that anyone can open a Page or publish an App on Facebook, and that not all admins of Pages are using their Page to promote sales. However, anyone wishing to use Facebook for marketing purposes will create a Page.

⁷ We were unable to double-check this because this part of the API was concealed at some point between our first and second readings, with documentation now available only to “partners” of the company.

⁸ We are aware of apps that enable Twitter users to see who unfollowed them. However, these apps work by comparing today's list of followers with yesterday's. They are not extracting data about unfollowers directly from the Twitter API.

⁹ We also take it as given that the SNS should not allow everyone access to everything that takes place, for to do so would be to violate its users' privacy.

References

- Baym, Nancy K. 2013. Data not seen: The uses and shortcomings of social media metrics. *First Monday* 18 (10).
- Bevan, Jennifer L., Jeanette Pfyl, and Brett Barclay. 2012. "Negative emotional and cognitive responses to being unfriended on Facebook: An exploratory study." *Computers in Human Behavior* 28 (4):1458-1464. doi: <http://dx.doi.org/10.1016/j.chb.2012.03.008>.
- Bodle, Robert. 2011. "Regimes of sharing: Open APIs, interoperability, and Facebook." *Information, Communication & Society* 14 (3):320-337.
- boyd, danah, and Kate Crawford. 2012. "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon." *Information, communication & society* 15 (5):662-679.
- Bruns, Axel. 2013. Faster than the speed of print: Reconciling 'big data' social media analysis and academic scholarship. *First Monday* 18 (10).
- Bruns, Axel, and Jean Burgess. 2016. "Methodological Innovation in Precarious Spaces: The Case of Twitter." In *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*, edited by Helene Snee, Christine Hine, Yvette Morey, Steven Roberts and Hayley Watson, 17-33. London: Palgrave Macmillan.
- Bucher, Taina. 2013. Objects of intense feeling: The case of the Twitter APIs. *Computational Culture* 3.
- Bucher, Taina, and Anne Helmond. 2017. "The affordances of social media platforms." In *The SAGE Handbook of Social Media*, edited by Jean Burgess, Alice E Marwick and Thomas Poell, n/a. London: SAGE.

- Burgess, Jean, and Axel Bruns. 2015. "Easy data, hard data : the politics and pragmatics of Twitter research after the computational turn." In *Compromised Data : From Social Media to Big Data*, edited by Ganaele Langlois, Joanna Redden and Greg Elmer, 93-111. London: Bloomsbury Publishing.
- Driscoll, Kevin, and Shawn Walker. 2014. "Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data." *International Journal of Communication* 8:20.
- Gaffney, Devin, and Cornelius Puschmann. 2013. "Data collection on Twitter." In *Twitter and society*, edited by Katrin Weller, Axel Bruns, Jean Burgess, Mary Mahrt and Cornelius Puschmann, 55-67. New York: Peter Lang.
- Gray, Mary L. 2017. "We Need More, Not Fewer, Collaborations With Tech Companies." *The Chronicle of Higher Education*. <http://www.chronicle.com/article/We-Need-More-Not-Fewer/240726>.
- Haimson, Oliver L., Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2016. "Digital Footprints and Changing Networks During Online Identity Transitions." *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, San Jose, California, USA.
- John, Nicholas A. 2016. *The Age of Sharing*. Cambridge, UK: Polity.
- John, Nicholas A., and Shira Dvir-Gvirsman. 2015. "'I don't Like you any more': Facebook unfriending by Israelis during the Israel-Gaza conflict of 2014." *Journal of Communication* 65 (6):953–974. doi: 10.1111/jcom.12188.

- Kwak, Haewoon, Hyunwoo Chun, and Sue Moon. 2011. "Fragile online relationship: a first look at unfollow dynamics in twitter." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Light, Ben, and Elija Cassidy. 2014. "Strategies for the suspension and prevention of connection: Rendering disconnection as socioeconomic lubricant with Facebook." *New Media & Society* 16 (7):1169-1184. doi: 10.1177/1461444814544002.
- Lomborg, Stine, and Anja Bechmann. 2014. "Using APIs for Data Collection on Social Media." *The Information Society* 30 (4):256-265. doi: 10.1080/01972243.2014.915276.
- Proctor, Robert. 2008. "Agnotology." In *Agnotology : the making and unmaking of ignorance*, edited by Robert Proctor and Londa L. Schiebinger, 1-36. Stanford, Calif.: Stanford University Press.
- Puschmann, Cornelius, and Jean Burgess. 2013. "The politics of Twitter data." In *Twitter and society*, edited by Katrin Weller, Axel Bruns, Jean Burgess, Mary Mahrt and Cornelius Puschmann, 43-54. New York: Peter Lang.
- Rieder, Bernhard. 2016. "Closing APIs and the public scrutiny of very large online platforms." *The Politics of Systems*, Last Modified 27 May 2016. <http://thepoliticsofsystems.net/2016/05/closing-apis-and-the-public-scrutiny-of-very-large-online-platforms/>.
- Rieder, Bernhard, Rasha Abdulla, Thomas Poell, Robbert Woltering, and Liesbeth Zack. 2015. "Data critique and analytical opportunities for very large Facebook Pages: Lessons learned from exploring "We are all Khaled Said"." *Big Data & Society* 2 (2):2053951715614980.

Ruths, Derek, and Jürgen Pfeffer. 2014. "Social media for large studies of behavior." *Science* 346 (6213):1063-1064.

Sibona, Christopher, and Steven Walczak. 2011. "Unfriending on Facebook: Friend request and online/offline behavior analysis." System Sciences (HICSS), 2011 44th Hawaii International Conference on System Sciences.

Skeggs, Beverley, and Simon Yuill. 2016. "The methodology of a multi-model project examining how facebook infrastructures social relations." *Information, Communication & Society* 19 (10):1356-1372. doi: 10.1080/1369118X.2015.1091026.

van Dijck, José. 2013. *The culture of connectivity : a critical history of social media*. Oxford ; New York: Oxford University Press.

Zhao, Xuan, Niloufar Salehi, Sasha Naranjit, Sara Alwaalan, Stephen Volda, and Dan Cosley. 2013. "The many faces of facebook: experiencing social media as performance, exhibition, and personal archive." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.