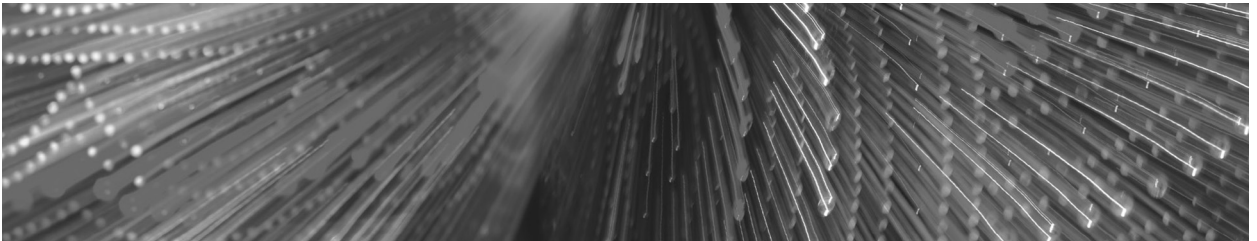


PART II

Approaches and Methods



Digital Methods for Cross-platform Analysis

Richard Rogers

DIGITAL METHODS AFTER SOCIAL MEDIA

Increasingly employed as an umbrella term for tool-based methods employed in the digital humanities and e-social sciences, digital methods have as their point of departure a series of heuristics with respect to how to study online media (Rogers, 2013b). The first historicises the web as an object of study, one that has undergone a transformation from a (virtual) site for the study of online culture specifically to a source of data about broader societal and cultural trends. Second, to extract the data one not only employs crawlers, scrapers, API logins and manual means, but also pays special attention to ‘query design’ and ‘search as research’ for creating tweet collections or sets of Facebook pages for social media analysis. To study those ‘natively digital’ source sets, digital methods learn from the methods of the medium, e.g., recommendation systems

such as trending topics or newsfeeds. How may platform treatments of retweets and likes (for example) be repurposed for studying the unfolding of historical events (on Twitter), or the most engaged with memes in a political campaign (on Facebook)? Digital methods, finally, consider the conditions of proof. When does it make sense to ground the findings (about the versions of historical events represented by Google search results, for example) in the particular characteristics and influences of online data, and when is ‘online groundedness’ less robust than mixed methods approaches?

One of the earliest digital methods maps the hyperlinking patterns between websites involved in the same social issue area so as to study the politics of association of actors from the purposively made as well as the missing links. The IssueCrawler, the software tool developed in the early 2000s or the so-called web 1.0 era, provides a ‘programmed method’ for studying associations in issue networks online, or clutches

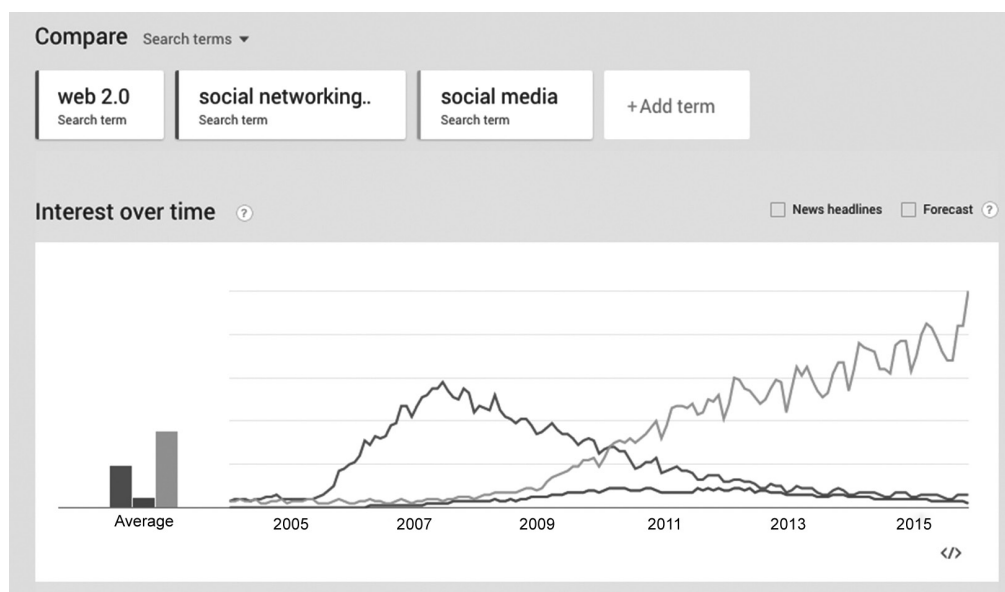


Figure 5.1 Comparison of search volume for [web 2.0], [social networking sites] and [social media], according to Google Trends, 19 November 2015

of NGOs, funders, think tanks, academics as well as databases, widgets and other online objects, working on or serving a particular issue (Bruns, 2007; Rogers, 2009a; Borra and Rieder, 2014). Once the links between actors have been found, one may begin to study association as well as the organisation of networked publics (Latour, 2005; Ito, 2008).

More recently, by calling for a move from 'so-called web 1.0 http or html approaches to 2.0 cross platform based methods,' Greg Elmer and Ganaele Langlois (2013: 45) argue that to study the web these days requires new methods that step past the hyperlink as the pre-eminent digital object tying it all together. They issue a much larger invitation to rethink the web more generally as an object of study, recognising its increasing platformisation, or the mass movement by web users to social media (Helmond, 2015). In the shift from an info-web (1.0) to a social web (2.0), recommendations are made through the participation of platform users rather than only by site webmasters (to use a throwback term).¹ That is, recommendations,

especially in the news feeds of platforms, follow from 'friends' activity, such as 'liking' and 'sharing'. The content recommendations thereby distinguish themselves epistemologically from those derived from site owners or webmasters' linking to another webpage for referencing or other purposes.² Following Tim O'Reilly, here the terms 'web 1.0' and 'web 2.0' have been used (or overused) to periodise not only the transition from the info-web to the social web, but also from the open web to the closed web or the walled gardens of platforms (O'Reilly, 2005; Dekker and Wolfsberger, 2009).

On the Web's 25th anniversary in 2014, Tim Berners-Lee, who 'slowly, but steadily' has come to be known as its inventor, called for its 're-decentralisation', breaking down new media concentration and near monopolies online working as walled gardens without the heretofore open spirit (Berners-Lee, 2014) (Agar, 2001: 371). The web's 'app-ification' is analogous. Next to increased government Internet censorship, mass surveillance and punitive copyright laws, Berners-Lee (2014)

lists ‘corporate walled gardens’ or social media platforms as grave concerns related to the very future of the web and its mobile counterpart.

Langlois and Elmer’s point, however, implies that one should not only periodise and critique the dominant phases of the web, but also do the same for its methods of study. There are those digital methods that rely on hyperlinks, and thereby are in a sense still committed to an info-web, and those that have taken on board ‘likes’, ‘shares’ and other forms of valuation and currency (such as ‘comments’ and ‘liked comments’) on online platforms. Indeed, this analytical periodisation is reflected in the much broader study of value online, reflected in the rise of the ‘like economy’ over the ‘link economy’ (Gerlitz and Helmond, 2013). As a case in point, Google’s Web Search once valued links higher than other signals (Hindman, 2008; Rieder, 2012). Through the rise of user clicks as a source adjudication measure, one could argue that Google Web Search, too, is valuing the social web over the document or semantic matching of the info-web (van Couvering, 2007). Metrification online, which starts with like counts and follower numbers and progresses towards Klout scores, similarly considers and makes rankings social. Thus the new analytics, both Google’s updated ones as well as Klout’s, are oriented to a web gone social.

The notion of web 2.0 (and the related idea of the social web) brought with it as its apparent forerunner web 1.0 (with a more informational set of metaphors), but beyond the versioning rhetoric web 2.0 itself has been supplanted first by ‘social network(ing) sites’ and ‘platforms’ and later just by ‘social media’ (boyd and Ellison, 2007; Beer, 2008; Scholz, 2008; Allen, 2013) (see Figure 5.1). The early distinction between social networking sites and social network sites, ushered in by boyd and Ellison, was normative as well as analytical. Social media users ought to have an interest to connect with others online other than for the purposes of ‘networking’, which would suggest a kind

of neoliberal activity of making sure that even one’s social life (online) is productive. In a sense, the authors also anticipated the nuancing of social media into platform types, such as the ones for business (LinkedIn), family (Facebook) and professional doings (Twitter), though social media user practices in each remain diverse. Whether for networking or to connect with one’s existing network, the analytical call made by boyd and Ellison seemed to be directed to the study of profiles and friends (together with friending).

The purposive use of the term ‘platform’, as Tarleton Gillespie (2010) has pointed out, could be viewed as particularly enticing for users to populate an otherwise empty database, thereby generating value for the companies. Platforms connote voice-giving infrastructure, where one can express one’s viewpoints (political or otherwise), rise up, and make an online project of oneself. Polishing the profile, friending, uploading videos and photos, and liking, sharing and commenting become not only newly dominant forms of sociality, but a kind of labour for a platform owned by others (Scholz, 2016). Cooperative, user-owned platforms would provide alternatives. Other critical calls for the analysis of Facebook have been made, certain of which have resulted in invitations to leave the platform, to liberate oneself or even to commit so-called Facebook suicide, which would allow you ‘to meet your real neighbours’, as suicidemachine.org’s software project’s slogan has it (Portwood-Stacer, 2013; Facebook Liberation Army, 2015).

As web 2.0 has given way to social network(ing) sites, platforms and, finally, social media, social media methods also have evolved. In particular, digital methods for social media analysis initially relied on social network analysis (the study of inter-linked friends) as well as profiles and the presentation of self. For example, Netvizz, the Facebook data extraction software, originally was considered a tool to map one’s own Facebook friend network (Rieder, 2013). The early digital methods work on social

networking sites similarly studied friends and profiles. Dubbed ‘post-demographics’, this approach to studying profiles considered preferences and tastes as a starting point of analysis as opposed to gender, age, education and such (Rogers, 2009b). One study examined the interests of presidential candidates’ MySpace ‘friends’. Did Barack Obama’s friends and John McCain’s friends share the same favourite television shows, movies, heroes, and books, or was there a distinctive politics to media taste and consumption? (For the most part, they did not share tastes and thus TV shows and the other preferences could be considered to have politics of consumption (Rogers, 2013b).) In the case of Netvizz friend-network mapping, as well as post-demographics, these methods could be called digital methods for social media 1.0, for they concerned themselves with profiles, friends and networking.

More recently, attention to social media in digital methods work has been directed towards events, disasters, elections and revolutions, first through the so-called ‘Twitter revolution’ surrounding the Iran election crisis (2009) and later the Arab Spring (2011–2012). Instead of starting with user profiles, friend networks or networking, such studies collect tweets containing one or more hashtags such as #iranelection (perhaps together with queried keywords), or focus on one particular Facebook page, such as We are all Khaled Said (Gaffney, 2010; Lotan et al., 2011; Rieder et al., 2015).

Many of the more recent methods to analyse platforms rest upon and also derive from the individual APIs that Twitter, Facebook, Instagram, YouTube and others have to offer. As data are increasingly offered and delivered by polling one API, and no longer screen-scraped or crawled from multiple websites (as in the days of the info-web), most work is a study of a page or multiple pages (and groups) on Facebook, or one concerning tweets containing one or more hashtags and keywords on Twitter. In social media analysis with digital methods, in

other words, ‘single-platform studies’ have become the norm.

If there were a significant turning point towards single-platform studies steered by the API (rather than by scraping), it may have been the critique of a 2008 social network analysis of tastes and ties that used college students’ Facebook data (Lewis et al., 2008; Zimmer, 2010; Marres and Weltevrede, 2013). It concerned a set of presumably anonymised users from a so-called renowned university in the northeast of the United States. Not so unlike the effects of the release of AOL user search histories in 2006, its publishing prompted detective work to uncover the identities of the users, who turned out to be Harvard College students from the graduating class of 2009 (Zimmer, 2008). Michael Zimmer, both in the detective work as well as in the reflection upon the way forward for social media method, entitled his critique, ‘But the data is already public’, echoing one of the remarks by an author of the study. In giving rise to a sharper focus on ethics in web studies more generally, coinciding with a decline in scraping, Zimmer argued that in the Harvard study users’ so-called contextual privacy was violated, for not only did they not give informed content, but they did not expect their publicly available data to be stored in a researcher’s database and matched with their student housing data for even greater analytical scrutiny of their ties and tastes (Nissenbaum, 2009). The actual data collection is described by the researchers as ‘downloading’ the profile and friend network data directly from Facebook, prior to the release of Facebook API 1.0 in 2010. In other words, the data were obtained or scraped in some non-API manner, albeit with permission from Facebook as well as Harvard for the project funded by the National Science Foundation and approved by the university’s ethics review board. Ultimately, in the evolution of its API to version 2.0 (in 2014), Facebook would remove permissions to access friends’ data such as ties and tastes (i.e., friends and likes, together with profiles),

thereby making (sociometric) social network analysis like the one performed in the Harvard study improbable, including even those of one's own network with all friends' privacy settings adhered to, as one would do with Netvizz (Facebook, 2016). 'Internal' studies still may be performed, which Facebook data scientists also took advantage of with their 'emotional contagion' experiment (Kramer et al., 2014). The data science study (of some 700,000 users with a corpus of 3 million posts) analysed the risks associated with the Facebook news feed. Is user exposure to positive or negative posts psychologically risky (Meyer, 2015)? The study found that negative posts run the risk of 'emotional contagion'. In order to make the findings, Facebook selectively removed negative posts from users' news feeds. The ethics of the study were similarly questioned, for the users were unaware (and not informed) that their news feeds were being altered and their moods measured, however seemingly impractical and obtrusive it would be to gain such permission (Puschmann and Bozdog, 2014). Among the ethical issues raised, one concerned whether researchers can rely on the terms of service as cover for the otherwise lack of informed consent. Are users agreeing to being analysed for more than improvement of the site and services, as is usually stated? To the letter, they are not.

It is worthwhile to recall from the AOL case that the 62-year-old search engine user told the *New York Times* that she never imagined that her queries would be made public, or that she would have to explain to anyone that her information-seeking about medical conditions was undertaken for her friends (Barbaro and Zeller, 2006). In joining a lawsuit brought against AOL at the Federal Trade Commission, the Electronic Frontier Foundation published highly personal and salacious query histories from unnamed individuals; another user's search engine query history was made into the mini-documentary, 'I Love Alaska: The heartbreaking search history of AOL user #711391', by the Dutch

artists and filmmakers Lernert Engelberts and Sander Plug (2009), who were asked subsequently by the broadcasting company to seek out the identity of the woman, now intimately known. (Ultimately, they did not.) Neither the study of Harvard College's 2009 graduating class nor the emotional contagion study appears to have led to the subjects being identified and in some way harmed through outing. It is also not straightforward to claim that informed consent would have been enough to preclude harm, given that the users may be unable to foresee the potential hazards of participation (van de Poel, 2009).

HASHTAG AND (LIKED) PAGE STUDIES

With the decline of scraping and the rise of issues surrounding human subject research in social media, the API-led studies (on events, disasters, elections, revolutions and social causes) rely increasingly on such content-organising elements as the hashtag (for Twitter) and the (liked) page (for Facebook). Each is taken in turn, so as eventually to discuss with which limitations one may study them concurrently across platforms.

The Twitter hashtag, put forward by Chris Messina in 2007, was originally conceived as a means to set up 'channel tags', borrowing from similar practices in Internet Relay Chat (IRC). The proposal was to organise 'group-like activity' on Twitter that would be 'folksonomic', meaning user-generated rather than an editorial or taxonomic practice by the company or its syndicated partners, as in Snapchat's 'Stories' (Messina, 2007). Messina also proposed to provide a ranked list of the channel tags by activity, i.e., most active ones in the past twenty-four hours, showing on the interface where the activity is. This feature is similar to trending topics which Jack Dorsey, co-founder of Twitter, described a year later as 'what the world considers important in this moment'

(Dorsey, 2008). With hashtags and trending topics, Twitter not only gained new functionality but became a rather novel object of study for what could be termed both on-the-ground and 'remote event analysis'. As such, it thus distinguishes itself from Dorsey's original Twitter, created to provide what he called 'personal immediacy – seeing what's happening in my world right now' (Dorsey, 2008). Dorsey himself, in the interviews he gave for the *Los Angeles Times* after his temporary ouster as CEO, acknowledged the shift away from this more intimate Twitter, saying Twitter thrives on 'natural disasters, man-made disasters, events, conferences, presidential elections' (Sarno, 2009). In the event, the study of Twitter as a space for ambient friend-following yielded, at least for a share of Twitter studies, to that of event-following, which is another way of distinguishing between digital methods for social media analysis 1.0 and 2.0 (Rogers, 2013a).

Not so unlike Google Trends that list the year's most sought key words (with a geographical distribution), Twitter's initial cumulative list of the year's trending topics, published in 2009, provides a rationale for the attention granted to the study of the single hashtag for events. In the announcement made by the Twitter data scientist, Abdur Chowdhury (who incidentally was head of AOL Research when the search history data were released), one notes how serious content began to take a prominent place in a service once known primarily for its banality. In 2009 'Twitter users found the Iranian elections the most engaging topic of the year. The terms #iranelection, Iran and Tehran were all in the top-21 of Trending Topics, and #iran-election finished in a close second behind the regular weekly favorite #musicmonday' (Chowdhury, 2009). Some years later the universal list of trending topics became personalised according to whom one follows and one's geographical coordinates, however much one may change one's location and personalise trending topics exclusively by new location. In some sense the change from

universal to personalised results (like Google Web Search's similar move in December 2009, which Eli Pariser (2011) relies upon for his notion the 'filter bubble') made trends more unassailable, for no longer could one call into question why a particular hashtag (like #occupywallstreet) was not trending when it perhaps should have been (Gillespie, 2012). Trending topics are in a sense now co-authored by the Twitter user, making them less compelling to study at least as a cultural barometer. (The exception is trending topics that are location-based only.)

While the single hashtag, or more likely a combination of hashtags and keywords, remain a prominent starting point for making tweet collections to study events, disasters, elections, revolutions and social causes, as well as subcultures, movements, stock prices, celebrity awards and cities, researchers have widely expanded their repertoire for assembling them, first through techniques of capturing follower, reply and mention networks, and subsequently using the 1% random sample made available by Twitter, geotagged tweets and the Twitter ID numberspace in combination with time zones to identify national Twitter spheres (Crampton et al., 2013; Gerlitz and Rieder, 2013; Bruns et al., 2014).

Network analysis remains a preferred analytical technique in digital methods work, and as such it endures in the transition to method 2.0, but one somewhat novel strand of work worthy of mention here concerns Twitter content studies, discussed by way of a brief analytical tool description (Venturini et al., 2014b; Kennedy and Hill, 2016).

The Twitter Capture and Analysis Tool (TCAT) can be installed on one's own server to capture tweets for analysis (Borra and Rieder, 2014). Researchers thereby make individual tweet collections, instead of having one or more larger databases that are collaboratory-like repositories. Such archival fragmentation could not be avoided, because Twitter, once rather open, changed its terms of service upon becoming a publicly traded

Figure 5.2 Netvizz output showing the share, like and comments count (as well as its sum of 'engagement') of two URLs on Facebook

<i>url</i>	<i>normalized_url</i>	<i>share_ count</i>	<i>like_ count</i>	<i>comment_ count</i>	<i>total_ count</i>
http://www.nytimes.com/2015/12/08/opinion/how-isis-makes-radicals.html	http://www.nytimes.com/2015/12/08/opinion/how-isis-makes-radicals.html	1775	2667	1087	5229
https://theintercept.com/drone-papers/	https://www.theintercept.com/drone-papers/	5995	5623	2396	14014

company, no longer allowing the sharing of tweet collections (Puschmann and Burgess, 2013). Thus researchers must curate their own. The TCAT tool, installed on a server (with GitHub instructions), enables tweet collection-making (gathered from both the streaming and the REST API) and provides a battery of network analyses: social graph by mentions, social graph by in_reply to status_id, co-hashtag, bipartite hashtag-user, bipartite hashtag-mention, bipartite hashtag-URL and bipartite hashtag-host. There are also modules, however, that direct attention towards forms of content analysis that are 'quanti-quali' and referred to as 'networked content analysis' (Niederer, 2016). By quanti-quali is meant that a quantitative, winnowing analysis (not so unlike sampling) is performed so as to enable not only a 'computational hermeneutics' but also a thicker description (Mohr et al., 2015). Quanti-quali is preferred over the more usual quali-quali moniker, owing to the order of the methodological steps (Venturini et al., 2014a). Departing from a collection of 600,000 tweets gathered through a single hashtag, an example of such an approach is the #iranelection RT project, which sought to turn Twitter into a story-telling machine of events on the ground and in social media by ordering the top three retweeted tweets per day, and placing them in chronological order, as opposed to the reverse chronological order of Twitter (Rogers et al., 2009). #iranelection RT relied on manual retweeting (where the user types RT in the tweet), whereas the TCAT module outputs, chronologically, 'identical tweet frequency', or narrowly defined 'native'

retweets. Other forms of quanti-quali content analysis with a tweet collection are hashtag as well as URL frequency list-making to study hierarchies of concern and most-referred-to content. It is the starting point for a form of content analysis that treats a hashtag as (for example) an embedded social cause or movement (#blacklivesmatter) and URLs a webpage such as a news story or YouTube video. The (often fleeting) 'hashtag publics' mobilise around a social cause not only phatically (and affectively) but also with content (Bruns and Burgess, 2011; Bruns and Burgess, 2015; Papacharissi, 2015). Networked content analysis considers how and to what substantive ends the network filters stories, mobilises particular media formats over others and circulates urgency (geographically), attracting bursty or sustained attention that may be measured. Techniques of studying social causes using hashtags in Twitter as well as Instagram are discussed below, including how to consider whether to downplay or embrace medium effects.

While, since June 2013, Facebook has included hashtags as proposed means of organising 'public conversations', the straightforward 'cross-platform analysis' of Twitter and Facebook using the same hashtags is likely fraught. The study of Facebook 'content' relies far more on other activities, such as liking, sharing and commenting, which is known as studying 'most engaged with content' (and is available in the Netvizz data outputs) (see Figure 5.2). For cross-platform work, the co-appearances of URLs (aka co-links) amplified perhaps by

'likes' (Facebook's as well as Twitter's hitherto favorites) may yield far more material for comparative resonance analysis.

From the beginning Facebook (unlike Friendster and MySpace before it) positioned itself as a social network site that would reflect one's own proper circle of friends and acquaintances, thereby challenging the idea that online friends should be considered 'friends' with quotation marks and thereby a problematic category worthy of special 'virtual' study. In a sense, such a friend designation could be interpreted as another mid-2000 marker of the end of cyberspace. Together with the demise of serendipitous (and aimless) surfing, the rise of national jurisdictions legislating (and censoring) the Internet and the reassertion of local language (and local advertising) as organising principles of browsing, Facebook also re-ordered the web, doing away with cyberspace in at least two senses. As AOL once did with its portal, Facebook sought to attract and keep users by making the web 'safe,' first as a US college website offering registration only to on-campus users with an .edu email address, and then later as it expanded beyond the colleges by ID-ing users or otherwise thwarting practices of anonymisation (Stutzman et al., 2013). This was an effort to prevent so-called 'fakesters', and thus distinguish itself from online platforms like MySpace, which were purportedly rife with lurkers and stalkers as well as publicised cases of sex offenders masquerading as youngsters (boyd, 2013). Facebook's web was also clean, swept of visual clutter. In contrast to MySpace, it did not offer customisation, skinning or 'pimping', so one's profile picture and the friend thumbnails would be set in a streamlined, blue interface without starry nights, unicorns and double rainbows surrounding the posts.

Facebook's safe and de-cluttered web brought a series of 'cyberspace' research practices down to earth as well, cleaning up or at least making seem uncouth such practices as scraping websites for data. For one, scraping social network sites for data became

a (privacy and proprietary) concern and also a practice actively blocked by Facebook. Data would be served on Facebook's terms through its API (as mentioned above), and the politics and practices of APIs (more generally) would become objects of study (Bucher, 2013). In this case, terms-of-service-abiding, non-scraping data extraction tools (such as Netvizz) would reside on Facebook itself as apps, and require vetting and approval by the company. Be it through the developers' gateway or a tool on Facebook, one would log in, and the data available would respect one's own as well as the other users' privacy settings, eventually putting paid to the open-ended opportunities social network sites were thought to provide to social network research. With the API as point of access, Facebook as an object of study has undergone a transition from the primacy of the profile and friends' networks (tastes and ties) to that of the page or group, and with it from the presentation of self to social causes (which I'm using as a shorthand for events, disasters, elections, revolutions, and so forth). In a sense the company's acquisition, Instagram, could be said to have supplanted Facebook as the preferred object of study of the self through its ambassadorship of selfie culture, however much its initiator would like the company to take the route of Twitter, at once debanalising and becoming a news and event-following medium, too (Goel, 2015; Senft and Baym, 2015).

If, with the API, Facebook analysis is steered towards the pages of social causes, 'liking' is no longer considered as frivolous, and like-based engagement analyses gain more weight. As a case in point liking a page with photos of brutal acts of violence requires the like button to be re-appropriated, as Amnesty International (and other advocacy organisations) are wont to do by asking one not to take liking lightly (or communicate only phatically) but to see liking as an act of solidarity with a cause or support for a campaign. While it has been dismissed as a form of slacktivism (which requires little or

no effort and has little or no effect), liking as a form of engagement has been studied more extensively, with scholars attributing to button clicking on Facebook distinctive forms of liking causes: '(1) socially responsible liking, (2) emotional liking, (3) informational liking, (3) social performative liking, (5) low-cost liking and 6) routine liking' (Brandtzaeg and Haugstveit, 2014: 258). In the event, low-cost liking would be especially slacktivist, though all forms of liking in the list also could be construed as a form of attention-granting with scant impact, as was once said of the 'CNN effect' when all the world's proverbial eyes are watching – but not acting (Robinson, 2002). The question of whether liking as a form of engagement substitutes for other forms, however, has been challenged, for social media activism, it is argued, aids in accumulating action and action potential (Christensen, 2011). It is also where the people are (online).

FROM SINGLE PLATFORM TO CROSS-PLATFORM STUDIES

Social movement, collective action and more recently 'connective action' researchers in particular have long called for multiple platform, and multi-media, analysis (to use an older term). In an extensive study based on interviews, Sasha Costanza-Chock (2014), for one, has deemed the immigrant rights movement in the United States a form of 'transmedia organising'. The cross-platform approach is a deliberate strategy, and each platform is approached and utilised separately for its own qualities and opportunities. Here one may recall the distinction made by Henry Jenkins (2006) between cross-media (the same story for all platforms) and trans-media (the story unfolds differently across platforms). Thus social media, when used as a 'collapsed category', masks significant differences in 'affordances' (Costanza-Chock, 2014: 61–66). (I return to a similar problem

concerning collapsed digital objects such as hashtags or likes across platforms with different user cultures.) If we are to follow Jenkins, as well as Costanza-Chock, a discussion of cross-platform analysis would be more aptly described as trans-platform analysis.

Researchers studying social causes on platforms have also called for 'uncollapsing' social media. Lance Bennett and Alexandra Segerberg, who coined the notion of 'connective action' as a counter-point to collective action, argue that to understand the forces behind social change one should study *those multiple platforms* that allow for 'personalized public engagement', instead of choosing one platform and its API in advance of the analysis (Bennett and Segerberg, 2012). It is, in other words, an implicit critique of the single-platform studies (as collapsed social media studies) that rely solely on Twitter for one issue (e.g., Fukushima in Japan) or Facebook for another (e.g., rise of right-wing populism), when one could have ample cause to study them across media. It is not only the silo-ing of APIs that prompts single-platform studies; as pointed out, the question of the comparability of the 'same' objects across platforms (likes, hashtags) is at issue.

One of Bennett and Segerberg's preferred tools is the IssueCrawler, developed at the Digital Methods Initiative, which could be described as web 1.0 analytical software, relying on the info-web's link and performing hyperlink analysis. For multiple-platform (and transmedia) analysis à la Bennett and Segerberg it could be employed as an exploratory instrument at the outset of a study of a cause (on the web), in order to ascertain which websites (including blogs) and platforms are the focus of attention. In other words, hyperlink analysis could be construed as a web 1.0 methodological starting point for multi-platform analysis. As described below, other 'inter-linkings' (broadly conceived) may be studied, such as co-linked and inter-liked content.

PLATFORM CULTURES OF USE

The purpose of the exercise here is to develop cross-platform methods, or digital methods for cross-platform studies, where one learns from medium methods and repurposes them for social and cultural research. It begins with a sensitivity to distinctive user cultures and subcultures, whereby hashtags and likes, digital objects used to organise and boost content (among other reasons), should not necessarily be treated as if they are employed equivalently across all platforms, even when present. For example, Instagram has inflated hashtag use compared to Twitter's, allowing up to 30 tags (and far more characters per photo caption post than Twitter grants for a tweet). That is, users may copy and paste copious quantities of hashtags in Instagram posts (see Table 5.1). Twitter recommends that one '[does not] #spam #with #hashtags. Don't over-tag a single Tweet. (Best practices recommend using no more than 2 hashtags per Tweet.)' (Twitter, 2016). While present, hashtags are under-utilised on Facebook.

A series of questions arises concerning the meaning of the term 'cross' in 'cross-platform analysis'. First, across which platforms are 'hashtags' worthy of study (Twitter, Instagram, Tumblr), which ones 'likes' (Facebook, Instagram, YouTube, Twitter, Pinterest), which ones 'retweets' or 'repins' (Twitter, Pinterest), which one '@mentions' (Twitter), which ones 'links', including shortened URLs (not Instagram), and so forth (see Table 5.1)? The point is that platforms have similar affordances, such as like buttons and hashtags, but one should not necessarily collapse them by treating them equally across platforms. More specifically, if one were to perform cross-platform analysis of the same hashtags across multiple platforms, how would one build into the method the difference in hashtag use in Twitter and Instagram? Because of hashtag proliferation on Instagram, does one devalue or otherwise correct for hashtag abundance on the one platform while

valuing it steadily on another? One could strive to identify cases of copy-and-pasting hashtag strings, and downplay their value, certainly if posts are being 'stuffed' with hashtags.

Second, certain platforms (and perhaps more so certain topics such as large media events on most any platform) may indeed have user cultures and automation activity that routinely befoul posts as well as activity measures. Hashtag hijacking is a case in point, especially when one is studying an event or a social issue and encounters unrelated hashtags purposively inserted to attract attention and traffic, such as when spammers monitor trending hashtags and use them tactically to promote their wares. Hashtag junk may distract at least the researcher.

Third, while a more complex topic, bots and the activity traces they leave behind are often similarly considered worth special consideration during the analysis (Marres, 2015). From a digital forensics point of view, bots that like and follow may have specific (network) signatures, e.g., they do not tend to be followed, or to be liked, meaning the bot often only has outlinks. For the purposes of this discussion, they may inflate activity in causes and such inflation may be considered artificial (though of course there are bots created for events and issues, too, and their activities are thereby purposive). Thus manipulation as well as artificiality are additional (intriguing) complications in both single-platform and cross-platform analysis.

Fourth, platforms have 'device cultures' that affect how one interprets the data from the API.³ That is, all platforms filter posts, showing particular content and letting other content slide downwards or off screen, so to speak (Eslami et al., 2015). Users thereby cannot 'like' all content equally. That which is liked may tend to be liked more often, and thus there may be power law and long tail effects that differ per platform. But we may not know how preferred posting affects activity measures. APIs will return like and share counts (for example) per post, but they do not let us know the extent to which all the content

has been equally visible to those who would be able to like, share, comment, and so forth. And filtering styles and thus visibility effects differ per platform.

Above a series of questions has been posed concerning the limitations of comparing evaluations of content, recommended with the same type of button on different platforms, given that the platforms may have different user, spamming, bot and device cultures. How to nevertheless undertake cross-platform analysis? When studying recommendations and the content that rises, metrically, to the top of the platforms, it may be instructive to begin by examining briefly which digital objects are available in each of the platforms (as above and in Table 5.1) and subsequently enquire into how dominant devices (or in this case metrics such as Klout) handle these objects. Subsequently, it is asked, how to repurpose the metrics?

CROSS-PLATFORM ANALYSIS: CO-LINKED, INTER-LIKED AND CROSS-HASHTAGGED CONTENT

Klout, as the term indicates, measures a user's 'clout', slang for influence, largely from data culled online, where the user is not only an individual but can be a magazine, institution, professional sports team, etc. Klout scores are measured on the basis of activity on Twitter, Facebook, YouTube, Google+, LinkedIn, Instagram, and Foursquare (Rao et al., 2015). It is an influence measure that takes into account particular appearance signals across the seven platforms (e.g., mentions on Twitter), and those mentions by highly influential user accounts grant more influence or clout to the user in question. It also grounds (and augments) the online appearance measures with 'offline factors' that take into account a user's 'real world influence' from Wikipedia as well as resonance in news articles (Rao et al., 2015: 3). Job titles, years of experience

and similar from LinkedIn are also factored in. It is also a computationally intensive, big data undertaking and an aggregated form of cross-platform analysis.

If one were to learn from Klout for social research, one manner would be to shift the focus from power (measures of increases or decreases in one's influence) to matters of concern (increases or decreases in attention, including that from significant others) – be these to events, disasters, elections, revolutions, social causes, and so forth. The shift in focus would be in keeping with how social media is often currently studied, as discussed above. That is, one could apply Klout's general procedure for counting user appearances, and ask, which causes are collectively significant across social media platforms, and which (key) actors, organisations and other users are linked to them, thereby granting them attention. Just as importantly, the attention granted to a cause by key actors, organisations and users may be neither undivided nor sustained. Such an observation would invite inquiries into partial attention as well as attention span, which together could begin to form a means to study engagement across social media.

When can so-called info-web methods based on the hyperlink still be applied to the study of the web and its platforms? By 'http or html approaches' to web 1.0, I mean software like the Issuecrawler and other hyperlink analysis tools, which, generally speaking, crawl a seed list of websites, locate hyperlinks either between them or between them and beyond them, and map the interlinkings, showing uni-directional, bi-directional as well as the absence of linking between websites (see Figure 5.3). Problems arise. Through automated hyperlink analysis, the researcher may miss relationships between websites which are not captured by hyperlinks, such as sites mentioning each other in text without linking. One may also miss links between websites because servers are down, or javascript or other code impenetrable to crawlers are employed on one or

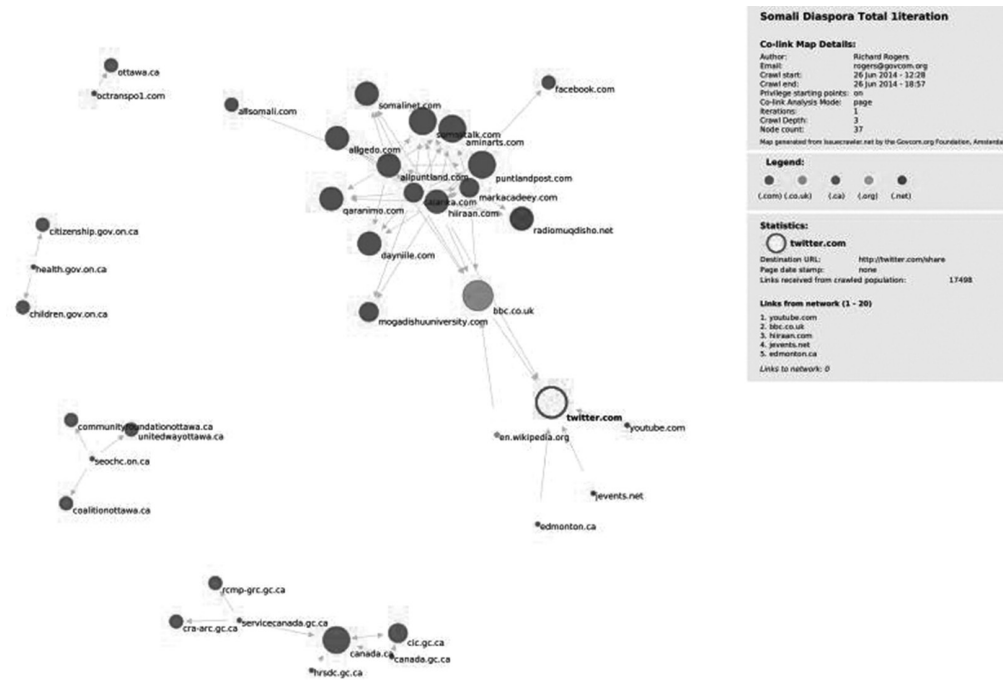


Figure 5.3 IssueCrawler map showing Twitter.com as significant node, albeit without showing individual, significant Twitter users

(Source: Issuecrawler.net, June 2014)

more websites in the network. Elmer and Langlois (2013) thereby proposed to follow keywords across websites as well as platforms.

As the info-web has evolved into a social web, hyperlink analysis generally captures

links between pages or hosts on the web, but not on social media platforms, where only the host is returned (Facebook.com) rather than individual user profiles, such as a Facebook account, page or group or an individual Twitter user. (Similarly, Google continually

Table 5.1 Elements of cross-platform analysis (adapted from Rieder, 2015)

	<i>Twitter</i>	<i>Facebook</i>	<i>Instagram</i>
Query design	Hashtag(s), keyword(s), location(s), user(s)	Group(s), page(s)	Hashtag(s), location(s)
Data capture	In advance (for overtime data); on demand (for very recent data)	On demand (for overtime and recent data)	On demand (for overtime location data and recent hashtag data)
Platform user accounts (with primary actions)	user (follow)	user (friend, follow), group (join), page (like)	user (follow)
Content (media contents and digital objects)	tweet (text, photo, video, hashtag, @mention, URL, geotag)	post (text, video, photo, URL)	photo, video (text, hashtag, geotag, @mention)
Activities (resonance measures)	like (fav), retweet	like, comment, share	like, comment

experiments with how its web search returns Twitter and Facebook content, although it still privileges web content.) These drawbacks have occasioned researchers to move in two directions at once: develop crawlers and hyperlink analytical machines that pinpoint deep links between social media platforms and websites as well as within platforms (such as the Hyphe project⁴), and to consider new means to study relationships between platforms as well as between platforms and the web that do not rely on hyperlinks only. Joining in part with the call by Elmer and Langlois (2013), here the proposal would be to study content across the platforms (and the web): which content is co-linked, inter-liked and/or cross-hashtagged?

Co-linked content are URLs (often shortened on social media) that are linked by two or more users, platform pages or webpages. Inter-liked content is content liked by users and pages across platforms. Cross-hashtagged content is content referred to by hashtags across platforms. As they are often embedded social issues (and events), the hashtags themselves could be considered the content.

RESEARCH STRATEGIES FOR CROSS-PLATFORM ANALYSIS

We might ask, then, how to perform cross-platform analysis, and which platforms may be productively compared. When discussing the kind of research done with social media, even with the shift to the study of social causes over the self, it is worthwhile to point out that one may emphasise medium research, social research, or a combination of the two. For medium research, the question concerns how the platform affects the content, be it its presence or absence as well as its orderings. Additionally, specific cultures of use per platform, and (strategic) transmedia deployment, may inform the medium research, as discussed above. For social research, the

question concerns the story the content tells, despite the platform effects. For a combination of medium and social research, the questions are combined; how does the platform affect the availability of content, and what stories do the content tell, given platform effects? Thus for cross-platform analysis, the following steps may be taken.

- 1 Choose a contemporary issue (revolution, disaster, election, social cause, and so forth) for cross-platform analysis. One may choose to follow an active or unfolding issue (an issue in motion, so to speak), or one from recent history (an issue from the past, where overtime analysis is desirable). Here one should consider which platforms provide overtime data (Facebook), and which do not without great effort (Twitter).
- 2 Design a query strategy. For social issues and causes, consider querying for a program and an anti-program (Rogers, 2017). For example, in the 2015 US Supreme Court ruling for same-sex marriage the competing Twitter and Instagram hashtags reflected hashtag publics forming around a program and an anti-program, #lovewins and #jesuswins, respectively. If hashtags are preferred, for an election, consider querying a set of candidates or parties, e.g., #Trump and #Hillary (perhaps together with additional hashtags as well as keywords). For a disaster (or tragedy), consider querying its name(s), e.g., #MH17. URLs and/or domain names can be used as queries for a number of platforms.
- 3 Develop an analytical strategy. For social issues and causes, consider which program or anti-program is finding favour (including among whom and where). Does it have a set of networked publics and a particular geography? For an election, consider creating portrayals of the candidates via the associated issues, or comparing their relative resonance with current election polls. For a revolution, consider its momentum and durability (including the subjects that continue to matter and those that do not endure). For a disaster, consider how it is (continually) remembered or forgotten, and to which extent it has been and still is addressed and by whom.
- 4 Consider the configuration of use. It may be instructive for the analysis to look into how the platform is configured and set up by the

- initiator(s). Is it a group or a page, with or without moderation? Is it centrally organised or a collective effort? Are comments allowed? Does the user have a distinctive follower strategy?
- 5 Cross-platform analysis. Undertake the platform analysis, according to the query design strategy as well as the analytical strategy discussed above, across two or more platforms. For each platform consider engagement measures, such as the sum of likes, shares, comments (Facebook), likes and retweets (Twitter) and co-hashtags (Instagram). Which (media) content resonates on which platforms? Consider which content is shared across the platforms (co-linked, inter-linked and cross-hashtagged), and which is distinctive, thereby enabling both networked platform content analysis as well as medium-specific (or platform-specific) effects.
 - 6 Discuss your findings with respect to medium research, social research or a combination of the two. Does a particular platform tend to host as well as order content in ways distinctive from other platforms? Are the accounts of the events distinctively different per platform or utterly familiar no matter the platform?

In practice certain platforms lend themselves to comparison more artfully than others, given both the availability of objects such as the hashtag or geotag as well as roughly similar cultures of use. Through the vehicle of the hashtag, Twitter and Instagram (as well as Tumblr) are often the subject of cross-platform analysis. One queries the APIs with such tools as TCAT (for Twitter) as well as relatively simple Instagram and Tumblr hashtag explorers made available by the Digital Methods Initiative, creating collections of tweets and posts for further quantitative and qualitative analysis. Take, for example, certain significant events in the so-called migration crisis in Europe, one concerning the death of refugee children (Aylan Kurdi and his brother) and another the sexual assaults and rapes on New Year's Eve in Cologne (Geboers et al., 2016). For each case Twitter and Instagram are queried for hashtags (e.g., #aylan), whereupon tweet and post collections are made. For Twitter, one 'recipe' to sort

through the contents of the collections would include the following:

- (a) Hashtag Frequency counts ascertain the other hashtags that co-occur, and is useful to explore the issue space. For the Cologne rape cases, the hashtag #einearmlänge co-occurs greatly, which was a trending topic referring to the remarks by the Cologne mayor that (as a solution) women should remain an arm's length away from so-called strangers.
- (b) Mention Frequency lists the usernames of those who tweet and who are mentioned so one notes which users may dominate a space.
- (c) Retweet Frequency provides a ranked list of retweeted tweets, showing popular or significant content.
- (d) URL Frequency is a ranked URL list showing popular or significant media (such as images and video). The most referenced media, especially images, become a focal point for a cross-platform analysis with Twitter.

For Instagram, hashtag frequency is undertaken together with image and video frequency. (One is also able to query Instagram for geo-coordinates, which is not undertaken here.) Ultimately, the means of comparison are hashtag as well as image and video use, where the former suffers somewhat from hashtag stuffing in Instagram.

The question of platform effects is treated in the qualitative analysis, where in both the Aylan as well as the Cologne New Year's Eve cases the incidence of news photos was much greater in Twitter than in Instagram, where there were more derivatives, meaning annotated, photoshopped, cartoon-like or other DIY materials with (implied or explicit) user commentary. Twitter thereby becomes a professional medium (with effects) and Instagram more a user-generated content medium, becoming a particular, user-led form of news-following platform to which its founder has been aspiring. The Aylan case, however, appears to reduce this medium-specificity, because there is a relatively greater amount of images which have been edited so as to come to grips with the tragedy of the drowned toddler.

CONCLUSIONS: DIGITAL METHODS FOR CROSS-PLATFORM ANALYSIS

In the call for methodological attention to the platformisation of the web, Elmer and Langlois (2013) discuss how analyses based on the hyperlink do not embrace the analytical opportunities afforded by social media. Hyperlink analysis, and its tools such as the IssueCrawler, rely on an info-web (aka web 1.0), where webmasters make recommendations by linking to another website (or non-recommendations through not making links, thereby showing lack of interest or affiliation). Focusing on links only misses the novel objects of web 2.0, social networking sites, platforms and social media (as the social web has been called), such as the like, share and tweet. While Elmer and Langlois (2013) called for the analysis of the keyword over the hyperlink, but also perhaps over other social media objects, around the same time as their publication the API had arrived (Facebook's version 1.0 in 2010, Twitter's in 2006), and gradually became the preferred point of access to data over scraping, which the platforms actively sought to thwart. The API is of course controlled by the service in question, be it Twitter, Facebook or others, and steers research in ways more readily palpable perhaps than scraping, for the data available on the interface (that could be scraped) and through the developer's entry point may differ considerably. The ethics turn in web research, bound up with the rise of the social web and its publicly available, personal data, in turn has shaped the accessibility of certain data on the APIs such that Facebook no longer allows one to collect friends' 'tastes and ties', or likes, profile interests as well as friends. Such unavailability comes on the heels of a critique of a study of the same name that collected (or scraped, albeit with permission) Facebook profiles and friends' data from Harvard students and enriched it with their student housing information, without their knowledge. Concomitant with the decline in the study of

the self in social media analysis with digital methods (given the increasing dearth of available data through API restrictions) has been the rise in attention to events, disasters, elections, revolutions and social causes. Not only is it in evidence in Facebook research on (Arab Spring) pages (and to an extent groups), but also in Twitter (revolutions), where Jack Dorsey, its co-founder, signalled the shift in the interviews in the *Los Angeles Times* in 2009, mentioning that Twitter did well events such as disasters, elections as well as conferences. Instagram, according to its founder Kevin Systrom, would like to follow the same trajectory, becoming a platform of substance and thereby for the study of events (Goel, 2015). The API, however, appears to have shaped social media studies beyond its selective availability of data. Rather, the APIs serve as silos for what I call 'single-platform studies', which are reflected in the available tools discussed. Netvizz is for Facebook studies, TCAT for Twitter studies, the Instagram hashtag explorer for Instagram, and so forth. Unlike the web 1.0 tools such as IssueCrawler, which find links between websites and between websites and platforms, the social web has not seen tools developed for cross-platform analysis. Where to begin?

The purpose here is to develop techniques for multiple platform analysis that bear medium-sensitivity. Stock is taken of the objects that platforms share, whereupon cultures of use are taken into consideration. In other words, Twitter, Facebook and Instagram share the hashtag, however much, on the one, no more than two are recommended, on another it is rarely used and on the third it is used in overabundance. The cross-platform approaches that are ultimately described rely on hashtags for making collections of tweets (in Twitter) and posts (in Instagram), whereupon the media format (images, but also videos) common to the two are compared in the study of events. During the European refugee crisis of 2015–2016, the death of the toddler, Aylan Kurdi, and the sexual assaults of

women in Cologne stand out as major (social media) events for analysis with a quanti-quali approach and a networked content analysis, which are forms of analysis with affinities with computational hermeneutics.

SUGGESTED RESOURCES

For tool tutorials, see the DMI ‘tools walk-through’ playlist on YouTube, www.youtube.com/playlist?list=PLKzQwIKtJvv9lwyYxh4708Nqo6YC6-YH4

1 Instagram

Instagram hashtag explorer, aka Visual Tagnet Explorer

<http://tools.digitalmethods.net>

Video tutorial for Instagram hashtag explorer, ‘Analyze Instagram Activity Around a Hashtag or Location.’ Note: since Instagram has blocked researcher use of its API in June 2016, one workaround is to locate and insert a token.

www.youtube.com/watch?v=o07aUKdRv0g

2 Twitter

DMI-TCAT (Twitter Capture and Analysis Tool)

<https://github.com/digitalmethodsinitiative/dmi-tcat/wiki>

Video tutorial for TCAT, ‘Overview of Analytical Modules’

www.youtube.com/watch?v=ex97eoorUeo

3 Facebook

Netvizz (Facebook Data Extraction Tool)

<https://apps.facebook.com/netvizz/>

Netvizz video tutorials:

‘Introduction to Netvizz 1.2+’

www.youtube.com/watch?v=3vkKPCn7V7Q

‘Downloading data and producing a macro view’

www.youtube.com/watch?v=df0YAPistYg

4 Gephi-related

Gephi (The Open Graph Viz Software)

<https://gephi.org>

‘Gephi Tutorial for working with Twitter mention networks’

www.youtube.com/watch?v=snPR8CwPld0

‘Combine and Analyze Co-Hashtag Networks (Instagram, Twitter, etc.) with Gephi’

www.youtube.com/watch?v=ngqWjgZudeE

Notes

- 1 There certainly were social aspects to the early web, however much its dominant devices (Google web search, and Altavista before it) were oriented less to sociality than information compared to online platforms of a later period.
- 2 More specifically, these days recommendations could be said to be co-authored by the user and the system, whereas previously they were made by the site owner.
- 3 ‘Device culture’ studies would inquire into the chain of interactions between user and platform that results in data collected and system-analysed so that ultimately content is recommended recursively back to the user (Rogers et al., 2013; Weltevrede, 2016).
- 4 See the Hyphe project at the MediaLab, Sciences Po, Paris, <http://hyphe.medialab.sciences-po.fr/>.

REFERENCES

- Agar, John (2001). ‘Review of James Gillies and Robert Cailliau, *How the Web was Born*. Oxford: Oxford University Press, 2000,’ *The British Journal for the History of Science* 34(3): 370–373.
- Allen, Matthew (2013). ‘What Was Web 2.0? Versions as the Dominant Mode of Internet History,’ *New Media & Society*, 15(2): 260–275.
- Barbaro, Michael and Tom Zeller Jr (2006). ‘A Face is Exposed for AOL Searcher no. 4417749,’ *The New York Times*, p. A1.
- Beer, David (2008). ‘Social Network(ing) Sites... Revisiting the Story so far: A Response to

- danah boyd & Nicole Ellison, 'Journal of Computer-Mediated Communication, 13(2): 516–529.
- Bennett, W. Lance and Alexandra Segerberg (2012). 'The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics,' *Information, Communication & Society*, 15(5): 739–768.
- Berners-Lee, Tim (2014). 'Tim Berners-Lee on the Web at 25: the past, present and future,' *Wired*, 23 August.
- Borra, Erik and Bernhard Rieder (2014). 'Programmed Method: Developing a Toolset for Capturing and Analyzing Tweets,' *Aslib Journal of Information Management*, 66(3): 262–278.
- boyd, danah (2013). 'White Flight in Networked Publics? How Race and Class Shaped American Teen Engagement with Myspace and Facebook,' in Lisa Nakamura and Peter A. Chow-White (eds.), *Race after the Internet*. New York: Routledge, pp. 203–222.
- boyd, danah and Nicole Ellison (2007). 'Social Network Sites: Definition, History and Scholarship,' *Journal of Computer-Mediated Communication*, 13(1), article 1.
- Brandtzaeg, Petter Bae and Ida Maria Haugstveit (2014). 'Facebook Likes: A Study of Liking Practices for Humanitarian Causes,' *International Journal of Web Based Communities*, 10(3): 258–279.
- Bruns, Axel (2007). 'Methodological for Mapping the Political Blogosphere: An Exploration Using the IssueCrawler Research Tool,' *First Monday*, 12(5).
- Bruns, Axel and Jean Burgess (2015). 'Twitter Hashtags from Ad Hoc to Calculated Publics,' in Nathan Rambukkana (ed.), *Hashtag Publics: The Power and Politics of Discursive Networks*. New York: Peter Lang, pp. 13–28.
- Bruns, Axel and Jean Burgess (2011). 'The use of Twitter hashtags in the formation of ad hoc publics,' Proceedings of the 6th European Consortium for Political Research (ECPR) General Conference 2011, University of Iceland, Reykjavik.
- Bruns, Axel, Jean Burgess and Tim Highfield (2014). 'A "Big Data" Approach to Mapping the Australian Twittersphere,' in Paul Longley Arthur and Katherine Bode (eds.), *Advancing Digital Humanities: Research, Methods, Theories*. Basingstoke: Palgrave Macmillan, pp. 113–129.
- Bucher, Taina (2013). 'Objects of Intense Feeling: The Case of the Twitter API,' *Computational Culture: A Journal of Software Studies*, 4.
- Chowdhury, Abdur (2009). 'Top Twitter Trends of 2009,' Twitter blog, 15 December, https://blog.twitter.com/official/en_us/a/2009/top-twitter-trends-of-2009.html.
- Christensen, Henrik Serup (2011). 'Political Activities on the Internet: Slacktivism or Political Participation by Other Means?' *First Monday*, 16(2).
- Constanza-Chock, Sasha (2014). *Out of the Shadows, Into the Streets! Transmedia Organising and the Immigrant Rights Movement*. Cambridge, MA: MIT Press.
- Crampton, Jeremy W., Mark Graham, Ate Poorthuis, Taylor Shelton, Monica Stephens, Matthew W. Wilson and Matthew Zook (2013). 'Beyond the geotag: situating 'big data' and leveraging the potential of the geoweb,' *Cartography and Geographic Information Science*, 40(2): 130–139.
- Dekker, Annet and Annette Wolfsberger (2009). *Walled Garden*. Amsterdam: Virtual Platform.
- Dorsey, Jack (2008). 'Twitter Trends & a Tip,' Twitter blog, 5 September, https://blog.twitter.com/official/en_us/a/2008/twitter-trends-a-tip.html.
- Elmer, Greg and Ganaele Langlois (2013). 'Networked Campaigns: Traffic Tags and Cross Platform Analysis on the Web,' *Information Polity*, 18(1): 43–56.
- Engelberts, Lernert and Sander Plug (2009). 'I Love Alaska: The Heartbreaking Search History of AOL User #711391,' Minimovies Documentary. Amsterdam: Submarine Channel.
- Facebook (2016). 'Facebook Platform Changelog,' Facebook for Developers, webpage, <https://developers.facebook.com/docs/apps/changelog>.
- Facebook Liberation Army (2015). *Directives*. Amsterdam: Waag Society, <http://fla.waag.org/downloads/FLA-Infographic.pdf>.
- Gaffney, Devin (2010). '#iranElection: Quantifying online activism,' *Proceedings of WebSci10*. New York: ACM.
- Geboers, Marloes, Jan-Jaap Heine, Nienke Hidding, Julia Wissel, Marlie van Zoggel and Danny Simons (2016). 'Engagement with

- Tragedy in Social Media,' Digital Methods Winter School '16, Amsterdam, <https://wiki.digitalmethods.net/Dmi/WinterSchool2016> EngagementWithTragedySocialMedia.
- Gerlitz, Carolin and Anne Helmond (2013). 'The Like Economy: Social Buttons and the Data-intensive Web,' *New Media & Society*, 15(8): 1348–1365.
- Gerlitz, Carolin and Bernhard Rieder (2013). 'Mining One Percent of Twitter: Collections, Baselines, Sampling,' *M/C Journal*, 16(2).
- Gillespie, Tarleton (2010). 'The Politics of "Platforms",' *New Media & Society*, 12(3): 347–364.
- Gillespie, Tarleton (2012). 'Can an Algorithm be Wrong?' *Limn*, 2.
- Goel, Vindu (2015). 'Instagram to Offer Millions of Current Events Photos,' *New York Times*, 23 June.
- Helmond, Anne (2015). 'The Web as Platform: Data Flows in Social Media,' PhD dissertation, University of Amsterdam.
- Hindman, Matthew (2008). *The Myth of Digital Democracy*. Princeton, NJ: Princeton University Press.
- Ito, Mizuko (2008). 'Introduction,' in Kazys Varnelis (ed.), *Networked Publics*. Cambridge, MA: MIT Press, pp. 1–14.
- Jenkins, Henry (2006). *Convergence Culture*. New York: New York University Press.
- Eslami, Motahhare, 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]": Reasoning about Invisible Algorithms in News Feeds,' *CHI 2015, Crossings*, Seoul, South Korea. New York: ACM.
- Kennedy, Helen and Rosemary Lucy Hill (2016). 'The Pleasure and Pain of Visualizing Data in Times of Data Power,' *Television & New Media*, first published 7 September, pp. 1–14, 10.1177/1527476416667823.
- Kramer, Adam D.I., Jamie E. Guilloryb and Jeffrey T. Hancock (2014). 'Experimental Evidence of Massive-scale Emotional Contagion through Social Networks,' *PNAS*, 111(24): 8788–8790.
- Latour, Bruno (2005). *Reassembling the Social*. New York: Oxford University Press.
- Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer and Nicholas Christakis (2008). 'Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com,' *Social Networks*, 30(4): 330–342.
- Lotan, Gilad, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce and danah boyd (2011). 'The Revolutions were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions,' *International Journal of Communication*, 5: 1375–1405.
- Marres, Noortje (2015). 'Why Map Issues? On Controversy Analysis as a Digital Method,' *Science, Technology, & Human Values*, 40(5): 655–686.
- Marres, Noortje and Esther Weltevrede (2013). 'Scraping the Social? Issues in Live Social Research,' *Journal of Cultural Economy*, 6(3): 313–335.
- Messina, Chris (2007). 'Groups for Twitter; or A Proposal for Twitter Tag Channels,' *Factory Joe blog*, 25 August, <http://factoryjoe.com/blog/2007/08/25/groups-for-twitter-or-a-proposal-for-twitter-tag-channels/>.
- Meyer, Michelle N. (2015). 'Two Cheers for Corporate Experimentation: The A/B Illusion and the Virtues of Data-Driven Innovation,' *Colorado Technology Law Journal*, 13(2): 273–332.
- Mohr, John W., Robin Wagner-Pacifici and Ronald L. Breiger (2015). 'Toward a Computational Hermeneutics,' *Big Data & Society*, 2(2).
- Niederer, Sabine (2016). 'Networked Content Analysis: The Case of Climate Change,' PhD dissertation, University of Amsterdam.
- Nissenbaum, Helen (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press.
- O'Reilly, Tim (2005). 'What is Web 2.0: Design Patterns and Business Models for the Next-Generation of Software,' O'Reilly, Sebastopol, CA: O'Reilly Media, www.oreilly.com/pub/a/web2/archive/what-is-web-2.0.html.
- Papacharissi, Zizi (2015). *Affective Publics*. New York: Oxford University Press.
- Pariser, Eli (2011). *The Filter Bubble*. New York: Penguin.
- Portwood-Stacer, Laura (2013). 'Media Refusal and Conspicuous Non-consumption: The Performative and Political Dimensions of Facebook Abstention,' *New Media & Society*, 15(7): 1041–1057.

- Puschmann, Cornelius and Jean Burgess (2013). 'The Politics of Twitter Data,' in Katrin Weller, Axel Bruns, Jean Burgess, Merja Mahrt and Cornelius Puschmann (eds.), *Twitter and Society*. New York: Peter Lang, pp. 43–54.
- Puschmann, Cornelius and Engin Bozdag (2014). 'Staking Out the Unclear Ethical Terrain of Online Social Experiments,' *Internet Policy Review*, 3(4).
- Rao, Adithya, Nemanja Spasojevic, Zhisheng Li and Trevor DSouza (2015). 'Klout Score: Measuring Influence across Multiple Social Networks,' *2015 IEEE International Big Data Conference – Workshop on Mining Big Data in Social Networks*. New York: ACM.
- Rieder, Bernhard (2012). 'What is PageRank? A Historical and Conceptual Investigation of a Recursive Status Index,' *Computational Culture*, 2.
- Rieder, Bernhard (2013). 'Studying Facebook via Data Extraction: The Netvizz Application,' in *Proceedings of the 5th Annual ACM Web Science Conference*. New York: ACM Press, pp. 346–355.
- Rieder, Bernhard (2015). 'Social Media Data Analysis,' lecture delivered at the University of Amsterdam, December.
- 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): 1–22.
- Robinson, Piers (2002). *The CNN Effect: The Myth of News, Foreign Policy and Intervention*. London: Routledge.
- Rogers, Richard (2009a) 'Mapping Public Web Space with the Issuecrawler,' in C. Brossard and B. Reber (eds.), *Digital Cognitive Technologies: Epistemology and Knowledge Society*. London: Wiley, pp. 115–126.
- Rogers, Richard (2009b). 'Post-demographic Machines,' in Annet Dekker and Annette Wolfsberger (eds.), *Walled Garden*. Amsterdam: Virtual Platform, pp. 29–39.
- Rogers, Richard (2013a). 'Debanalizing Twitter: The Transformation of an Object of Study,' *Proceedings of WebSci13*. New York: ACM.
- Rogers, Richard (2013b). *Digital Methods*. Cambridge, MA: MIT Press.
- Rogers, Richard (2017). 'Foundations of Digital Methods: Query Design,' in Mirko Schaefer and Karin van Es (eds.), *The Datafied Society: Studying Culture through Data*. Amsterdam: Amsterdam University Press.
- Rogers, Richard, Esther Weltevrede, Erik Borra, Marieke van Dijk and the Digital Methods Initiative (2009). 'For the ppl of Iran: #iranlection RT,' in Gennaro Ascione, Cinta Massip and Josep Perello (eds.), *Cultures of Change: Social Atoms and Electronic Lives*. Barcelona: Actar and Arts Santa Monica, pp. 112–115.
- Rogers, Richard, Esther Weltevrede, Sabine Niederer and Erik Borra (2013). 'National Web Studies: The Case of Iran Online,' in John Hartley, Axel Bruns and Jean Burgess (eds.), *A Companion to New Media Dynamics*. Oxford: Blackwell, pp. 142–166.
- Sarno, David (2009). 'Jack Dorsey on the Twitter Ecosystem, Journalism and How to Reduce Reply Spam: Part II,' *Los Angeles Times*, 19 February.
- Scholz, Trebor (2008). 'Market Ideology and the Myths of Web 2.0,' *First Monday*, 13(3).
- Scholz, Trebor (2016). 'Platform Cooperativism. Challenging the Corporate Sharing Economy,' New York: Rosa Luxemburg Stiftung.
- Senft, Theresa and Nancy Baym (2015). 'What Does the Selfie Say? Investigating a Global Phenomenon,' *International Journal of Communication*, 9: 1588–1606.
- Stutzman, Fred, Ralph Gross and Alessandro Acquisti (2013). 'Silent Listeners: The Evolution of Privacy and Disclosure on Facebook,' *Journal of Privacy and Confidentiality*, 4(2), article 2.
- Twitter (2016). 'Using Hashtags on Twitter,' Help Center, San Francisco, CA: Twitter, Inc., <https://support.twitter.com/articles/49309>
- Van Couvering, Elizabeth (2007). 'Is Relevance Relevant? Market, Science, and War: Discourses of Search Engine Quality,' *Journal of Computer-Mediated Communication*, 12(3), article 6.
- Van de Poel, Ibo (2009). 'The Introduction of Nanotechnology as a Societal Experiment,' in Simone Araldi, Andrea Lorenzet and Federica Russo (eds.), *Technoscience in Progress: Managing the Uncertainty of Nanotechnology*. Amsterdam: IOS Press, pp. 129–142.
- Venturini, Tommaso, Dominique Cardon and Jean-Philippe Cointet (2014a). 'Présentation – Méthodes digitales: Approches quali/quantitatives des données numériques,' *Réseaux*, 188(6): 9–21.

- Venturini, Tommaso, Nicolas Baya Laffite, Jean-Philippe Cointet, Ian Gray, Vinciane Zabban and Kari De Pryck (2014b). 'Three Maps and Three Misunderstandings: A Digital Mapping of Climate Diplomacy,' *Big Data & Society*, 1(2).
- Weltevrede, Esther (2016). 'Repurposing Digital Methods: The Research Affordances of Engines and Platforms,' PhD dissertation, University of Amsterdam.
- Zimmer, Michael (2008). 'More on the "Anonymity" of the Facebook Dataset – It's Harvard College,' Michaelzimmer.org blog, <http://michaelzimmer.org/2008/10/03/moreon-the-anonymity-of-the-facebook-dataset-its-harvard-college/>
- Zimmer, Michael (2010). 'But the Data is Already Public: On the Ethics of Research in Facebook,' *Ethics and Information Technology*, 12(4): 313–325.