

New Methodologies for Researching News Discussion on Twitter

Axel Bruns & Jean Burgess

ARC Centre of Excellence for Creative Industries and Innovation

Queensland University of Technology

Brisbane, Australia

a.bruns / je.burgess @ qut.edu.au – @snurb_dot_info / @jeanburgess – <http://mappingonlinepublics.net/>

Abstract

Twitter has become a major instrument for the rapid dissemination and subsequent debate of news stories. It has been instrumental both in drawing attention to events as they unfolded (such as the emergency landing of a plane in New York’s Hudson River in 2009) and in facilitating a sustained discussion of major stories over timeframes measured in weeks and months (including the continuing saga around *Wikileaks* and Julian Assange), sometimes still keeping stories alive even if mainstream media attention has moved on elsewhere.

More comprehensive methodologies for research into news discussion on *Twitter* – beyond anecdotal or case study approaches – are only now beginning to emerge. This paper presents a large-scale quantitative approach to studying public communication in the Australian Twittersphere, developed as part of a three-year ARC Discovery project¹ that also examines blogs and other social media spaces. The paper will both outline the innovative research tools developed for this work, and present outcomes from an application of these methodologies to recent and present news themes.

Our methodology enables us to identify major themes in *Twitter*’s discussion of these events, trace their development and decline over time, and map the dynamics of the discussion networks formed ad hoc around specific themes (in part with the help of Twitter #hashtags: brief identifiers which mark a tweet as taking part in an established discussion). It is also able to identify links to major news stories and other online resources, and to track their dissemination across the wider Twittersphere.

Introduction

Even though *Twitter* itself is only five years old, and significant questions over its long-term financial sustainability persist, it has become a mainstream and important, if contested, addition to the toolboxes of both journalists and journalism researchers (Ahmad, 2010). Indeed, its trajectory has mirrored that of previous social media platforms. Just as blogs have become established as a platform both for the occasional first-hand reporting of news, and for extensive follow-on curation, commentary, and discussion of news stories (cf. Bruns, 2006), so too *Twitter*’s popular uses have been extended beyond everyday “lifesharing” and interpersonal communication, to a range of similar journalistic, para-journalistic, and quasi-journalistic activities (Kwak *et al.*, 2010; Subasic & Berendt, 2011). Even the accompanying claims of “over-hype” or concerns about *Twitter*’s possible adverse affects on journalistic practice (Farhi, 2009) are in fact a sign of widespread awareness and take-up, and an indication that for journalists and journalism researchers, when it comes to *Twitter*, there is something at stake. This paper addresses the challenges and opportunities which arise for researchers as they endeavour to study how *Twitter* is used in the context of specific journalistic

¹ This paper draws on the project *New Media and Public Communication: Mapping Australian User-Created Content in Online Social Networks*, funded from 2010-2013 through the Australian Research Council’s Discovery scheme. The project Website is at <http://mappingonlinepublics.net/>.

activities. First, however, it is useful to catalogue in broad terms the possible uses of *Twitter* in relation to specific news events which they may encounter.

First and most obviously, perhaps, *Twitter* is used for the first-hand reporting of events as they occur; indeed, the simple format of *Twitter* messages and the near-ubiquitous accessibility of the *Twitter* network (which provides tweet-via-SMS functionality as an ultimate fallback) have combined to make live tweeting a more important practice on *Twitter* than comparable live activities have been for previous social media platforms, including blogs. Further, such live tweeting activities now include not only the reporting of events by actual eyewitnesses on the ground, from the much-publicised live coverage of the emergency water landing of a plane on New York's Hudson River (Subasic & Berendt, 2011) to the inadvertent (and only retrospectively discovered) live account of the raid on Osama bin Laden's compound by an Abbottabad local (Hill, 2011), but also the second-hand live discussion of unfolding events as they are covered by other media – such as the worldwide public sharing of news on the Japanese earthquake and tsunami in March 2011, which generated some 140,000 tweets per hour containing the word 'tsunami' in the hours immediately following the earthquake (see Bruns, 2011a) or the appearance of Rupert and James Murdoch before the British Parliament's Culture, Media, and Sports Committee in July 2011 (Richards *et al.*, 2011).

An additional driver of such tweeting activity is the ease with which – using helper Websites whose functionality is now often embedded in *Twitter* clients – materials well beyond *Twitter*'s 140 character limit (links, photos, video, audio) can also be shared. This aids both the dissemination of first-hand, user-generated material documenting unfolding events as directly experienced by the user, as well as the sharing of existing material in the form of links, screen captures, or even photos of TV screens. These additions have extended *Twitter*'s affordances far beyond the 140 character limit, adding a rich media layer to the tweets themselves – and these multimedia materials now often also make their way into the mainstream media coverage of these events.

Second, and consequently, *Twitter* is also used widely for ongoing discussion – indeed, a kind of instant evaluation – of newsworthy events. Such discussion may be triggered variously by mainstream media reporting, first-hand coverage by *Twitter* users themselves, or the established interests of specific communities of *Twitter* users; especially for the core of these communities of interest, it may often employ processes similar to what Bruns (2005) has described as "gatewatching": highlighting, sharing, and evaluating the relevant material released by other sources, preferably online, in an effort to develop a more comprehensive understanding of the event or issue currently under discussion. What is being shared here, and how it is framed in discussion, provides an interesting alternative to conventional 'vox pop' statements; 'what *Twitter* thinks' about an issue has become a staple feature of mainstream media coverage, therefore.

It is important to understand the structure of *Twitter*'s social network(s), and the extent to which its functionality affords its users the ability to form both *ad hoc* and persistent communicative communities – or, where such communities are directly related to specific themes and topics, issue publics; we will return to this question of network structure in more detail later. One important aspect of news discussion practices on *Twitter* is its use for *curating* information related to specific news stories: especially some of the more active *Twitter* users frequently engage in gathering and sharing what they perceive to be relevant materials, for example by tweeting links to further information (or retweeting relevant posts of other users) to their own followers or to *Twitter* communities formed around topical hashtags.

Finally, beyond such discussion informed by curated evidence, *Twitter*'s coverage of newsworthy events also consists of significant amounts of broader commentary on current events, reflecting mainly the tweeting user's own perspective and intended more as a marker of that perspective than as a formal contribution to debate. It is here, perhaps, that the "ambient" function of *Twitter*, which Hermida (2010) and Burns (2010) have both highlighted, emerges most clearly: for the average user, the majority of their *Twitter* use on any given day may be taken up with non-news-related communication – much of it mundane and phatic – with *Twitter* thus serving as an ambient, always-on, always-in-the-background medium akin to ambient background music (Crawford, 2009). However, when important news breaks, discussion of such news spreads across the Twittersphere, passed along from user to user at a speed which reflects each user's level of interest in the

story – and this shift in the tone and topical focus of tweets coming in from a user’s network of *Twitter* connections may then cause that user to pay attention to the story (again, much as a sudden shift in musical style may cause background music no longer to be perceived as merely ambient).

This shift from ambient to central is measured, in a simple way, by *Twitter*’s own table of ‘trending topics’ (provided worldwide, as well as for specific geographic regions), and appears to be used by a growing number of newsrooms as an additional measure of the newsworthiness of specific stories and events; more detailed analysis not only of keyword and hashtag trends, but also of the relative attention paid to particular sources, political and other actors, or indeed to the information shared by different *Twitter* users, is also possible, however. A recurring feature of the *Guardian*’s coverage of the *News of the World* voicemail hacking scandal, for example, have been visualisations of how *Twitter* users’ mentions of key figures and organisations involved in the scandal have unfolded over time (see e.g. Richards *et al.*, 2011). In this context, especially, journalism practitioners and journalism researchers share a common interest: the development of more sophisticated methods and metrics for describing, analysing and representing *Twitter*’s response to news events.

Twitter as Social Networking Site and Information Stream: Followers and Hashtags

Twitter is the most prominent example of a recent shift in social media: one which has seen the convergence of explicit social networking practices (‘friending’, ‘following’ and interpersonal communication) with original content (the ‘broadcasting’ of updates), and large-scale information sharing and propagation. Crucially, it is *through* the social network that news and information spreads: *Twitter* is both a social networking site and an ambient information stream. This convergence between social network and news propagation (cf. Kwak *et al.*, 2010) is what underpins *Twitter*’s unique significance for journalism; and any serious evaluation of user activities on *Twitter* must reflect on the structural aspects of this convergence. Indeed, explicit networking structures among users determine – to an important extent, though not exclusively – which tweets will be visible to what subset of the total userbase (which has recently been officially estimated at over 200 million worldwide; cf. White, 2011). In the main, *Twitter*’s communicative structure is determined by two overlapping and interdependent networks – one long-term and relatively stable, and based on follower-followee relationships; and one relatively short-term, mercurial and emergent, and based on shared interest in a topic or event, often coordinated by a common hashtag.

First, and most obviously, *Twitter* users are able to ‘follow’ one another, which means that all tweets originating from the followee will automatically be visible to the follower, in an update feed which combines the tweets from all followed users.² *Twitter* follower/followee relationships will usually be based on a more long-term interest in the updates originating from the followee, therefore; studies examining the *Twitter* follower network at large scale have indicated that the structure of the overall network shows unsurprising tendencies to cluster around key interests or attributes shared by communities of followers (Java *et al.*, 2007), much as has been observed for other social networks including the blogosphere (e.g. Adamic & Glance, 2005; Kelly & Etling, 2008). Our own, continuing investigation of follower networks in the Australian Twittersphere shows that even an as yet incompletely mapped network of some 440,000 Australian *Twitter* users already tends towards clustering around shared professional and cultural interests (such as political discussion, specific sports, or professions) as well as shared geography (Bruns, 2011c).

To the extent that *Twitter* users consciously understand this network structure, their responses to newsworthy events must also be understood as attempts variously to address and interact with their own immediate community of followers, and followers of followers, and to overcome the barriers dividing specific clusters in the network from one another. It is likely, in the first place, that the imagined immediate audience

² Generally, it is possible for a user to follow any other *Twitter* user; however, users may set their accounts to ‘private’, in which case they will have to approve any follow requests.

for any one tweet is the user's network of followers, and that receiving responses from, or being retweeted by, these followers is seen as possible and perhaps even desired. Indeed, some tweets carry explicit encouragements to 'pls RT', pointing to the original author's desire for their message to be spread well beyond their own network of followers.

Similarly, then, while retweeting practices vary widely in meaning and intention (boyd *et al.*, 2010), retweeting an existing message reflects an implicit understanding of *Twitter's* network structure: a recognition that unaided, the original message will reach only a limited number of users, and that further passing-along will amplify its visibility. Users retweeting messages may even see themselves to some extent as a kind of information broker, bridging two distinct communities of interest on *Twitter* by passing on tweets from one network cluster to the next. Retweeting is also likely to be interpreted as a form of endorsement for both the message and the originating user, unless additional commentary is added by the retweeter during the retweeting process; especially where the original message stems from a prominent *Twitter* user who may already be safely assumed to be widely visible across the overall Twittersphere, this latter motivation may well outweigh the desire to amplify and disseminate the original tweet.

Twitter follower/followee structures are by no means static: to add another user to one's list of connections is a matter of a few clicks. Nonetheless, it is unlikely that such structures will fluctuate wildly for any one user: substantial changes to one's list of followees would also significantly impact on the incoming stream of updates which a user will receive. Analogous to some extent to blogrolls, *Twitter* follower/followee networks represent the long-term interests, rather than the short-term foci, of individual *Twitter* users.

The network structures described by follower/followee relationships, however, are also overlaid by another mechanism for coordinating communication using the *Twitter* platform: hashtags. Hashtags (an innovation originally introduced by the *Twitter* user community) consist of brief keywords or abbreviations, prefixed by the hash symbol, which are included in tweets – for example, #auspol for a discussion of Australian politics, or #NotW for exchanges about the *News of the World* scandal. Hashtags are simply entered manually by users as they compose their messages; so they can be created *ad hoc*, and often emerge almost instantaneously as news breaks. Hashtags make topical tweets more visible: drawing simply on *Twitter's* search functionality, users are able to find (and even subscribe to) all the tweets marked with the same hashtag – regardless of whether these tweets originate from their established followees or from previously unknown users.³

For *Twitter* users, following and posting to a hashtag conversation makes it possible for them to communicate with an *ad hoc* community of interest around the hashtag topic without needing to go through the process of establishing a mutual follower/followee relationship with all or any of the other participants. Arguably, indeed, it is this flexibility and ability to rapidly form discursive communities around breaking news which has contributed most notably to *Twitter's* recognition as a platform for news dissemination and discussion. At the same time, a drawback of the *ad hoc* and non-supervised emergence of hashtags is that competing hashtags may emerge in different regions of the Twittersphere (for example, #eqnz as well as #nzeq for coverage of the Christchurch earthquakes in 2010 and 2011), or that the same hashtag may be used for vastly different events taking place simultaneously (for instance, #spill for the BP oil spill in the Gulf of Mexico in 2010, as well as for the leadership change in the Australian Labor Party on 23 June 2010). *Twitter* users themselves will often work to resolve such conflicts quickly as soon as they have been identified.

Using a hashtag in one's tweet can be seen as an explicit attempt to address an imagined community of users following and discussing a specific topic, therefore – and the network of *Twitter* users which is formed from this shared communicative practice must be understood as separate from follower/followee networks. At the same time, the two network layers overlap: tweets marked with a specific hashtag will be visible *both* to the user's established followers, *and* to anyone else following the hashtag conversation. Users from the follower network who respond and themselves include the hashtag in their tweets thereby also become part of the hashtag community, if only temporarily, while responses to or retweets of material from the hashtag

³ One remaining limitation is that hashtagged tweets from accounts marked as 'private' will not be included in the search results.

conversation are also visible to the follower network (similarly, some users may retweet topical tweets from their followers while adding a hashtag in the process, thereby making those tweets visible to the hashtag community as well). Each user participating in a hashtag conversation therefore has the potential to act as a bridge between the hashtag community and their own follower network.

At the same time, not all users posting *to* a hashtag conversation also *follow* that conversation itself: they may include a topical hashtag to make their tweets visible to others following the hashtag, thereby increasing its potential exposure, but may themselves continue to focus only on tweets coming in from their established network of followers (this is especially likely for very high-volume hashtagged discussions). Conversely, not all relevant conversations following on from hashtagged tweets will themselves carry the hashtag: to hashtag a response to a previous hashtagged tweet, in fact, may be seen as performing the conversation in front of a wider audience, by comparison with the more limited visibility which a non-hashtagged response would have. It should be noted that hashtags are also used in other contexts: for example as an ironic marker of one's mood – #tired – or to highlight key terms – #Australia –, but our focus in the following discussion is purely on *topical* hashtags.

Finally, a unique communicative feature of the *Twitter* platform should also be noted: excepting 'private' accounts, tweets are generally visible also to non-users. Such users may use the *Twitter* Website to follow the tweets made from specific accounts, or use *Twitter's* search functionality to find all tweets containing particular keywords or hashtags; *Twitter* search will also alert them automatically as further tweets containing their search terms are made. This further adds to the value especially of hashtags as a mechanism for coordinating news discussion and information curation.

Methods for Researching *Twitter*

Both for journalism scholars and indeed for journalism practitioners themselves, there is a clear need to develop further methods for researching *Twitter's* response to news and current events. Like other modern social media platforms, *Twitter* provides significant levels of access to data on user activities through its Application Programming Interface (API): an interface which is designed predominantly for use by *Twitter* clients, but which can also be used for tracking current activity by particular users or groups of users, or on specific keywords and hashtags.

That said, it must also be noted that there are substantial limitations to what is available directly through the API, and that these limitations are gradually tightened as *Twitter* seeks sources of income for its service. Changes made late in 2010, for example (Melanson, 2011), mean that even for the purposes of publicly funded, non-commercial research, it is no longer possible to gain access to the full 'firehose' of *all* tweets, or even of a substantial subset of this full feed; indeed, tracking the ongoing public activities of more than a maximum of 5,000 identified *Twitter* users is now only possible by working with *Twitter's* licenced third-party API provider Gnip, at commercial access costs which are likely to be well beyond the funding available to the vast majority of research institutions. In practice, this rules out any academic studies that aim to track the current activities and thematic interests of large, representative samples of *Twitter* users.⁴

Similarly, due to access speed restrictions set by *Twitter*, even to establish a full picture of follower/followee networks on *Twitter* is a time-consuming exercise: while (except for 'private' accounts) information on a user's connections is readily available, the retrieval of that information from *Twitter* is artificially slowed down to limit stress on its data servers, and accessing such data for several hundreds of thousands of users may take months. Our current project (see Bruns & Burgess, 2011) is attempting to establish a snapshot of the structure of the Australian Twittersphere, whose size we estimate at between 500,000 and two million users; we hope to complete this mapping exercise within six months.

⁴ *Twitter* has entered into an agreement with the U.S. Library of Congress to make its full 'firehose' feed available to the Library, to be accessed by researchers after a six-month delay period (Raymond, 2010); however, the six-month embargo means that research into *current* coverage and news discussion on *Twitter* remains impossible, and the modalities of access to the *Twitter* dataset at the Library are as yet unknown.

By contrast, a more immediately achievable direction for *Twitter* research focussing on news and current events is to track the *Twitter* community's use of keywords and/or hashtags. The *Twitter* API makes it possible to automatically capture any tweets containing given keywords (including hashtags), with comparatively few limitations, and a number of tools for doing so are now readily available to interested researchers; of these, the leading open-source solution to date is the *yourTwapperkeeper* (*yTK*) platform.⁵ With minor modifications to improve its data export functionality (see Bruns, 2011b), *yTK* can be used to track a substantial number of keywords simultaneously, and further processing and evaluation of the datasets generated in this way can reveal a number of important patterns in the data.

Such processing draws on a number of additional tools: the command-line data processing tool Gawk, statistics packages such as Excel or SPSS for data analysis, textual processing software such as Leximancer or WordStat for keyword analysis, and the open-source package Gephi for network analysis and visualisation. This paper can necessarily only provide an overview over the tools and methods which may be brought to bear on these datasets; it does not intend to act as a hands-on methodological guide. However, our project Website at <http://mappingonlinepublics.net/> offers substantial practical advice on processing *yTK* data, as well as providing access to the custom-made open source software tools we have developed for this purpose.

In the following, we offer an overview of the insights that this data-driven approach to the study of *Twitter* interactions can provide. It is important to note in this context that a keyword- or hashtag-based study of *Twitter* activity is necessarily limited, and will not capture the parts of the discussions of any particular news event which do not include the chosen keywords or hashtags in their tweets. Therefore, the data captured following these methods must be understood as a more or less representative sample rather than a comprehensive dataset of *Twitter* activities around the event – indeed, hashtag datasets in particular may be weighted considerably towards the most engaged subset of *Twitter* users (that is, those users engaged enough to include the hashtag in their tweets), and thus towards a comparatively élite group. This does not invalidate studies undertaken using these methods, however, any more than a study of the tenor of letters to the editor in a leading newspaper would be invalidated by that newspaper's specific audience composition. Additionally, the overall make-up of the *Twitter* userbase, and its lack of correlation with the overall demographic patterns of most countries, means that the views of *Twitter* users can never simply be regarded as representative for the general population.

Development over Time

Amongst the simplest forms of analysis which are possible for a *yTK* dataset is the study of activity patterns over time. This may simply chart the number of tweets made during any one time period, or also break them down into a number of distinct categories. Fig. 1, for example, depicts the daily number of tweets made to the #ausvotes hashtag in the lead-up to and immediate aftermath of the 2010 Australian federal election, showing a gradual ramping-up of activity and a major spike on election day itself; another minor early spike occurs on 25 July, the day of the single televised debate between the two party leaders.

By contrast, Fig. 2 shows minute-by-minute *Twitter* activity under the #royalwedding hashtag, following the wedding of Prince William and Kate Middleton on 29 April 2011; here, activity is broken

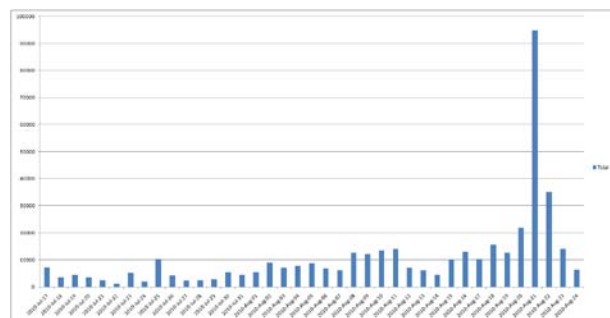


Figure 1: #ausvotes tweets, 17 July - 24 Aug. 2010

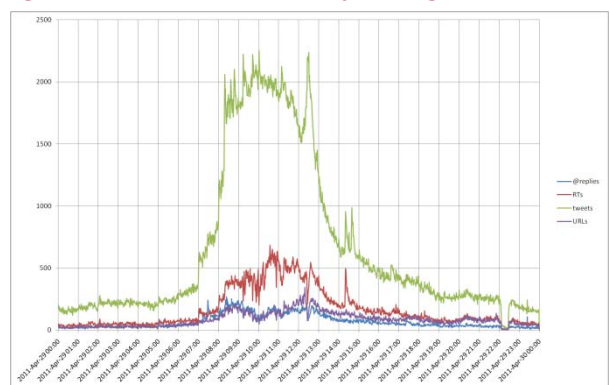


Figure 2: #royalwedding tweets, 29 Apr. 2011

⁵ See <http://your.twapperkeeper.com>

down into the overall number of tweets (shown in green), retweets (red), @replies (blue), and tweets containing URLs (purple). The graph clearly points to the nature of the wedding as a televised event, with significant increases in volume at 7:00, 8:00, and 8:15 GMT, as major networks in the U.K. and elsewhere switched to their live coverage; it also enables researchers to identify key moments of the day, as *Twitter* activity spikes around 12:25, for example, during the newlyweds' first kiss (while retweet, @reply, and URL levels simultaneously drop, indicating a heightened level of original tweets).⁶

Although they are relatively simplistic, these metrics nonetheless provide us with clear insights into what moments in an unfolding story appeared to have particular resonance for the participating *Twitter* community; they may be usefully correlated with key developments as they occurred. Inferences about *Twitter* users' media consumption practices may also be drawn: both figs. 1 and 2 point to the use of *Twitter* as a backchannel for television, for example (key moments in #ausvotes, too, correlate with major televised events, from the leaders' debate to the televised tally room vote count), and to the continuing resonance of well-established mainstream media tropes in the coverage of the respective events (debates and TV appearances, vows and kisses) even in social media environments.

Key Users

In addition to establishing these basic activity statistics, it will usually be useful to identify the key participants in the topical discussion. Given the obvious ethical concerns with highlighting the activities of individual users, the goal here should not be to engage in detailed profiling of participants, however, but rather in establishing the overall structure of the community (to the extent that a 'community' in the full sense of that term actually exists around a specific keyword or hashtag). In doing so, it will often be less important to examine the total number of tweets *sent* by each user (since the volume of contributions made in itself does not provide a reliable approximation of the impact of those tweets on the continuing discussion), but rather to focus on the number of responses (that is, public @replies) and retweets received by participating users.⁷

It should also be expected that the majority of messages received by an account in a hashtag conversation will be retweets, rather than @replies: as noted above, direct @reply responses to a hashtagged tweet will not usually include the hashtag, unless the respondent is deliberately aiming to make their response visible to the wider hashtag community as well. Much of the follow-on conversation around the hashtagged topic will therefore be absent from the dataset.

In spite of these limitations, an analysis of responses and manual retweets received provides a useful indication of the overall visibility of each account: as discussed above, retweets themselves are a means of amplifying the reach of a tweet, and thus of increasing the visibility of a tweet and its sender (and there is no obvious reason to believe that there should be any significant discrepancy between the number of manual and the number of 'button' retweets, meaning that manual retweets provide a proportional representation of overall retweeting patterns), while the volume of @replies received by a user can be seen as a direct result of the visibility of their tweets, and thus of themselves (*Twitter* users can only @reply to messages and users they are aware of in the first place). In combination, then, these two measures are a sign of a user's visibility.

⁶ Higher-resolution versions of these and most other figures presented in this paper are available online – see the Index of Figures below for URLs.

⁷ It should be noted that *yourTwapperkeeper* only captures *manual* retweets ("RT @user ..."), not retweets made using *Twitter*'s more recently introduced 'retweet button' functionality. This, however, can also be seen as a benefit, as – because they can be edited before sending – many manual retweets serve a significantly more conversational function than 'button' retweets; for example, users will often retweet part of an earlier message in order to add their own, original commentary. 'Button' retweets, on the other hand, constitute merely a verbatim passing-along of the original message, but do not enable retweeting users to include any additional comments with the retweeted message. While a tracking of the amount of button retweets for each individual message captured in the dataset might provide an interesting additional dimension to the analysis, then, it does not have significant relevance to the analysis of actively *discursive* interaction.

day. What emerges here is both an overall bias in the *Twitter* conversation about the election as it took place on #auspol, in comparison with mainstream media coverage, towards themes related to information and communication technology, as well as a shift in discursive focus after 10 August, from the mandatory Internet filter proposed by the Labor government (in red) to the National Broadband Network project opposed by its conservative challengers, and criticised by Abbott in his television interview (in blue).

Advanced Network Analysis

Finally, in addition to statistical and textual analysis of *Twitter* activity patterns, network analysis provides further possibilities. An immediately obvious approach in this context is the analysis of the (temporary) social networks which are evident in response and retweeting patterns; such analysis adds further insight to the statistical identification of the most visible accounts which we have discussed above. Beyond merely counting retweets and @replies, social network analysis adds a number of further possible metrics, which can highlight not simply those accounts which are most active in sending and receiving tweets, but also point to those participants, for example, with the greatest ‘betweenness’ – that is, who have the greatest importance as connectors of otherwise distant parts of the network: such users may not be major net sources or recipients of information, but are highly instrumental in ensuring the movement of information *across* the network.

Fig. 7, for example, shows the network of retweets and @replies in the #spill hashtag community discussing rumours of a leadership challenge in the Australian Labor Party during the evening of 23 June 2010. While not actively participating, the account of the then Prime Minister, @KevinRuddPM, is featured prominently as an addressee of @replies, while some journalists covering the unfolding event are prominent both as senders and receivers of @replies and retweets. In addition to such static network maps, indeed, it is also possible to visualise their gradual development over time in a series of snapshots or as a dynamic animation of the available social network data (see Bruns, 2010a/b).

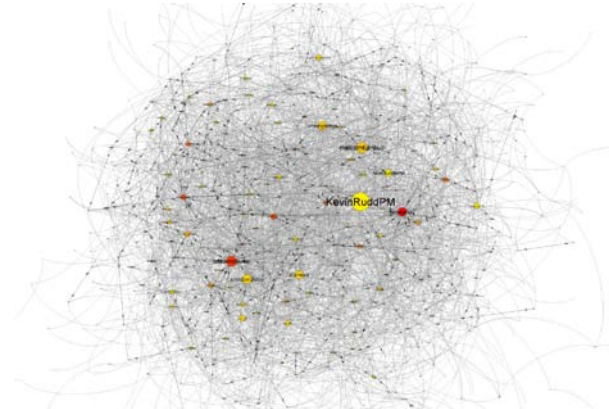


Figure 7: #spill @reply network, 23 June 2010
Node size = indegree; node colour = outdegree

Beyond social network analysis, further opportunities exist for the visualisation of hybrid networks, linking participating users to the key concepts they reference, or showing the relative closeness or distance between different terms and actors in the textual corpus. An overview of the complete range of possibilities in this space is well beyond the scope of this paper, but as one example, Fig. 8 depicts a network extracted from tweets discussing the SBS television miniseries *Go Back to Where You Came From*, which utilised the reality TV format to tackle Australia’s persistent debate about asylum seekers. The show sent six Australians with outspoken views about ‘illegal’ immigration on a reverse voyage from Australia to a number of the main countries of origin of asylum seekers, and generated substantial discussion on *Twitter* under the #GoBackSBS hashtag.

Fig. 8 identifies mentions of the six main ‘contestants’ of the series, and connects *Twitter* users to those contestants whom they mention most often. Users mentioning only one of the contestants are shown in red, while users discussing a high number of contestants are shown in green; the overall volume of mentions is also indicated by the size of each user’s node. The contestants themselves are shown in dark red. What emerges from this is

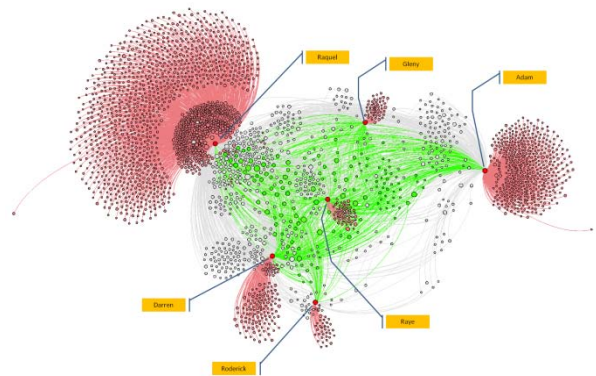


Figure 8: Mentions of #GoBackSBS ‘contestants’

that two contestants, Raquel and Adam, attract a particularly high number of one-off comments, which may indicate them to be especially rich characters and/or polarising personalities, while other contestants are either hardly mentioned at all, or discussed mainly by *Twitter* users who did venture beyond one-off comments on any one specific contestant.

Conclusion

The methods and research possibilities outlined here are well within the grasp of journalism researchers, and provide clear and detailed insights into how the *Twitter* community responds to and engages in newsworthy events. Clearly, such responses should not be seen as simply representative for society at large, any more than letters to the editor or vox-pop statements from random passers-by are representative for the wider population. But that is not the point: rather, how the *Twitter* userbase reacts to specific events is relevant in its own right, and an equally valid indication of popular sentiment as these other non-representative indicators.

These insights are of direct value to journalism practitioners; additionally, especially where it is possible to use our methods to examine the interrelation between *Twitter* use and other forms of (mainstream) media participation and consumption, there are significant further benefits for journalism scholars: compared to studies which build on such relatively artificial devices as media usage surveys or diaries, for example, tracking *Twitter* activities provides a considerably more *in situ* picture of media engagement, comparatively unaffected by the very act of observation itself.

Additionally, many of the analyses outlined here can be performed virtually in real time, as data are gathered; this enables researchers to respond very quickly to current events, and to examine the popular response on *Twitter* even while these events are still unfolding. While detailed analysis may still require considerable mental effort, of course, delays associated with the establishment of a substantial dataset in the first place, at least, are no longer relevant.

This kind of research does rely on the continued availability of relevant data through the *Twitter* API, of course – and as researchers, we are treated as no more than uninvited guests by *Twitter* at this point, with no guarantees of continued access, and a gradually increasing push by *Twitter* towards a commercialisation of large-scale data access as part of a broader trend in online business models toward the direct monetisation of social media data (Steele, 2011). It would be most unfortunate for this trend to continue, since commercialisation would effectively rule out much publicly funded research into *Twitter* use – but it is precisely these scholarly studies which clearly document the important role which *Twitter*-based communication plays in the coverage and discussion of news and current events as they unfold.

Even if *Twitter* were to become an increasingly hostile environment for researchers, however, the time spent developing *Twitter* research methods remains time well spent. It is inconceivable for any future social media platform not to offer an API comparable to that of *Twitter*, and while technologies may change, the methods which we have outlined here will be transferable – *mutatis mutandis* – to the study of other online social networks as well.

Index of Figures

URLs for higher-resolution versions of the figures included in this paper:

- 1: <http://www.mappingonlinepublics.net/2010/11/24/election-2010-the-view-from-twitter/>
- 2: unpublished
- 3: <http://www.mappingonlinepublics.net/2011/07/12/emergency-management-conference/>
- 4: <http://www.mappingonlinepublics.net/2011/07/12/emergency-management-conference/>
- 5: <http://www.mappingonlinepublics.net/2010/11/24/election-2010-the-view-from-twitter/>
- 6: <http://www.mappingonlinepublics.net/2010/11/24/election-2010-the-view-from-twitter/>
- 7: <http://www.mappingonlinepublics.net/2010/11/24/election-2010-the-view-from-twitter/>
- 8: <http://www.mappingonlinepublics.net/2011/07/07/tweeting-at-the-tv-some-observations-on-gobacksbs/>

References

- Adamic, L. A., and N. Glance (2005) "The Political Blogosphere and the 2004 U.S. Election: Divided They Blog." 2nd Annual Workshop on the Weblogging Ecosystem: Aggregation, Analysis and Dynamics, Chiba, Japan, 10 May 2005. <http://www.blogpulse.com/papers/2005/AdamicGlanceBlogWWW.pdf>, accessed 1 Aug. 2011.
- Ahmad, Ali Nabil (2010) "Is Twitter a Useful Tool for Journalists?" *Journal of Media Practice*, 11(2), pp. 145-55.
- boyd, danah, Scott Golder and Gilad Lotan (2010) "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter." Proceedings of 43rd Hawaii International Conference on System Sciences, IEEE. <http://doi.ieeecomputersociety.org/10.1109/HICSS.2010.412>, accessed 1 Aug. 2011.
- Bruns, Axel (2005) *Gatewatching: Collaborative Online News Production*. New York: Peter Lang.
- (2006) "The Practice of News Blogging." In Axel Bruns and Joanne Jacobs (eds.), *Uses of Blogs*. New York: Peter Lang, pp. 11-22.
- . (2010a, 30 Dec.) "Visualising Twitter Dynamics in Gephi, part 1." *Mapping Online Publics*, <http://mappingonlinepublics.net/2010/12/30/visualising-twitter-dynamics-in-gephi-part-1/>, accessed 1 Aug. 2011.
- . (2010b, 30 Dec.) "Visualising Twitter Dynamics in Gephi, part 2." *Mapping Online Publics*, <http://mappingonlinepublics.net/2010/12/30/visualising-twitter-dynamics-in-gephi-part-2/>, accessed 1 Aug. 2011.
- . (2011a, 10 Feb.) "Twitter Events in Perspective (updated)." *Mapping Online Publics*, <http://www.mappingonlinepublics.net/2011/02/10/twitter-events-in-perspective/>, accessed 1 Aug. 2011.
- . (2011b, 21 June) "Switching from Twapperkeeper to yourTwapperkeeper." *Mapping Online Publics*, <http://www.mappingonlinepublics.net/2011/06/21/switching-from-twapperkeeper-to-yourtwrapperkeeper/>, accessed 1 Aug. 2011.
- . (2011c, 4 Aug.) "A First Map of Australia." *Mapping Online Publics*, <http://www.mappingonlinepublics.net/2011/08/04/a-first-map-of-australia/>, accessed 1 Aug. 2011.
- Bruns, Axel, and Jean Burgess (2011) *Mapping Online Publics*, <http://mappingonlinepublics.net/>, accessed 1 Aug. 2011.
- Burns, Alex (2010, May) "Oblique Strategies for Ambient Journalism." *M/C Journal* 13(2). <http://journal.media-culture.org.au/index.php/mcjournal/article/view/230>, accessed 1 Aug. 2011.
- Crawford, Kate (2009) "Following You: Disciplines of Listening in Social Media." *Continuum: Journal of Media & Cultural Studies* 23(40): pp. 525-35.
- Hermida, Alfred (2010, May). "From TV to Twitter: How Ambient News Became Ambient Journalism." *M/C Journal* 13(2). <http://journal.media-culture.org.au/index.php/mcjournal/article/view/220>, accessed 1 Aug. 2011.
- Hill, Amelia (2011) "Osama bin Laden Killed - How a Live Blogger Captured the Raid." *The Guardian*. <http://www.guardian.co.uk/world/2011/may/02/osama-bin-laden-live-blogger>, accessed 1 Aug. 2011.
- Farhi, Paul (2009) "The Twitter Explosion." *American Journalism Review* April/May. http://ajr.org/article_printable.asp?id=4756, accessed 1 Aug. 2011.
- Java, Akshay, et al. (2007) "Why We Twitter: Understanding Microblogging Usage and Communities." Paper presented at *WebKDD/SNA-KDD '07*, New York. <http://portal.acm.org/citation.cfm?id=1348556>, accessed 1 Aug. 2011.
- Kelly, J., and B. Etling (2008) "Mapping Iran's Online Public: Politics and Culture in the Persian Blogosphere." Boston: Berkman Center for Internet & Society. http://cyber.law.harvard.edu/publications/2008/Mapping_Irans_Online_Public, accessed 1 Aug. 2011.

- Kwak, Haewoon, *et al.* (2010) "What Is Twitter, a Social Network or a News Media." *WWW'10: 19th International Conference on World Wide Web*. <http://portal.acm.org/citation.cfm?id=1772751>, accessed 1 Aug. 2011.
- Melanson, Mike. (2011, 11 Feb.) "Twitter Kills the API Whitelist: What It Means for Developers & Innovation." *ReadWriteWeb*. http://www.readwriteweb.com/archives/twitter_kills_the_api_whitelist_what_it_means_for.php, accessed 1 Aug. 2011.
- Raymond, M. (2010, 25 Apr.) "How Tweet It Is!: Library Acquires Entire Twitter Archive. *Library of Congress Blog*. <http://blogs.loc.gov/loc/2010/04/how-tweet-it-is-library-acquires-entire-twitter-archive/>, accessed 1 Aug. 2011.
- Richards, Jonathan, *et al.* (2011, 13 July) "How Twitter Tracked the News of the World Scandal." *The Guardian*. <http://www.guardian.co.uk/media/interactive/2011/jul/13/news-of-the-world-phone-hacking-twitter>, accessed 1 Aug. 2011.
- Steele, Julie (2011) "Data Markets Aren't Coming, They're Already Here." *O'Reilly Radar*. <http://radar.oreilly.com/2011/01/data-markets-resellers-gnip.html>, accessed 1 Aug. 2011.
- Subasic, Ilija, and Bettina Berendt (2011) "Peddling or Creating? Investigating the Role of Twitter in News Reporting." In P. Clough, C. Foley, C. Gurrin, G. Jones, W. Kraaij, H. Lee, and V. Mudoch (eds.), *Advances in Information Retrieval*. Berlin: Springer, pp. 207-13.
- White, Charlie (2011) "Reaching 200 Million Accounts: Twitter's Explosive Growth." *Mashable*. <http://mashable.com/2011/07/16/twitter-accounts-200-million/>, accessed 1 Aug. 2011.