

This article is downloaded from



CHARLES STURT
UNIVERSITY

CRO

CSU Research Output
Showcasing CSU Research

<http://researchoutput.csu.edu.au>

It is the paper published as:

Author(s): Albright, K.H. ; Henninger, M. ; Scifleet, P.A.

Title: When social media are your source

Journal: Information Research: an international electronic journal

ISSN: 1368-1613

Year: 2013

Pages: 1 - 14

Volume: 18

Issue: 3

Abstract: Introduction. This study investigates the challenges that arise when social media are acquired as an information resource to be maintained, studied, re-purposed and re-used. The aim of the research is to understand how the information architecture of the messages contributes to social interactions; how the record of an event might be collected and understood; and how knowledge of architecture can contribute to other social studies. In doing so, the research contributes to an understanding of the ...

URLs:

FT: <http://www.informationr.net/ir/18-3/colis/paperC41.html#.U0t9ecfBEZw>

PL: http://researchoutput.csu.edu.au/R/-?func=dbin-jump-full&object_id=55954&local_base=GEN01-CSU01

When social media are your source

Introduction. *This study investigates the challenges that arise when social media is acquired as an information resource to be maintained, studied, re-purposed and re-used. The aim of the research is to make sense of the information architecture of social messages to understand, how it contributes to social interactions; how the record of an event might be collected and understood; and how knowledge of architecture can contribute to other social studies. In doing so, the research contributes to an understanding of the custodianship of an increasingly important cultural document.*

Method. *The research undertakes a qualitative media analysis of real time social media feeds documenting a significant national event, the flooding of the Riverina in Australia during March 2012. By combining a qualitative content analysis of communications with analysis of the materiality and form, the investigation presents an integrated approach for understanding the social dimensions of information architecture.*

Analysis & findings. *During the flood, communications were sampled from Facebook, Google Plus & Twitter activity streams encapsulated in standardised metadata, allowing the same questions to be asked across each stream. The architecture of messages, communicative intent of authors and topics of conversation are analysed to report on the organisation of knowledge in social media documents.*

Conclusion. *The research departs from the direction of other studies in social media and crises management, with their attention on improving communications, to investigate the social arc of communication between people, raising new and important directions for library and information science.*

Keywords. *Documentation, information architecture, knowledge organisation, social media*

1 Introduction

Social media has changed the information landscape across all environments and communities of practice. Information professionals, particularly within library practice, face many challenges and issues when social media becomes a source to be acquired, managed and re-used. This paper reports on a study undertaken by the University of Technology, Sydney and Charles Sturt University, Australia, to investigate how information sourced from social media is being incorporated into scholarship.

The large amount of research that is taking place around the world by collecting and analysing topical sets of social media is a clear indicator of the global significance of the evidentiary role of social media; with studies analysing human behaviour, election campaigns, natural disasters and events like the Arab Spring of 2011, it is hard to argue otherwise. The announcement from the U.S. Library of Congress in August 2010 that it was to become the custodial institution for the Twitter archive underlines this point (Eversley, 2011).

For researchers, working with the massive flows of information presented in social media activity streams is challenging: What kinds of questions can be answered? Which social media feeds are suitable and how do they inform? What tools can be used to collect, manage and work with social media? What types of reporting are valuable?

This study is motivated to understand how information practitioners can participate in making sense of these sources of information. Our paper reports on a pilot study undertaken between March 2012 and March 2013, to investigate the difficulties faced in discovering, collecting and analysing social media messages. The perspective we bring to the research is that of documentary practice. We ask what issues, challenges and assumptions about the richness, validity and reliability are faced when social media becomes a documentary source to be brought under custodial control.

2 The importance of social media research

With very few boundaries on the kind of prospecting that might be undertaken, social media analysis is often associated with trends in 'big data'. There is no apparent limit to the number of messages that might be acquired or the dimensions of society that might be investigated: "we no longer have to choose between data size and data depth. We can study exact trajectories formed by, billions of cultural expressions, experiences, texts, and links" (Manovich, 2012, pp. 462-463). There are high expectations about what this kind of research might empower social scientists to do (Dutton & Jeffeys, 2010), but without research and development activities designed to set the foundations by discovering what is available in the message, working with the continuous streams of communications is likely to remain challenging and outside the reach of most social scientists.

The foundations for working with public available social media are being set now. With new mediators entering a market for provisioning access to social media directly to end users, questions of custodianship, curation, preservation and access are moot points. In the

prevailing doxa, publicly available social media messages are treated as though they are published online and available for use like any other publication (Neuhaus & Webmoor, 2012). That may be so, however ethical questions of privacy, ownership and use remain as challenges to be addressed. Library and information science must ask if its moral and ethical foundations in fair use, freedom of access, conditions for use, the protection of intellectual property and privacy and custodianship have a role to play.

2.1 Types of social research

Social media scholarship is emerging from disciplinary and methodological approaches building on traditions as diverse as computer science and social studies (Boyd & Ellison, 2007). Much of this research focuses on social media as a communication phenomenon. Both quantitative hypothesis-driven, and qualitative interpretive approaches are widespread and incorporate everything from statistical techniques for data mining and natural language processing to classification for content, genre and sentiment analysis.

Research founded on the analysis of social media messages spans all social fields, including education (Simon, Davis, Griswold, Kelly, & Malani, 2008), economics and business (Riemer & Scifleet, 2012), health sciences (Oh, 2012), linguistics (Zappavigna, 2011), sociology (Boyd, Golder, & Lotan, 2010), media and communication (Papachariss, 2012), and political science (Woolley, Limperos, & Oliver, 2010). In business, it is an important ingredient in monitoring brand health and identifying sales channels (Etlinger & Li, 2011; McGuire, 2012). Journalists are using social media to identify breaking news, present data visualizations, and quote views or opinion in their stories (Boyd & Ellison, 2007; Bruns & Highfield, 2012). Within information science, work -including this study- is taking a sense-making approach (Dervin, (1983) (Heverin & Zach, 2012) to understand what is taking place. However, more work needs to be done to understand how the public document might constitute a corpus of digital information sources.

2.2 The research aim

Collecting social media messages is akin to filling a bucket of water from an ocean, the evidence that results is not *the social media*, it is an artefact purposefully collected as a record of public communication about topics, events and themes. A judgement is made about what is been documented (Frohmann, 2009). Access to that evidence requires an understanding of the information architecture of the document that results. We view information architecture as the elements of a document its broadest sense, encompassing the text content, structure, form and context of a message (Lund, 2010; Pédaque, 2003), and aim to understand *firstly*, how the messages' architecture contributes to the record of social interactions taking place; *secondly*, how that architecture might be utilised in the collection and management of social media as an information resource; *finally*, how the architecture might be applied for analysis and reporting in all manner of social studies. By doing so the paper aims to contribute to an understanding of the custodianship of an increasingly important document.

3 Social media as social document

The view we bring to this study is one of documentary practice as the set of techniques, including processes for the selection, synthesis and interpretation of the material form of documents and their content, meaning and context, that librarianship brings to the organisation and management of knowledge (Briet, Day, Martinet, & Anghelescu, 2006; Pédaque, 2003). Current emphases in social media research on ‘big data’ and quantitative analysis are distracting from the significant role social media has to play as a record of social significance that should be brought into public custody for future use.

In its multiple manifestations, social media is “a new kind of cultural artefact” (Lyman & Kahle, 1998, para 15), as was the World Wide Web when Brewster Kahle set up the Internet Archive, reasoning that “in future it may provide the raw material for a carefully indexed, searchable library” (Kahle, 1997, p. 82). While the issues associated with digital archiving are not easily resolved, libraries and information scientists have not shied away from the challenge. National libraries, including the Library of Congress and the National Library of Australia have been developing processes for the selective archiving of documents from the Internet since the late 1980’s (Pymm & Lloyd, 2007); yet Charnigo & Barnett-Ellis, in their survey of library uses of Facebook, noted that while no administrator or librarian would dare weed a collection [of year books] or find its presence irrelevant, “online social networks are dynamically documenting the here and now of campus life and shaping the future of how we communicate” (Charnigo & Barnett-Ellis, 2007).

It is time for librarians to take up the challenge that interactive social media documents present. There are many questions to be answered, about what the duty to collect and preserve is, what kinds of social media records should be prioritised and, what selection and management criteria should be applied. Document description and knowledge organisation are themselves activities that we use to make sense of complex sets of information because it tells us about them—it seems sensible then that our methodological starting point needs to be an interpretation of what the social media document presents.

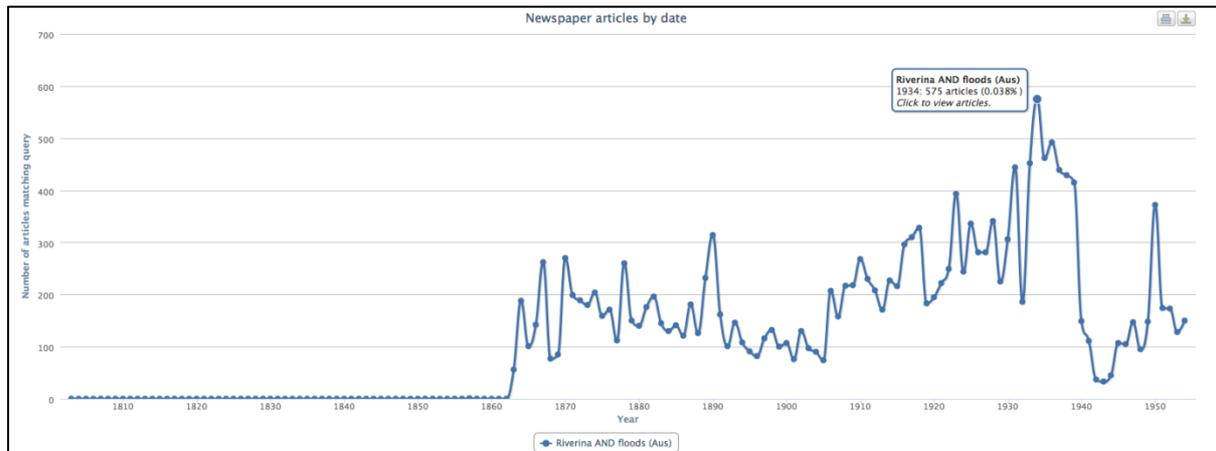
4 Case study: The Riverina flood, March 2012

Australia’s Riverina, in southwestern New South Wales is one of the country’s most important agricultural regions. The Murray and Murrumbidgee rivers that supply water from the Snowy Mountains to the river basin give the district its name. Although the Australian Bureau of Meteorology classifies the region as a hot dry zone with severe droughts yet, as Figure 1 shows flooding is also a feature of life.

During the first week of March 2012, the Riverina catchment experienced a flood that was unprecedented. On 5 March a state of emergency was declared; residents of the region’s largest city, Wagga Wagga (abbreviated to ‘Wagga’), were evacuated following predictions that the Murrumbidgee would breach the 10.7 metre high levee with catastrophic impact. This study tracks the evacuation of Wagga as it was documented in three social media services, Facebook, Google Plus and Twitter, over a period of 48 hours from the time

evacuation orders were given at 9pm on Monday 5 March, 2012 until the city was reoccupied on Wednesday 7 March.

Figure 1 Newspaper articles for Riverina and Floods, 1850-1950.



Source: dHistory workbench, February, 2013: <http://dhistory.org/querypic/65/>

The opportunity social media provides to study the patterns of human communication during crises like this are important; new knowledge is generated through research that pays dividends by mitigating risks and saving lives. Most current studies aim to improve the communication strategies of emergency services during a crisis (Bruns, Burgess, Crawford, & Frances, 2012), by monitoring communications to identify events as they commence (Li et al., 2011), and developing techniques to collect public communication in real time to support management decision-making during a crisis (Cameron, Power, Robinson, & Yin, 2012). The approach to these studies is varied, ranging from the development of statistical models for automatic classification (Cameron, et al., 2012) to counting keyword frequencies and their distribution over time (Li, et al., 2011). Of the three studies, Brun's is closest to our own; in this study researchers used a combination of automated data extraction followed by content analysis to categorise tweets for the Queensland floods and Queensland's Police force following a serious flood in Queensland, Australia in 2011. While our study also presents a qualitative content analysis supported by descriptive statistics, it makes a significant departure from the objectives of other studies. Where preceding studies have focused on analyzing social media with the objective of improving crisis management, our study follows the communication arc of a social document constructed by people caught up in the event. Our intention here is to learn more about the kind of cultural document that social media presents.

5 Research design and analysis

5.1 GNIP enterprise data collector

The study reports on a pilot project undertaken with GNIP to collect and analyse real time social media feeds. Like any number of services now populating the public data marketplace, GNIP is a mediator between primary social media (e.g. Twitter) and those requiring access to social media for secondary purposes, such as data analysis and application development.

That this mediation must affect the document and its interpretation, while not the focus of the pilot study, is important to it. GNIP is not only the largest service provider of its type, providing access to standardized feeds from Facebook, Google Plus, Newsgator, Reddit, Twitter, YouTube, Wordpress and others, it also delivers the Twitter archive to the Library of Congress and therefore has a significant position in any study asking questions about collecting social media.

GNIP operates by configuring data from native social media feeds into standardized JSON or XML based activity streams (<http://activitystrea.ms/specs/json/1.0/>), that are accessed via an application programming interface (API). GNIP's processes do not change the original message nor its structure; they simply check and ensure the descriptive metadata encapsulating the messages is normalized and compliant with the activity stream specification. This results, for example, in a Facebook document that uses the same descriptive model as a Twitter document, despite differences in the service architecture. Access to standardized, readable, messages is important in our consideration of document curation because it allows us to work with comparable documents from multiple sources that combine to form a social document authored by everyday people.

Figure 2 Data streams processes.

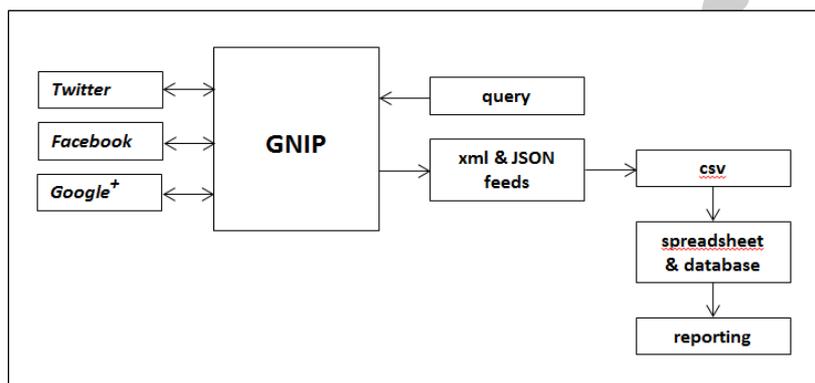


Figure 2 presents the data acquisition process for this study, beginning with the submission of our search terms to GNIP's Enterprise Data Collector and concluding with the return of matching real time communications as XML activity streams. We next used Google Refine¹ to convert the streams from an XML to a CSV file format so that the files could be imported into Microsoft Excel for data cleansing, analysis and reporting.

GNIP's Enterprise Data Collector is a limited service, and this study did not have access to the advance search operators (boolean operators, delimiters and filters) available through GNIP's premium services—a significant limitation. The messages returned in our study result from simple keyword correspondence between the query and the source documents.

¹ Now Open Refine

Table 1: Search terms used and results.

GNIP Search Terms	Facebook	Google Plus	Twitter
Wagga evacuation	1	20	8
Wagga Wagga evacuation	15	28	50
Wagga emergency services			
Wagga state emergency service	5	16	
Wagga flood	9	54	77
#Waggaflood			11
#flood #wagga	34		
Wagga Murrumbidgee	30	9	5
Wagga rescue	8	11	1
Wagga Wagga	564	221	509
Wagga levy	10		4

5.2 Qualitative media analysis

The methodology applied in this research is an interpretive and qualitative media analysis (QMA). We commenced from the viewpoint that the information architecture framing a message is integral to its communication and combined a qualitative analysis of the messages with an analysis of their materiality and form. The approach we take builds on the work of qualitative content analysts (Altheide, 1996) by categorizing and reporting on the *communicative intent* of the social media messages (i.e. the purpose that message was produced for) and the *topic(s)* described in the messages, e.g. the communicative intent of a message to share a photo or a news story; the topic might be the Wagga floods.

We support our findings with descriptive statistics about the communications. However, unlike quantitative counterparts in content analysis our approach moves beyond ‘counting’ the manifest contents of the media to analyzing the symbolic and latent elements of communication as they relate to the information architecture and context of communication. Taking this approach is essential for understanding the complexity of social media messages where structural elements are not easily separated from the (symbolic and latent) meaning of the message; for example, whether a #hashtag is being used to classify the message subject or as an affective device can only be decided through qualitative analysis.

5.3 Sampling and analysis

Although GNIP provides access to 10,000 social media messages every 24 hours, the data sample for this study was restricted to a small sample size consistent with the objectives of an in-depth qualitative investigation. We commenced sampling activity streams from Facebook, Google Plus and Twitter at 11:30 pm on 5 March 2012, two and a half hours after the New South Wales State Emergency Service issued the order for citizens to evacuate Wagga and continued until reoccupation. We sampled 100 messages from each service at approximately 6-hour intervals for 48 hours, concluding at 12:10am on the morning of 8 March. This resulted in 700 messages each for each Facebook, Google Plus and Twitter.

Messages in the sample required some cleansing; the character sets for a number of messages were unreadable and these were removed from the study. Facebook and Google Plus messages comprise multiple objects and a small number of duplicates (exact copies) were

taken out. A keyword error in the Google Plus stream during collection meant only 359 of the 700 messages were about the Riverina floods. As our study is a rich qualitative case we made the decision to include the 359 messages and account for the different sample size in analysis. GNIP archives messages for 24 hours and as a result our sample tracks reporting to a point before evacuation commenced even though data collection started after evacuation.

Table 2: Summary of messages retrieved through GNIP.

	Facebook	Google+	Twitter
Number of polls collected	7	7	7
Number of activities collected	700	359	700
Number of activities in sample (duplicates and/or errors removed)	676	203	688
Collection commenced	5 March 2012, 11:30pm	5 March 2012, 11:30pm	5 March 2012, 11:30pm
Collection concluded	7-8 March, 11:50pm–12:08am	7-8 March, 11:50pm–12:08am	7-8 March, 11:50pm–12:08am

The analysis was done in Microsoft Excel by two researchers with a third researcher acting as a discussant. The work commenced with an initial phase of familiarization involving the review of the metadata and an investigation of descriptive elements available to support an understanding of the message. Each message was interpreted in the context of its media (Facebook, Google Plus and Twitter) with codes developed iteratively and used to describe the communicative intent and topic before any cross media comparison was considered. From that point researchers worked iteratively, coding from the bottom-up and verifying topics and themes with each other.

6 Findings

6.1 What is available in the document

Every social media message received is described with the same encapsulating metadata. There are variations among those of Facebook, Google Plus and Twitter because of the way each service captures and publishes the information; however the core documents created in each share the same descriptive metadata and are comparable. All messages analysed in this study include fields for the author, content and type of message posted; whether the author posted a note, bookmark or image file; and whether the message is liked, mentioned or shared is uniformly described and available. This core information is enriched with additional supporting, descriptive, administrative and structural metadata for managing digital documents. These are familiar in information science: time and date of creation and revision, uniform resource locators and identifiers (URL and URI), unique identifying numbers, folksonomic tags and signifiers (# and @ in Twitter), links to related information (shared items, homepages, biographies, and avatars), the device the message was created on, and geo-location tracking identifying the place where the message was created.

Table 3 The major metadata elements (architecture) across the three services sampled

Service	Author	Content	Type of post	Time & date	Supporting metadata	Device	Geo location*
Facebook	✓	✓	✓	✓	✓	✓	✓
Twitter	✓	✓	✓	✓	✓		✓
Google Plus	✓	✓	✓	✓	✓		✓

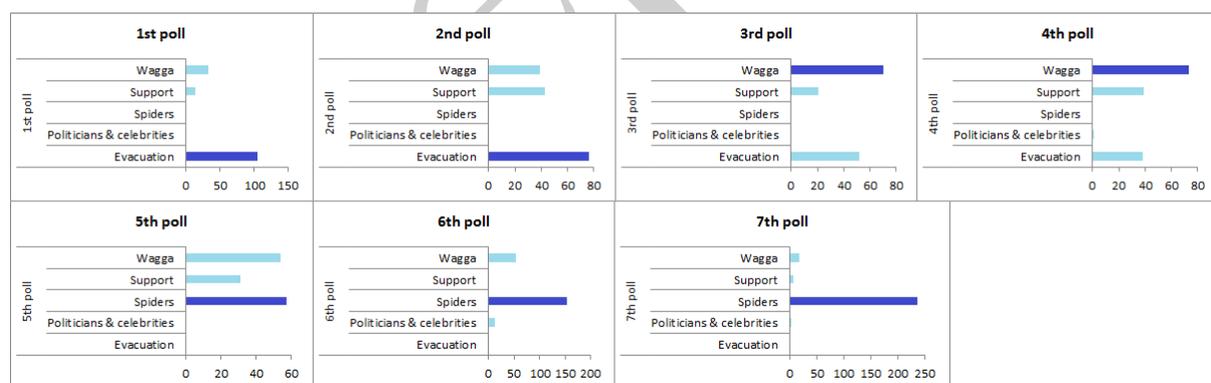
*Geo-location is available only with author permission; in our sample most people had not granted permission

The uniformity of the metadata encapsulating social media activity streams has passed by public collecting agencies practically unnoticed. The possibility of designing a high-level metadata standard for social media, to assist collecting institutions like the Library of Congress, is important. Further research into taxonomy development and metadata crosswalks that improve access and management is needed so that others can make use of publicly available APIs and work with social media.

6.2 The social arc of communication

Figure 3 presents the arc of a straightforward and compelling story told by social media by showing the most prevalent topics overtime. As Wagga received the news to ‘evacuate’ this became the overarching communication theme in the first 24 hours of the crisis. Each of these topics was joined by smaller range of topics as people discussed the issues of the moment—would the levee hold, at what height would the river peak? They turned to news media for information and they posted messages of support to family, friends and the community-at-large. The conversation centred on what was happening in the city, the state of emergency, evacuation centres and community spirit shown by filling sandbags or offering aid to those in need.

Figure 3: Evacuation to reoccupation: prevalent topics over time.



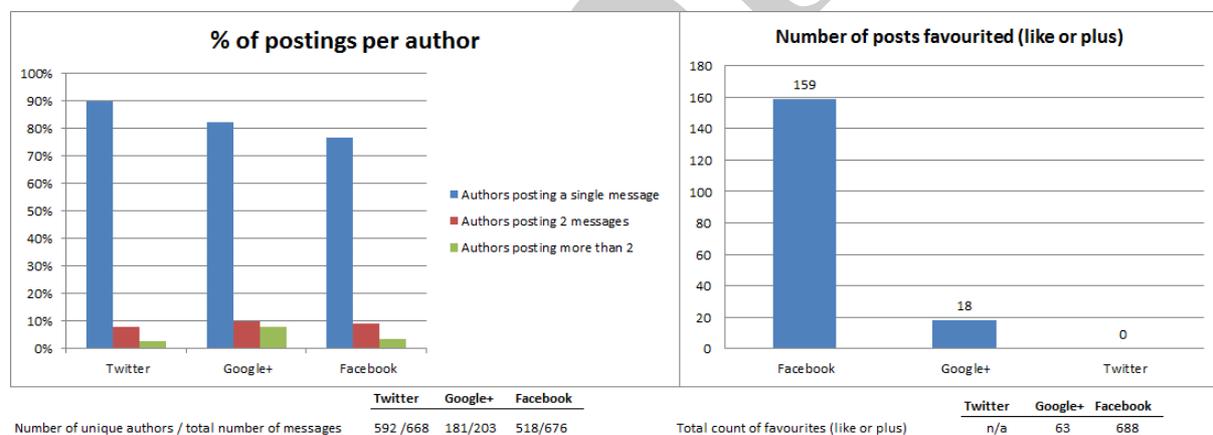
As the tide of the event turned and it became clear that the levee would hold, communications took a lighter tone; humour is a constant response and is present in messages about a visit made by Prime Minister Julia Gillard. On 6 March an unusual natural phenomenon began as millions of usually unnoticed wolf spiders rose from the ground to seek safety. The spiders cast webs above the wet ground with the stunning visual effect of fields that seemed covered by snow. By 7 March ‘spiders’ was the trending topic with people expressing reactions of amazement, awe and fear. A clear sign the danger had passed was the reporting of the floods in the Huffington Post “Weird News” section (Campbell, 2012).

The arc from impending danger to safety is the social document—the evidence of a community’s reaction to a crisis that makes sense when the record is brought together in sequence from the time stamps available in the messages. Linking the evidence between records together in this way presents one of a number of social graphs that could be used for analysis. Our study could just as easily explored the network of relationships between news stories or photographs that were shared within the data set. Independently, each service, each record and the links between records contribute to our understanding of the social document in different ways.

Time, authorship and intention

In any collection of social media, start and end points are arbitrary markers bracketing a discourse that the record ceases to be a part of. What we observed after collection is a document about social communication it is no longer social media. This becomes apparent when taking stock of Facebook ‘likes’ and Google Plus ‘+1’ (Figure 4). ‘Liking’ is a fluid, interactive form of social validation that takes place between people in social media; it can be given and taken away. In our study, it seemed to be used to signal support or appreciation; it was prevalent in messages of *support, thoughts and prayers* or when interesting images were shared. However on returning to our analysis months later, the number of ‘likes’ had often increased; and, in this case, our sample may have been captured too soon to comment accurately on the extent of validation taking place.

Figure 4 Postings across all services.



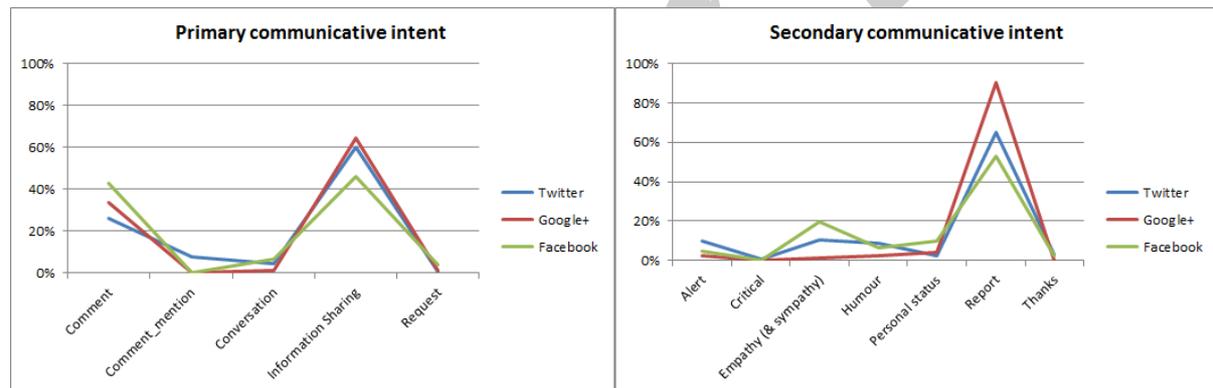
The news to ‘evacuate’ moved quickly in Twitter with 37 of the first 40 messages from the NSW State Emergency Service or the police. These posts were quickly followed by a group of messages from the Australian Broadcasting Authority’s (ABC Radio) Riverina service resending emergency announcements. It is clear that news media and government agencies are using Twitter effectively as a mechanism for broadcasting information that is picked up and shared by a large number of people passing the message along, most of whom communicated only once.

We coded communications across all three services for their primary and secondary intent, that is, to identify the main and ancillary reason for the communication. The analysis clearly showed the main reason was to share information about the floods (Figure 5, primary chart),

usually by reporting current conditions and associated happenings (Figure 5, secondary chart). At least 15.5% of all Facebook messages (including photographs) were attributable to either News media or government sources. A further 5.18% of Facebook and 5.44% of Twitter messages about news coverage indicates the cognitive authority news media receives is significant. Indeed one advantage of collecting the social media record may be its ability to stand as an historic record documenting which news (events and topics) resonated most with people at the time.

The full extent of information sharing in Facebook is hard to gauge since Facebook ‘internalises’ links by bringing data from a linked page directly into Facebook. While title text and the domain name for the linked page are available it is not possible to track back to its origin. The amount of link sharing to images or news stories in Twitter and Google Plus is also very high; in our sample 68.9% of Twitter messages and 46.8% of Google Plus contained links. With the important issue raised by SalahEldeen & Nelson (2012) concerning the number of broken links to web pages, videos and photographs documenting the Egyptian revolution as files disappeared from the Internet, perhaps the views and opinions expressed in the message that carried a link may become the most important documentary evidence of an event.

Figure 5 Communicative intent



Hashtags and rhetorical intent

Twitter’s familiar signifiers (@ #) can simplify the search, however understanding their purpose requires detailed analysis of the rhetorical intent. We coded messages that used the @ symbol to explicitly communicate to another account holder as a *conversation*, for example, “@name we all ok evac of central wagga”. However another purpose of the @ is to refer to a person, for example “The only sandbags @person and I have are under our eyes #waggaflood”; in these cases it was coded *comment_mention*. The code *comment* was used when the primary intent was simply to comment about an event or situation (with no obvious intention of passing on information), for example in Facebook, to express empathy (& sympathy) “Good luck Wagga ... May your levy hold its banks tonight ... And keep all thoughts that matter to me nice and dry!!!”

Over 23% of messages in our Twitter sample included hashtags, however their role in the document is complex. Approximately 51% of messages included a descriptor deliberately

designed to support identification and discovery—the floods’ official tag used by the ABC, NSWSES and NSW police was #nswfloods, and it consistently appears in many early messages. However over time the way the flood is described is determined by the public domain, and #wagga and at #waggafloods became the two the most used tags.

Table 4 Use of # hashtags in Twitter

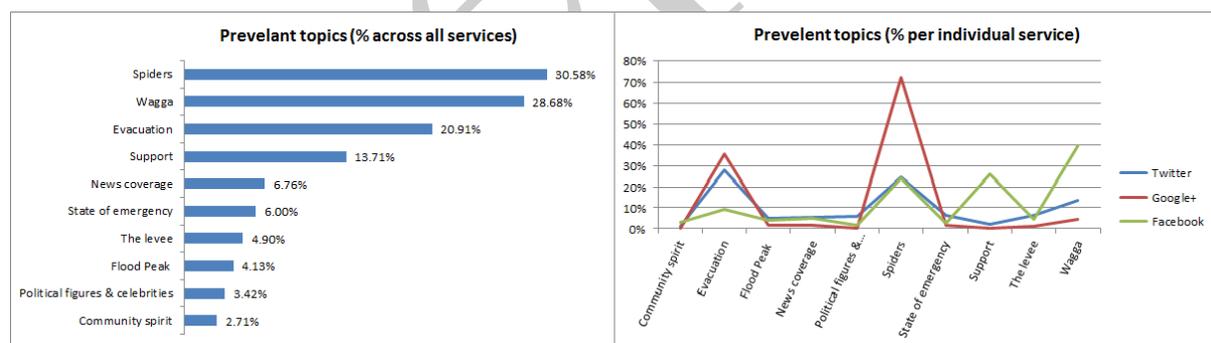
Type	Description	% occurrence
Classify	deliberately assigned to describe a message for the purpose of identification and discovery e.g. #nswfloods	51.6%
Content	describes the message but not intentionally for identification and discovery e.g #ballooning used in association with ‘spiders’	26.7%
Affective	describes the message content affectively, often with rhetorical intent like humour or fear, e.g #freakingmeoutalittle, used in association with ‘spiders’	11.8%
Extend	used expressly for the message/extending the content of the message e.g. : Politicians and journos inundate Wagga Wagga. Expected to exceed #MadeleinePulver #bombhoax levels by 6pm	2.5%

Equally important is the informal way authors improvise hashtags as a rhetorical device to provide emphasis, express emotion, or to share a joke. The documentary role is far from what was intended.

Relationship between the type of media and the topic of conversation

An important observation in the findings is the noteworthy difference among the ways Facebook versus Twitter versus Google Plus are used. These differences have significant implications for social research as the type of social media source examined may impact the construction of an event or social issue.

Figure 6 Prevalent topics



While 13% of all messages posted during the 48 hours expressed *support, thoughts and prayers* for family, friends and the community-at-large, in Facebook one quarter of the messages was about this topic; with more space to write, people are engaging more personally with family and friends. In Twitter they are sharing news and facts in brief 140 character bites.

Evidence from the much smaller Google Plus sample suggests behaviour may be different again with communication among circles being most significant, especially with topics of interest like the wolf spiders. In Google Plus, 31.5% of activities were posts shared from just

seven authors. A single message about spiders posted in Indonesia was shared 64 times, representing 17.8% of the sample, this suggests a different social role for Google Plus with people, expressing points of view or sharing interesting and amusing messages purposefully within their networks.

7 Conclusion

In this paper we have demonstrated the importance of social media as a significant social and cultural record that is used widely in research. We have raised the possibility of selective acquisition and management of social media, as a document of specific events and topics, as an alternative to the Library of Congress's whole-of-archive approach with Twitter. Our study reveals how different social media may document an event differently, thus presenting significant implications for social research. The implication of the importance of basing collection management around the document we intend to create, rather than the platform the message is delivered from, is clear. By undertaking a detailed and in depth qualitative analysis of the Riverina floods, in March 2012, we have demonstrated how the information architecture of social media messages and of the subsequent social document can support our making sense of it for the purposes of knowledge organisation and custodianship.

8 References

- Altheide, D. L. (1996). *Qualitative Media Analysis*. Thousand Oaks, California: Sage Publications Inc.
- Boyd, d. m., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer Mediated Communication*, 13(1), 210-230.
- Boyd, D. M., Golder, S., & Lotan, G. (2010). *Tweet, tweet, retweet: Conversational aspects of retweeting on twitter*. Paper presented at the HICSS-43, January 6, Kauai, Hawaii.
- Briet, S., Day, R. E., Martinet, L., & Anghelescu, H. G. B. (2006). *What is documentation?: English translation of the classic French text: Scarecrow Pr.*
- Bruns, A., Burgess, J., Crawford, K., & Frances, S. (2012). #qldfloods and @QPS: Crisis Communication on Twitter in the 2011 South East Queensland floods *Meia Ecologies Project* (pp. 1-58).
- Bruns, A., & Highfield, T. J. (2012). Blogs, Twitter, and breaking news : the produsage of citizen journalism. In R. A. Lind (Ed.), *Produsing Theory in a Digital World : The Intersection of Audiences and Production in Contemporary Theory* (pp. 15-32). New York: Peter Lang Publishing Inc.,.
- Cameron, M. A., Power, R., Robinson, B., & Yin, J. (2012). *Emergency situation awareness from twitter for crisis management*. Paper presented at the Proceedings of the 21st International Conference Companion on World Wide Web, April 16–20, Lyon, France.
- Campbell, A. (2012). Spiders Blanket Wagga Wagga, Australia After Flooding *Huffington Post*. Retrieved from http://www.huffingtonpost.com/2012/03/09/spiders-wagga-wagga-australia_n_1335332.html
- Charnigo, L., & Barnett-Ellis, P. (2007). Checking out Facebook. com: The impact of a digital trend on academic libraries. *Information technology and libraries*, 26(1), 23-34.
- Dervin, B. (1983). *An overview of sense-making research: Concepts, methods, and results to date*. Paper presented at the Annual Meeting of the International Communications Association, May 1983, Dallas, TX.
- Dutton, W. H., & Jeffeys, P. W. (2010). *World Wide research: Reshaping the Sciences and Humanities*. Cambridge, MA.: MIT Press.

- Etlinger, S., & Li, C. (2011). *A Framework for Social Analytics: including six use cases for social media measurement*. Altimeter Group. San Mateo, CA. Retrieved from <http://www.altimetergroup.com/research/reports/a-framework-for-social-analytics>
- Eversley, M. (2011, 12 July). Library of Congress to store tweets based on Twitter deal, *USA Today*. Retrieved from <http://content.usatoday.com/communities/ondeadline/post/2011/12/1...of-congress-to-store-tweets-based-on-twitter-deal/1-.T4y1k3gWH8s>
- Frohmann, B. (2009). Revisiting “what is a document?”. *Journal of documentation*, 65(2), 291-303.
- Heverin, T., & Zach, L. (2012). Use of microblogging for collective sense-making during violent crises: A study of three campus shootings. *Journal of the American Society for Information Science and Technology*, 63(1), 34-47.
- Kahle, B. (1997). Preserving the internet. *Scientific American*, 276(3), 82-83.
- Li, L. T., Kavanaugh, A., Fox, E. A., Sheetz, S. D., Shoemaker, D., Whalen, T., & Srinivasan, V. (2011). *Twitter use during an emergency event: the case of the UT Austin shooting*. Paper presented at the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times, June 12-15, 2011, College Park, MD.
- Lund, N. W. (2010). Document, text and medium: concepts, theories and disciplines. *Journal of Documentation*, 66(5), 734-749.
- Lyman, P., & Kahle, B. (1998). Archiving digital cultural artifacts. *D-lib Magazine*, 4(7).
- Manovich, L. (2012). Trending: The Promises and the Challenges of Big Social Data. In M. K. Gold (Ed.), *Debates in the Digital Humanities* (pp. 460–475). Minneapolis.
- McGuire, S. (2012). *Social Media and Markets: The New Frontier*. GNIP. Boulder, CO. Retrieved from <http://gnip.com/success-stories/gnip-smm-nf.pdf>
- Neuhaus, F., & Webmoor, T. (2012). Agile ethics for massified research and visualization. *Information, Communication and Society*, 15(1), 43-65.
- Oh, S. (2012). The Characteristics and Motivations of Health Answerers in Social Q&A. *Journal of the American Society for Information Science and Technology*, 63(3), 543-557.
- Papachariss, Z. (2012). Without You, I’m Nothing: Performances of the Self on Twitter. *International Journal of Communication*, 6, 1989-2006.
- Pédauque, R. T. (2003). *Document: form, sign and medium, as reformulated for electronic documents*. Retrieved from <http://hal.archives-ouvertes.fr/>
- Pymm, B., & Lloyd, A. (2007). Dealing with digital collections: interviews with the National library and selected state libraries of Australia. *Australian Academic & Research Libraries*, (September), 167-179.
- Riemer, K., & Scifleet, P. (2012). *Enterprise Social Networking in Knowledge-intensive Work Practices: A Case Study in a Professional Service Firm*. Paper presented at the ACIS 2012: Location, location, location. Proceedings of the 23rd Australasian Conference on Information Systems, 3-5 December, 2012, Geelong, Vic., Australia.
- SalahEldeen, H., & Nelson, M. L. (2012). *Losing my revolution: How many resources shared on social media have been lost*. Paper presented at the Theory and Practice in Digital Libraries, 23 - 27 September, Cyprus.
- Simon, B., Davis, K., Griswold, W. G., Kelly, M., & Malani, R. (2008). *Noteblogging: Taking Note Taking Public*. Paper presented at the SIGCSE 08, The 39th ACM Technical Symposium on Computer Science Education, March 12-15, 2008, Portland, Oregon, USA.
- Woolley, J. K., Limperos, A. M., & Oliver, M. B. (2010). The 2008 presidential election, 2.0: A content analysis of user-generated political Facebook groups. *Mass Communication and Society*, 13(5), 631-652.
- Zappavigna, M. (2011). Ambient Affiliation: A linguistic perspective on Twitter. *Journal of New Media and Society*, 13(5), 788-806.