Insights from a Decade of Twitter Monitoring for Emergency Management

Robert Power
CSIRO Data61
robert.power@data61.csiro.au

Bella Robinson
CSIRO Data61
bella.robinson@data61.csiro.au

Mark Cameron
CSIRO Data61
mark.cameron@data61.csiro.au

ABSTRACT

The Emergency Situation Awareness (ESA) tool began as a research study into automated web text mining to support emergency management use cases. It started in late 2009 by investigating how people respond on Twitter to specific emergency events and we quickly realized that every emergency situation is different and preemptively defining keywords to search for content on Twitter beforehand would likely miss important information. So, in late September 2011 we established location-based searches with the aim of collecting all the tweets published in Australia and New Zealand. This was the beginning of over a decade of collecting and processing tweets to help emergency response agencies and crisis coordination centres use social media content as a new channel of information to support their work practices and to engage with the community impacted by emergency events. This journey has seen numerous challenges overcome to continuously maintain a tweet stream for an operational system. This experience allows us to derive insights into the changing use of Twitter over this time. In this paper we present some of the lessons we've learned from maintaining a Twitter monitoring system for emergency management use cases and we provide some insights into the changing nature of Twitter usage by users over this period.

Keywords

Crisis Coordination, Disaster Management, Situation Awareness, Social Media, System Architecture, Twitter

INTRODUCTION

The Emergency Situation Awareness (ESA) tool provides all-hazard situation awareness information for emergency managers using tweets obtained from the public Twitter API. It collects and processes tweets in near-real-time, enabling effective alerting for unexpected incidents and monitoring of emergency events as they progress with results accessible via an interactive website.

ESA was developed in close collaboration with partners from emergency management agencies in Australia to ensure fitness-for-purpose for the tasks they perform. ESA processes large volumes of tweets and identifies trends and unusual topics using language models (Cameron et al., 2012). A burst detector generates alerts for unexpected high frequency words that are filtered using text mining techniques and machine learning algorithms to identify tweets of interest to users (Yin et al., 2012).

The ESA tool was initially developed as a research prototype in late 2009 and by the end of 2011 we had deployed an online, web accessible operational system that was available 24/7 and is still in operation today over a decade later, see https://esa.csiro.au.

In order to maintain this tool, over the past 10 plus years we have had to deal with normal business as usual activities expected from supporting an operational software system (for example managing scheduled downtimes and unscheduled outages, responding to hardware faults, migration of servers to new IT infrastructure, maintaining regular software updates and third party software patches and upgrades), managing user engagement (responding to user enquiries, providing help desk functions, engaging with high profile clients during emergency events) as well as ensuring its longevity within a research organization amongst competing project activities (championing the continued investment of the tool to senior management, providing support outside normal business hours, and ‘keeping the lights on’ during times of funding shortfalls).

We have also had to respond to changes from Twitter itself. Over this time there have been changes to the Twitter API used to collect tweets, changes in the structure of the JSON data returned by the API, the removal of content from tweets and the inclusion of new elements in the tweet JSON. All these changes have required updates to the ESA software and careful management to ensure continuous 24/7 operation. We have also updated the ESA offering by including new features and upgrading existing functionality over time, such as adopting new machine learning technologies that become best practice, upgrading the text processing third party libraries used and making enhancements to the user interface to maintain an active user engagement.

One of the core functions of the ESA tool is the ability to easily define new tweet collection tasks by defining keywords to search for, users to follow or geographic regions to target. These tweets are then automatically processed in near real-time to find information relevant to specific use cases. We have used this feature in other social media monitoring related projects and as a result the volume of tweets collected over the past decade has expanded beyond our original emergency management scenarios.

The rest of the paper is organized as follows. First background information is presented about the ESA tool, including a summary of how tweets are collected, and an overview of the software elements used. We then review some of the technical challenges encountered in maintaining the tweet collection over this time. Insights into the tweets collected over this period are then presented, including a discussion of some of the trends seen. Finally, we conclude with a summary of our findings.

SYSTEM OVERVIEW

The Early Years

ESA was developed as part of a research platform to investigate how social media content could be used to help the work practices of emergency management agencies in Australia and New Zealand. Our first active case study was in March 2010 when we collected tweets using the search API focused on the impact area as Tropical Cyclone Ului (Wikipedia contributors, 2022) made landfall in Queensland. Similarly, after the two large Christchurch earthquakes on 4 September 2010 and 22 February 2011, tweets from this area were collected and used as test data to refine our tools. Twitter was used since the Application Programming Interface (API) provided a useful and versatile method of obtaining public crowdsourced content and its uptake in Australia at that time had been steadily growing (Digital Marketing Lab, 2010).

These investigations indicated that there was potentially useful information being reported by the public on Twitter. The challenge was to find the information, ensure it was relevant and provide a timely and concise summary to interested people. This preliminary work was encouraging and helped establish a working relationship with the national Crisis Coordination Centre (CCC) in Australia where general-purpose emergency management use cases were defined (Cameron et al., 2012; Power et al., 2014; Power et al., 2015a).

These first tweets were collected using specific keywords expected to be relevant to the incident being monitored. Doing this for the general needs of ‘all-hazard’ social media monitoring for various events of interest, from anticipated natural disasters (floods, bushfires, cyclones) to unexpected emergency events (flash flooding, public unrest, terrorist incidents), would require an ongoing curation of keywords in order to ensure important information would not be missed.

Therefore, in late September 2011, we established location-based searches using the Twitter search API (Twitter, 2022a) for the regions shown in Figure 1 by providing a search location with geographical coordinates and a search radius. These queries are repeated every twenty seconds to obtain the recent tweets. As tweets are obtained from Twitter, they are processed using text mining and machine learning techniques (Power et al., 2013; Robinson et al., 2013a; Robinson et al., 2013b; Yin et al., 2012). From October 2010 through to the end of June 2022, we have collected over 14 billion unique tweets, approximately 6.6 billion original tweets and 7.7 billion retweets. A profile of the tweets collected is described below.

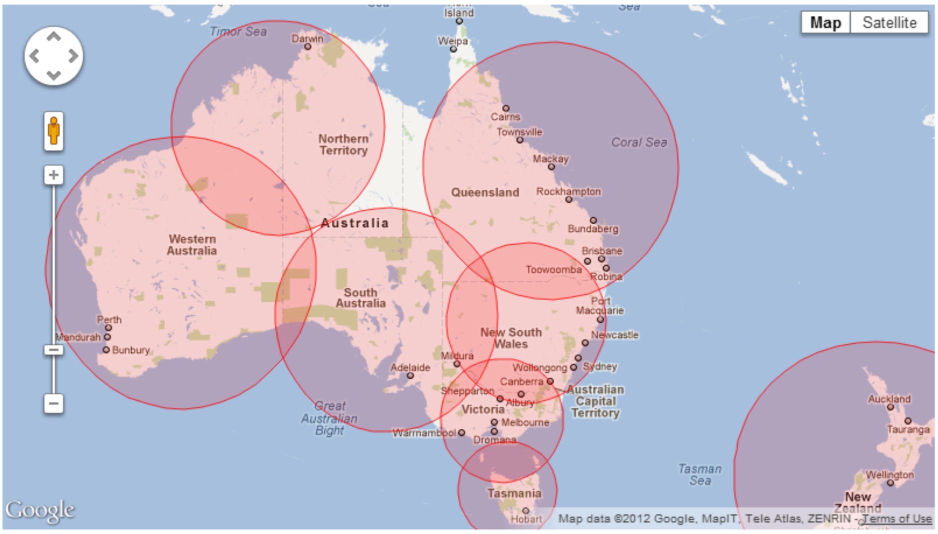

Figure 1: ESA Tweet collection regions.

System Architecture

A schematic of the ESA system architecture is shown in Figure 2. Tweets are gathered from Twitter using various API endpoints and sent to a Java Messaging Service (JMS) instance which makes them available for various backend consumers. The original tweet JSON is filtered with a subset of the contents cached in a database. The tweets are also processed by a Burst Detector which generates alerts that indicate words that have a statistically significant high usage with respect to the Language Model constructed from previous tweets collected (Yin et al., 2012). These alerts are also published to the JMS instance where subsequent consumers store them in a database and process them to target specific keywords of interest which may generate user notifications. Examples are earthquake and bushfire detectors that are triggered when specific alert keywords (for example 'earthquake' or 'fire') are found (Power et al., 2013; Power et al., 2015b; Robinson et al. 2013a; Robinson et al., 2013b).

In Figure 2, the elements represented as clouds are tools we adopt, ovals are Java software we have written, cylinders are databases we maintain, solid rectangles are processing pipelines (burst detection and event detection) and rounded rectangles are the user interfaces available at the web site https://esa.csiro.au.

There are multiple burst detectors (one for each state/territory in Australia, New Zealand, and Australia and New Zealand) and there are multiple event detectors implemented using machine learning classifiers, currently to detect fires and earthquakes. We have in the past maintained multiple repositories as redundant backups, but the current deployment only has one database instance. This repository acts as a cache of recently collected tweets so the original tweet content can be readily displayed on the various user interfaces as required without the extra overhead of subsequent Twitter API calls.

By early 2012, a comprehensive toolset had been developed that included:

- a statistical language model that characterizes the expected discourse on Twitter.
- a burst detector based on the language model to identify deviations from the expected discourse.
- an alerting system that targets specific bursting keywords which generates user notifications.
- clustering techniques for condensing and summarizing information content.
- interfaces supporting forensic analysis tasks.

A key element of this processing is determining the location of the user. This is the role of the Location Mapper component and is achieved in various ways. When ESA started collecting tweets, the meta data included a time zone and UTC offset which was useful to narrow down the location. A small percentage of tweets are geocoded, so this is used when present. We also investigated extracting location cues from the tweet text but predominantly, the user's location field in the tweet JSON was used.

Originally, this self-described location text content was sent to Yahoo Geoplanet/PlaceFinder, a service that determined a location from a text string. These locations could range in geographic extent, from 'Australia', 'New South Wales', 'Sydney' down to a specific place such as 'Sydney Opera House'. We stored the results of the lookup text and location result in a database. This allowed us to progressively build our own cached copy of the Yahoo service results for previously seen locations. This Yahoo service was migrated to a different offering around 2014 based on Yahoo Query Language (YQL) APIs which were then retired around 2019. Since then, we have been using our cached copy of text/location results.

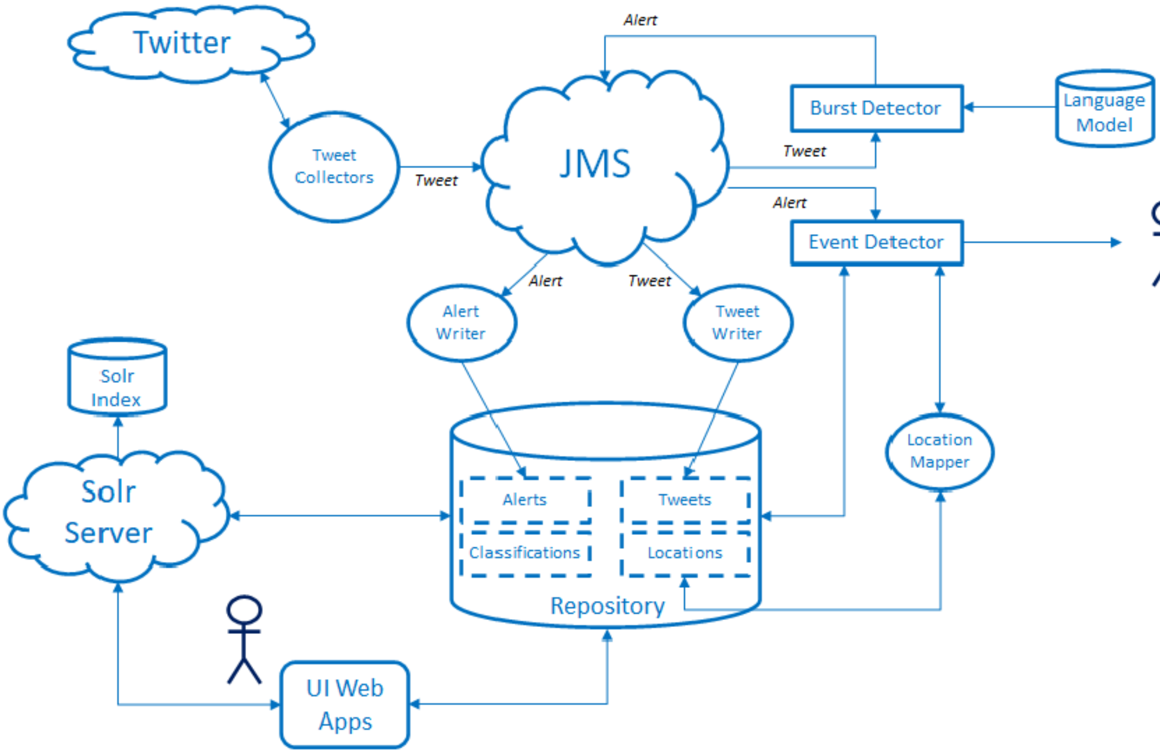

Figure 2: ESA System Schematic

Tweet Collection

Figure 3 shows the number of tweets collected per day for the period 1 October 2011 through to 30 June 2022. In total, 14,381,908,726 unique tweets have been collected during this time with 6,616,304,127 of these being original tweets and 7,765,604,599 are retweets. This period corresponds to 3,926 days and there was only one day where we did not collect any tweets: 16 October 2016, a Sunday.

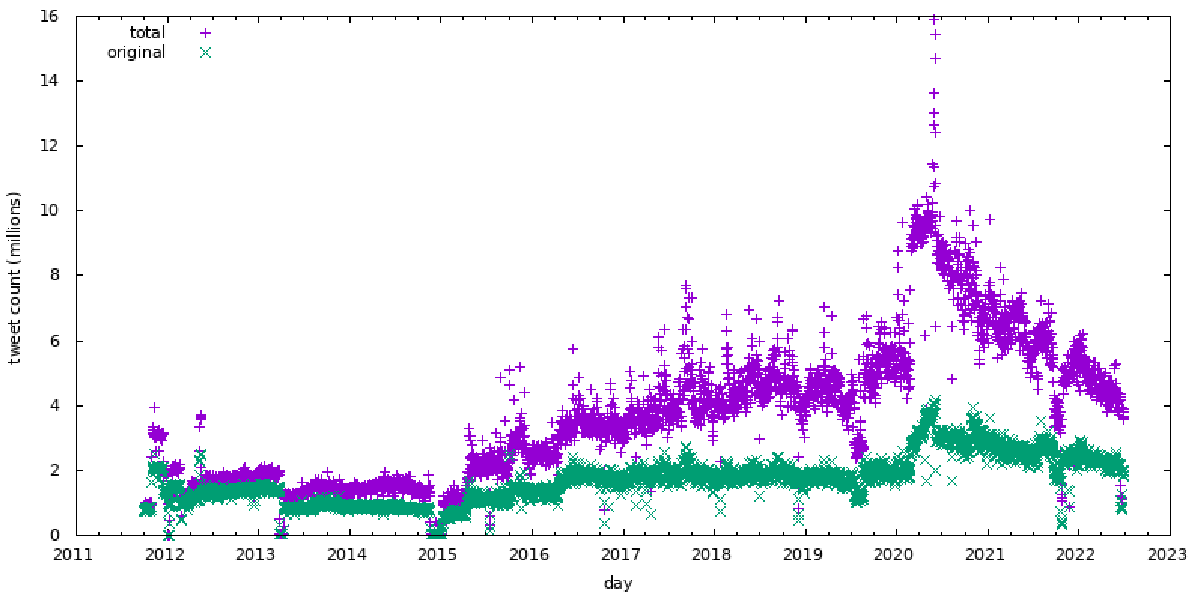

Figure 3: Daily Tweets Collected.

There were a further 56 days where we collected less than 10,000 tweets per day. These days correspond to three periods where we had issues with the location-based search APIs, specifically:

A. 3 January 2012 – 8 January 2012

B. 28 March 2013 – 9 April 2013

C. 22 November 2014 – 4 January 2015.

The location-based Twitter search API, where geographic coordinates and a search radius are provided to obtained tweets from that region, had issues during these periods. In late November 2014 it broke completely. This was discussed on developer forums and Twitter were aware of the problem, but it was not resolved until April 2016, almost a year and a half later.

To maintain the flow of tweets from Australia and New Zealand for ESA, we had to find different ways of collecting content. This was done using other Twitter API endpoints:

- Follow specific users by their id.
  We compiled a list of Twitter accounts managed by emergency services agencies in Australia, news media outlets, politicians, and government departments. This 'authoritative' content is focused on the discourse topics of interest for ESA. We also reviewed our previously collected tweets to curate a list of relatively prolific tweeters and followed these users as well.

- Track keywords.
  Collect tweets using the filter stream that include given keywords. A list of the keywords used for specific emergency events by ESA can be found at: https://esa.csiro.au/aus/help-public.html.

- Geotagged tweets.
  Collect geotagged tweets for specific bounding boxes.

The impact of this change in tweet collection strategy can be seen in Figure 3. There is a dip in tweet volume at the end of 2014, which lifts in early 2015 after we established the new collection methods. When the search API resumed operation in April 2016, the tweet volume lifts again. The volume of original tweets then remains steady through to mid-2020 when there's a noticeable increase, which corresponds to the arrival of COVID-19. Our tweet collection use cases had expanded from emergency management to include health (Sparks et al., 2017) and this resulted in an increase of tweets being collected. There is also a downward trend from the peak of mid-2020 onwards as the COVID-19 discussion abated.

Table 1 provides an overview of the main tweet collection strategies we have used in the past decade, noting the volume of original tweets and retweets collected. This table reflects both the need to use different Twitter API endpoints and the changing use cases of focus for ESA, such as politics, sport, news media, agriculture, science and health. In a couple of instances, we also established tweet collectors for high profile events, such as the Lindt café siege in Sydney and tropical cyclone Maria in Queensland.

As well as focusing on keywords and topics of interest, we also filter out tweets and retweets that are not of interest. Examples are international pop stars such as Justin Bieber, Cody Simpson, Harry Styles, BTS, 5SOS, 1D. We also aim to remove spam and adult content however this has been difficult to achieve and requires constant attention to refine the exclusion filters.

Since we are using multiple collection methods, we can obtain the same tweet multiple times. When including these duplicates, the total number of tweets collected by ESA has been 15,575,816,806 which means we retrieve approximately 8% of duplicate tweets from Twitter. However, in terms of downstream tweet processing we only process unique tweets, for example for alert detection and event notification.

TECHNICAL CHALLENGES

The following sections provide an overview of some of the technical challenges we've overcome in maintaining the tweet processing pipeline and operational system for ESA.

Blacklisting

One of our early explorations to collect tweets was done by modifying a publicly available Twitter client on GitHub, called TwitterClient (Gist, 2010). We changed the code to use the sample stream (the original code used the filter stream) and we ran it for a month to collect tweets and it one day suddenly stopped working. Investigations revealed that the program would no longer work from a specific machine: its IP address had been blacklisted. When we contacted Twitter support, they indicated there was an issue with multiple simultaneous connections being open in violation of the Twitter Rules (Twitter, 2022b). A review of the code changes we made to the original software confirmed that we were doing the wrong thing, which was remedied.

This was an important lesson learnt early: follow the rules or you could be banned. Although the Twitter API is a publicly available resource, its use can, and will be, revoked by Twitter if you do the wrong thing.

Table 1: Tweet collection summary

Num collectors	Started	Finished	Description	Tweets (millions)	Original (millions)
8	Sep 2011	-	Australia/NZ region collectors	3,749	2,546
1	Oct 2011	-	Emergency Management users	187	55
1	Aug 2013		Australia Mexico/Cuba	440	262
1	Dec 2014	Jan 2015	Sydney Lindt Café siege	0.9	0.4
5	Jan 2015	-	Stream collection keywords: fire, flood, earthquake	3,622	1,349
1	Feb 2015		Health keywords	2,207	690
1	Feb 2015	Mar 2015	Tropical Cyclone Marcia, Queensland	2	1
8	Apr 2015	Aug 2019	Following high volume users by state	792	364
3	Sep 2015	-	Stream keywords topics: politics, sport, transport	2,745	1,088
1	Oct 2015	-	Australia/NZ sample stream using time zone	74	32
1	Dec 2016		Australia Day	633	230
1	Feb 2017		Australian media organizations	129	59
1	Sep 2017	-	Australian agriculture	101	36
1	Apr 2018	-	Science topics	24	9
1	Apr 2020	-	Geocoded tweets from Australia and NZ	101	101

Twitter API Changes

Unsurprisingly, the Twitter API has undergone several changes over the past decade. New features have been introduced, new meta data has been included in the tweet JSON and some elements have been removed. Importantly for us in terms of trying to determine the location of the tweet user, the UTC offset and time zone fields were disabled in 2018.

As part of tweet processing, we check for new fields in the tweet JSON and note new ones as they appear. This allows the data to inform us of changes as they occur and highlights new data elements that could be useful. The last time a new JSON field was introduced was recorded on 25 May 2018. The two previous times new fields were introduced, as recorded by our system, was in August 2017 and October 2016. It seems the tweet content has been stable for the last four years, although we have been using API version 1.1 and version 2 is out now and may contain more or different content.

Detecting Earthquakes

One of the early successes of the ESA system was earthquake detection. We developed an email notification system that was activated when tweets were identified from people experiencing an earthquake (Robinson et al., 2013a). This system was based on the combination of tweet heuristics and a machine learning text classifier which had an F1 score of 0.881 and accuracy of 96.85% (Robinson et al., 2013b). We also developed a similar classifier for tweets discussing bushfires, however this was less successful partly due to the ambiguity of language: when people discuss ‘fires’ they may not be referring to natural disasters or an emergency event. However, when people are discussing ‘earthquake’ they are more likely to be referring to an actual earthquake. Or so we thought.

Initially, the main task of the earthquake classifier was to distinguish tweets from people experiencing an earthquake as opposed to a general discourse about earthquakes, for example referencing past events. We did this by manually curating tweets in classes of positive and negative examples and training a machine learning text classifier to distinguish the two classes (Robinson et al., 2013a; Robinson et al., 2013b). The heuristics used were to check that the tweets were sent by people geographically close together (when it is possible to determine the location from the tweet) and that the tweets were sent around the same time. The classifier used several features including the text length, the use of punctuation and the number of retweets. Tweets about an earthquake usually contain few words, include exclamations and are not retweets.

The ESA earthquake classifier was operating successfully for over 12 months until we had two notable incorrect notifications. There was a racehorse called 'Earthquake' which was tweeted about in February and April 2014. These tweets generated two false positive earthquake notifications from our software. We considered the first event a one off but after the second, we needed to revise our classifier by including these false positive tweets as negative examples of earthquake related tweets. The classifier hasn't generated a horse related earthquake detection after this. Table 2 shows a selection of tweets that contributed to these incorrect notifications.

Table 2: Example False positive earthquake tweets in 2014

Tweet text	Date
Favourite Earthquake draws 15. #bluediamond	18 Feb
Earthquake barrier...15 ouch	18 Feb
Barrier 15 for Earthquake! #ouch	18 Feb
Earthquake draws barrier 13. Barrier 13 has had no winners #GoldenSlipper	1 Apr
Earthquake - barrier 13 - wow - makes it tough	1 Apr
Racing: #Slipper Draw:Fav Earthquake draws 13 & eases Mossfun 11; Ghibellines 2; Unencumbered 14; Risen From Doubt 6; Oakleigh Girl 10;Law 7	1 Apr

These tweets are similar in structure to tweets posted by people experiencing an earthquake: they are from people geographically close together around the same time, in this case proximity to a horse racing event.

Tweet Identifiers

Tweet identifiers are numbers consisting of 18 digits in length. When working with some tools, for example Excel, these tweet ids are truncated if they are represented as numbers. For example, a common task in building a text classifier is to curate tweets for human review to allocate them to positive and negative classes as part of the classifier training process. We would export a CSV file that would include the tweet id, text and possibly other information and do the labelling process using Excel. Unfortunately, the tweet ids would be truncated and so the link back to the tweet source would be lost.

As an example, the tweet id '435566610384822272' is truncated to '435566610384822000' and '435568602423373824' to '435568602423373000'and so on. The solution is simply to treat the ids as strings by enclosing the ids in quotes and prefixing with text, for example: 'id: 435566610384822272'. Putting quotes around the number is not enough to resolve this problem. This was probably an issue with other applications and programming languages as well since Twitter introduced the field "id_str" in the tweet JSON. For example, such large integers can cause problems with JavaScript[1].

Hashtags

As noted above, one of the reasons we adopted a tweet collection strategy of 'collect everything' was that it is difficult to know beforehand the language people will use when describing an event of interest for emergency management use cases. This is also true of hashtags before they become established. Soon after the Christchurch earthquake of October 2010, various hashtags were used by those tweeting about it, such as: #earthquake, #quakenz, #nzquake, #nz, #christchurch, #chch. One hashtag started to dominate: #eqnz and this became the hashtag of choice from then on for all earthquakes in New Zealand (Potts et al., 2011).

How the community converge on a given hashtag for a specific event or for an event category is difficult to

[1] https://developer.twitter.com/en/docs/twitter-ids

anticipate. For example, when Tropical Cyclone Marcia struck Queensland in 2015, it was discussed extensively on Twitter. For the approximate 4-week period over February and March 2015 several hashtags were used to describe the impact of the cyclone across the state. The most popular was #TCMarcia. This is a common format used by the community in Australia: prefix the cyclone name by 'TC' for tropical cyclone. However, #CycloneMarcia was also a popular hashtag as were the major town names impacted by the event, such as #Rockhampton, #Mackay, #Gympie, #Gladstone, #Bundaberg. The general hashtag #weather was also used as well as a less serious one: #thisisqueensland.

Location Ambiguity

The initial location-based tweet collectors were focused on the Australian states and New Zealand as shown in Figure 1. When reviewing the content of the location field of the tweets being collected, a notable absence were locations with the text 'Australia'. This was a common location setting for Twitter users from Australia, but we weren't getting these tweets from the location-based tweet collectors. After investigating this, we discovered that searching for tweets in Mexico collected tweets with a location value of 'Australia'. It turns out there is a suburb in Mexico in the region of Saltillo called Australia!

To obtain these tweets, we established a location-based collector with coordinates and radius for this location in Mexico and included a language filter for English tweets. This is shown in Table 1 with the description 'Australia Mexico/Cuba' since Twitter again changed where it thinks Australia is from somewhere in Mexico to Cuba[2]. This location ambiguity is a consequence of the opaque geocoding method Twitter uses to determine the location of tweets. This process is occasionally updated and needs to be reviewed periodically to ensure the tweets being collected correspond to the expected outcome.

TWITTER INSIGHTS

As noted above we have collected and processed over 14 billion unique tweets for the period 1 October 2011 to 30 June 2022 with over 6 billion of these being original tweets and the remainder retweets. Here we present further analysis of the tweets collected to provide an insight to the trends over the period.

Twitter Users

The tweet JSON includes a user id and a screen name of the Twitter user who made the tweet. The screen name can change over time, but the user id remains the same. This allows us to note the number of users who have made tweets collected by the ESA system. This is shown in Figure 4 below where we distinguish between users who have made original tweets and all tweets including retweets.

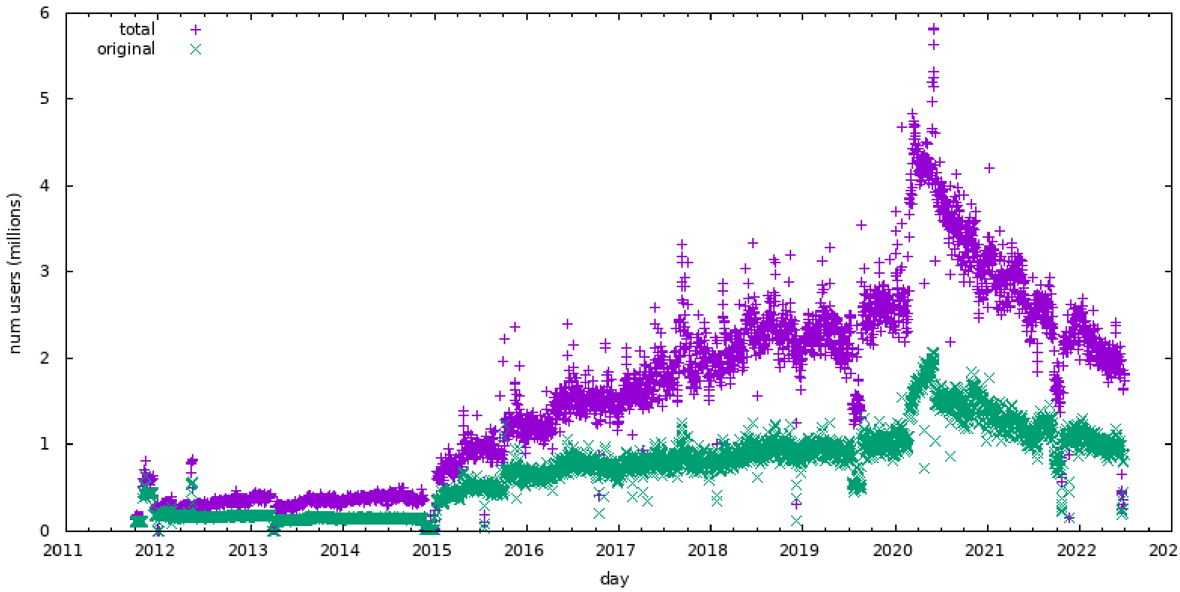

Figure 4: Number of users per day.

[2] This can be seen on Google Maps at: https://www.google.com/maps/place/22%C2%B030'02.9%22N+81%C2%B008'10.0%22W

Figure 4 shows that there are more users who simply retweet rather than provide original content. The impact of the different tweet collection strategies can also be seen in this plot. After the search API broke in late 2014, we switched to a different tweet collection strategy focused on keywords and specific users early in 2015. This resulted in more tweets being collected and interestingly, more users who simply retweet over time.

Tweets Per User

Since the number of users recorded by ESA over time depends somewhat on how the tweets were collected, an alternative insight is to look at the ratio of the number of tweets per user collected each day. This is shown in Figure 5, again distinguishing between original tweets and all tweets, including retweets. These ratios need to account for the category of tweet (retweet or not) and so the ratios are: total tweets/total users and original tweets/total 'original users' (the number of users who posted an original tweet on that date).

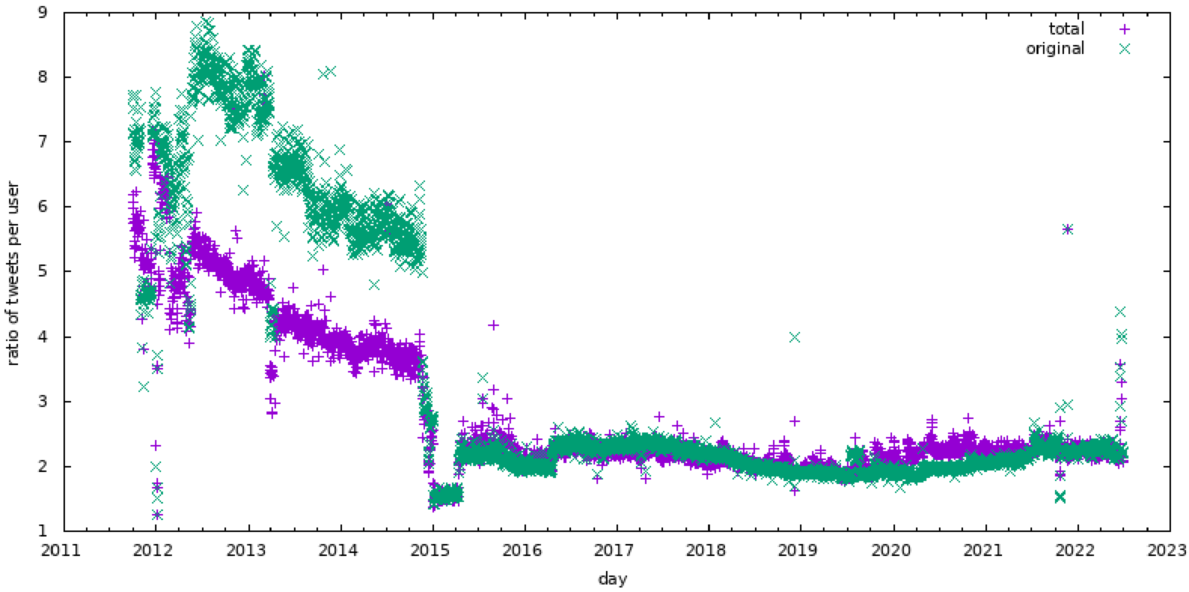

Figure 5: Ratio of tweets per user.

The impact of the different collection strategies is significant in this plot. After April 2015, the tweets per user have been steady and is similar for original tweets and all tweets. Prior to this however, there was a significant difference. The people posting original content on Twitter were doing so more often that those retweeting however this was in steady decline leading up 2015.

Tweet Length

In November 2017, Twitter doubled the character limit from 140 characters to 280 characters. In the tweet JSON object obtained using the Twitter API, the tweet text can be found in the field labelled 'text'. We originally assumed the longer tweet text would simply appear in the same JSON field. This was not the case. A new field called 'full text' was introduced containing the longer tweet text. We didn't discover this initially and it was a few weeks until we made the necessary changes to our tweet collection software to extract the full text.

Figure 6 shows the average tweet length per day for original tweets. The impact of the doubling of the allowable number of characters from 140 to 280 in late 2017 is clearly seen. This figure also indicates why this change was introduced – the average tweet length was steadily increasing and after the extension it seems to have settled to around an average of 120 characters.

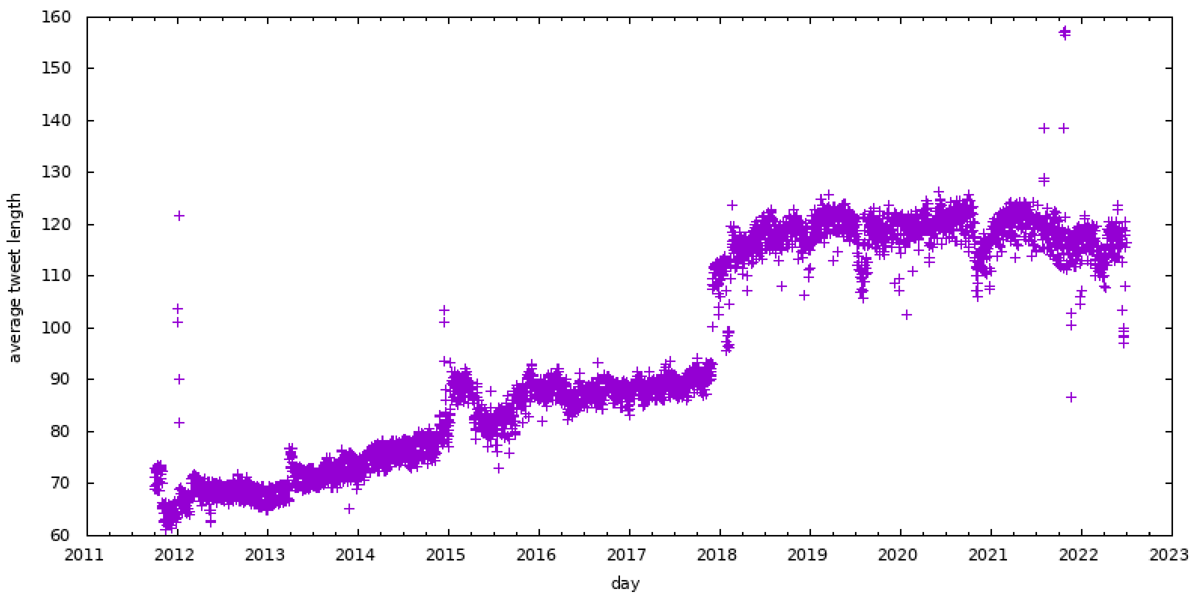

Figure 6: Average (original) tweet character length per day.

CONCLUSION

The ESA tool has been collecting and processing tweets about specific topics, mostly focused on emergency management use cases, for over a decade. Keeping the system running during this time has been a challenge not just from a software and operational systems perspective, but also from a corporate and governance one. A large part of the continued success has been the dedication of those involved. Without them it would not have lasted so long.

This experience has allowed us to reflect on some of the insights and learnings we have accumulated along the way. In summary:

- Know what you're doing, or you may be blacklisted.
- The Twitter developers have been responsive to questions, but fixes may take some time.
- Be adaptable in your tweet collection strategies.
- Review the content you collect to ensure it's what you're expecting.
- It's difficult to know beforehand how people will describe an emergency event.
- Be careful using numeric tweet ids when exporting tweets.
- When using the Twitter APIs, keep up to date with the changes to the Tweet JSON structure.
- Don't rely on third party services to be around forever.

ACKNOWLEDGMENTS

The authors thank the contributions of our CSIRO colleagues who have worked on ESA over the years: John Colton, Sarvnaz Karimi, Michael Kearney, Andrew Lampert, Peter Marendy, Claire Mason, David Ratcliffe, Saguna, Brooke Smith, Ross Sparks, Gavin Walker, Allan Yin and Jie Yin.

ESA has been financially supported by numerous government departments, funding bodies, emergency services agencies and benefited from collaborations with advocates within the Australian emergency management community and championed by senior leadership within CSIRO. Without this support our journey would not have lasted as long as it has.

REFERENCES

Cameron, M., Power, R., Robinson, B. and Yin, J. (2012) Emergency situation awareness from twitter for crisis management, *Proceedings of the 21st international conference companion on World Wide Web, WWW '12 Companion*, New York, NY.

Digital Marketing Lab (2010) 2010 Australian Social Media Compendium, *DML*. Retrieved July 6, 2022, from http://digitalmarketinglab.com.au/2010/03/28/2010-australian-social-media-compendium/

Gist, Inc. (2010) TwitterClient. A Java client for the Twitter Streaming API, *GitHub*, Retrieved July 4, 2022, from https://github.com/gistinc/TwitterClient.

Potts, L., Seitzinger, J., Jones, D., and Harrison, A. (2011) Tweeting disaster: Hashtag constructions and collisions, *Proceedings of the 29th ACM International Conference on Design of Communication*, New York, NY.

Power, R., Robinson, B. and Ratcliffe, D. (2013) Finding Fires with Twitter, *Proceedings of the Australasian Language Technology Association Workshop 2013*, Brisbane, Australia.

Power, R., Robinson, B., Colton, J. and Cameron, M. (2014) Emergency situation awareness: Twitter case studies. *International conference on information systems for crisis response and management in Mediterranean countries (ISCRAM Med)*, Toulouse, France.

Power, R., Robinson, B. and Wise, C. (2015a) Using Crowd Sourced Content to Help Manage Emergency Events. In: Nepal, S., Paris, C., Georgakopoulos, D. (Eds) *Social Media for Government Services*, Springer, Cham. https://doi.org/10.1007/978-3-319-27237-5_12

Power, R., Robinson, B., Colton, J and Cameron, M. (2015b) A Case Study for Monitoring Fires with Twitter, *Information Systems for Crisis Response and Management, ISCRAM 2015*, Kristiansand, Norway.

Robinson, B., Power, R. and Cameron, M. (2013a) A sensitive Twitter earthquake detector, *Proceedings of the 22nd international conference companion on World Wide Web, WWW '13 Companion*, Rio de Janeiro, Brazil.

Robinson, B., Power, R. and Cameron, M. (2013b) An Evidence Based Earthquake Detector Using Twitter, *Proceedings of the Workshop on Language Processing and Crisis Information 2013*, Nagoya, Japan.

Sparks, R., Robinson, B., Power, R. Cameron, M. and Woolford, S. (2017) An investigation into social media syndromic monitoring, *Communications in Statistics - Simulation and Computation*, 46, 8, 5901–5923.

Twitter (2022a) Standard search API, *Twitter Development Platform*, Retrieved July 4, 2022, from https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets

Twitter (2022b) The Twitter Rules, *Twitter Help Centre*, Retrieved July 4, 2022, from https://help.twitter.com/en/rules-and-policies/twitter-rules.

Wikipedia contributors (2022) Cyclone Ului, *Wikipedia, The Free Encyclopedia*, Retrieved July 4, 2022, from https://en.wikipedia.org/w/index.php?title=Cyclone_Ului&oldid=1113982358

Yin, J., Lampert, A., Cameron, M., Robinson, B. and Power, R. (2012) Using social media to enhance emergency situation awareness, *IEEE Intelligent Systems*, 27, 6, 52-59.