

# Social Media Monitoring for Health Indicators

Bella Robinson, Ross Sparks, Robert Power and Mark Cameron

*Commonwealth Scientific and Industrial Research Organisation*

*Email: [robert.power@csiro.au](mailto:robert.power@csiro.au)*

**Abstract:** Social media has been recognised as a new source of information from the general public to help achieve positive social outcomes. Some examples are detecting earthquakes, monitoring ongoing disaster events, tracking public opinion, marketing, human behaviour research and public health issues. Given the large volume of information available on numerous social media platforms currently in use, a significant challenge is to extract meaningful and relevant information for these different purposes.

In the area of health research, social media has been investigated to provide health information to the community for the purposes of early warning or intervention, preparedness and targeted health advice. Crowd source content has also been used for disease mapping, see for example Google flu trends, <https://www.google.org/flutrends/au/#AU>, while information published on social media has been identified as an indicator for public health issues, such as detecting influenza epidemics (Aramaki et al. 2011). The importance of early detection of large-scale contagious disease outbreaks and the ability to understand how a population is reacting to such events, whether naturally occurring or as a result of bioterrorism, is of interest to governments world-wide.

Health monitors and decision makers need credible early signals of disease outbreaks. Although this is difficult due to the variability of health monitoring capabilities, early warnings combined with available key data could be used for a number of improved population health outcomes such as estimating the spatio-temporal spread of diseases, severity of disease outbreaks, projected peak time and duration of disease outbreaks, the use and effect of early mitigation measures and the targeted deployment of limited medical resources. This has the potential to augment and complement existing information to reduce the cost of information gathering and analysis to increase the productivity, responsiveness and planning for health agencies to achieve a new perspective on population health for government agencies and health professionals.

In Australia, CSIRO have been investigating these techniques using statistical data mining methods and natural language processing procedures, such as text classification and unsupervised clustering, applied to messages published on Twitter to identify content of relevance to emergency managers. A large collection of tweets from Australia and New Zealand have been processed since late 2011 to identify unexpected emergency incidents and to monitor ongoing disaster events (Yin et al. 2012; Power et al. 2014).

This previous work has been adapted to develop an investigative tool using content published on Twitter to provide indicators of population health and well being. The aim was to conduct a preliminary feasibility study to better understand the potential for detecting and alerting on medical symptoms in on-line communities using social media postings. The following two key questions were investigated:

1. Is it feasible and valuable to detect and alert on unusual variations in medical symptoms within online Australian communities monitored through social media?
2. Can social media monitoring, equipped with novel statistical and online data mining algorithms, provide reliable early evidence of disease outbreak?

This paper reports on our experience to date which includes preliminary positive results indicating that health issues such as colds, influenza and fever expressed by the general public can be identified from tweets originating from Australia. These results need to consider the issue of selection bias inherent in the Twitter data source before population inferences can be made.

**Keywords:** *Health monitoring, social media, syndromic surveillance, Twitter*

## 1. INTRODUCTION

The main task of the project was to determine if evidence of infectious disease outbreaks could be detected using Twitter data. The process undertaken was to use symptom keyword counts in tweet messages as indicators of potential disease outbreaks. The advantage of tweet data is that information can be identified about infections that are not sufficiently severe enough for the sufferer to present for medical attention at hospital emergency departments. This allows near misses in severe infections to be identified which provides different information to official emergency department data.

In summary, the process undertaken was to: define symptom words and phrases; collect Twitter data as evidence; define statistical data models and perform analysis; develop an interactive demonstrator for users to explore the findings; and report results, analysis and recommendations. The structure of the paper reflects these steps.

## 2. SOCIAL MEDIA MONITORING FOR SYNDROMIC SURVEILLANCE

### 2.1. Symptom words

When people tweet about feeling unwell they are likely to express the symptoms they are experiencing. These symptoms are expected to be a result of health conditions they have. The conditions of focus for the investigation are: cough/cold, diarrhoea, fever, influenza (flu), stomach flu, unwell and vomiting. These conditions were chosen as being of interest for syndromic surveillance of public health risks as a result of influenza, food poisoning and other infectious diseases. The real-time nature of Twitter combined with the objective of identifying health risks suggests that content published on Twitter should be analysed for symptoms not conditions.

A list of symptom keywords and phrases was prepared for each condition. Note that some phrases can be symptoms for multiple conditions. At the same time an attempt was made to eliminate any keywords or phrases that might be used to reflect something other than an individual's well-being (e.g. 'that person makes me feel sick' rather than 'I feel sick'). As an example, the phrases used for influenza are shown in Table 1 below. Note that matching is case in-sensitive and a '\*' is a pattern 'wild card' that matches zero or more characters. For example, 'have flu\*' matches 'have flu', 'have flu-like' and 'have flu!'. However it also matches, for example, 'have flute', 'have fluid' and 'have fluffy'. For the seven conditions examined, there were a total of 228 phrases used, with 28 containing the '\*' wild card.

**Table 1.** Symptom phrases for influenza.

am getting the flu	got pestilence	got the scourge	flu coming on	got grippe	have flu*
have the affliction	have pestilence	flu is coming on	got the grip	have grippe	got grip
got the infestation	got pneumonia	have the curse	have the grip	influenza	feel flu
have the infestation	have pneumonia	got the curse	getting flu	got flu	have grip
have the pestilence	have the scourge				

This approach of using keywords is a common method for identifying candidate tweets that may provide evidence of people reporting they are suffering a health condition, see for example Sadilek et al. (2012), Signorini et al. (2012), Zuccon et al. (2015). We have chosen phrases that aim to specifically target people self reporting that they have symptoms, hence the use of words such as 'am', 'got' and 'have'.

### 2.2. Data Gathering

The tweets used were obtained from the CSIRO Emergency Situation Awareness (ESA) tool. ESA collects tweets published in Australia and New Zealand and has previously been used to explore how to effectively identify tweets of interest for emergency coordinators during times of natural disasters and emergency events, such as finding bushfires (Power et al. 2013) and detecting earthquakes (Robinson et al. 2013). Note that only the tweets from Australia were used for this study.

The data used for this investigation were tweets collected during the fifteen month period of 1 July 2013 through to 30 September 2014, spanning two winter seasons. Text mining was performed on the tweet repository to find the number of tweets that contained any of the symptom keywords or phrases associated

with each condition during the target period. These counts were aggregated by hourly intervals generating time series count data for each condition based on the corresponding symptoms for each. In total, 81,236 tweets were collected, an average of just under 178 tweets per day. There were 42,549 tweets found for 'unwell', the most tweets for any condition tested, while 'diarrhoea' only found 1,660 tweets.

### 2.3. Statistical Data Models

Statistical process control (SPC) methods were applied to this Twitter data to examine if the onset of unusual public health events could be detected. The hourly symptom tweet counts represent the frequency of occurrence of the associated underlying health condition which can be monitored using SPC. In particular, when there is no potential outbreak of a specific health condition, the frequency of counts would be expected to demonstrate the characteristic of a stable process (not necessarily at the zero level) over time. Symptom counts that are stable, and thus predictable, are defined to be in-control and therefore would be expected to be manageable with the generally available public health resources. The goal of the SPC monitoring is to detect when the symptom counts become unstable and increase beyond what would be expected for an in-control process. This would occur when there is a significant increase in the number of tweets that contain symptom keywords which could suggest an outbreak of the corresponding health condition that may require additional public health resources.

Three SPC approaches were applied to the data and assessed for their ability to detect significant increases in the expected health condition counts. In particular, dynamic bi-plots, adaptive exponentially weighted moving averages (AEWMA) and adaptive cumulative sums (ACUSUM) were chosen for their applicability to real-time data of this nature and their capability to facilitate early detection for relatively small tweet counts per time interval.

There are two rationales for applying these different detection methods. Firstly, it is reasonable to assume that the conditions in the data may be present at the same time in the same tweets (for example fever and flu) due to an individual experiencing symptoms associated with multiple conditions. In this case there may be additional power for detection of an out-of-control event by using a multivariate detection method such as the dynamic bi-plot (Sparks et.al. 1997). In addition, the multivariate method has the advantage of controlling the overall false detection rate. However, if the incidence counts associated with the different conditions are not highly correlated, it may be as effective and possibly simpler to consider the conditions separately for the purpose of detecting out-of-control events. Consequently, in this case we consider the AEWMA and the ACUSUM for detection as outlined in Sparks et.al. (2010).

Note that all of these approaches are adapted to be effective with small tweet counts as could be expected during the early onset of symptoms. The aim is to demonstrate the efficacy of these methods for use in the early detection of the conditions being considered.

### 2.4. Analysis

In summary, the data was split into a training period (1 July 2013 to 30 June 2014) where retrospective surveillance was applied to determine the in-control data for designing the prospective surveillance period (1 July 2014 to 30 September 2014). The findings suggest that moderately small potential outbreaks can be detected within a few days of first surfacing and for large potential outbreaks within hours. This suggests that it could be useful for monitoring and controlling adverse bioterrorism events. The methodology used provides the ability to view multivariate trends and diagnose its nature in the same plot, a dynamic bi-plot, which is useful for describing the nature of outbreaks when they present simultaneously with more than one symptom. These findings are described in detail in Sparks et al. (2015) for flu.

The dynamic bi-plot is useful when there are correlations between condition counts in the tweets. If the condition counts have very little correlation, there may not be much advantage in monitoring them together unless we are trying to diagnose what symptoms appear to be correlated when an increase in condition counts occurs. Since the dynamic bi-plots have been described in detail elsewhere, Sparks et al. (2015), we focus here on the uni-variate processes instead.

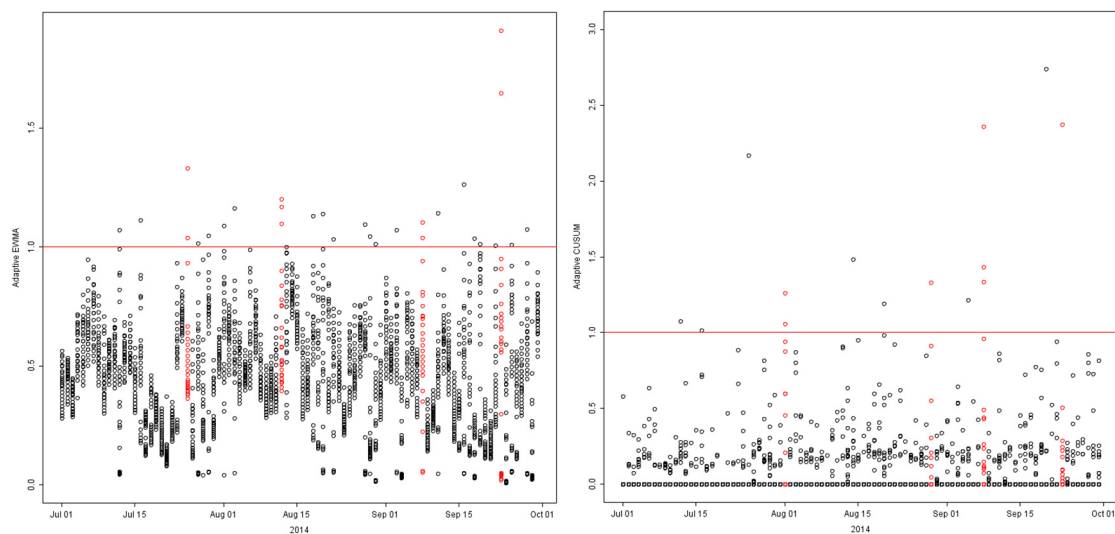
#### *Adaptive Exponentially Weighted Moving Averages*

The Adaptive Exponentially Weighted Moving Averages (AEWMA) is one method of detecting increases in the counts independently. AEWMA is applied to the forecasted counts for each condition after fitting an empirical model to the raw data of counts for that condition. This model utilizes a separate Poisson model for each condition and accounts for the natural sources of variation in the data for each condition due to the hour of the day, date (season) and day of week in order to remove broad based trends and cycles while allowing

the capture of local changes in the residuals. The model utilizes harmonic terms to capture the within day variation and any annual variation and indicator variables for day of the week effect. There are more than a dozen terms in the model, values for which were obtained through a process of fitting the model using the training data. Full details can be obtained by contacting the second author and this work forms part of an upcoming publication.

In applying the adaptive monitoring plans to hourly counts, the test statistic is reset to its initial state after each outbreak flag. In other words the adaptive EWMA is reset to zero, and the process is started again each time the adaptive statistic exceeds the value of 1. It is not that unusual for the test statistic to be greater than 1 once in a day – it is expected to occur once in four days. However, to have two or more such observations in one day is unusual. We use ‘two or more values greater than one in a day’ as the criterion to detect a potential outbreak event which results in a false detection rate of less than 0.0048 per day, which is approximately one false alarm every 208 days.

An example for flu is shown on the left in Figure 1. The red markers are where a potential outbreak has been detected; there are four of them on 25 July, 12 August, 8 September and 23 September.



**Figure 1.** Hourly Adaptive EWMA for flu (left) and Hourly Adaptive CUSUM for flu (right).

### *Adaptive Cumulative Sums*

The Adaptive Cumulative Sum (ACUSUM) plan is also based on Sparks et al. (2010) and is only optimal for the zero state process. In this case the reference value is defined as the optimal zero state Cumulative Sum plan for homogeneous Poisson counts and it adapts to changes in the expected counts if these expected values are non-homogeneous. Optimality only kicks once an increase in counts is large enough, and then the plan adapts to the optimal plan for detecting of that average size. The model is defined to deliver a low false alarm rate of 1 in 100. With hourly data this amounts to a false alarm roughly every 4 days ( $24 \times 4 = 96$ ) which may be too frequent. However the option of using an average run length of 1000 rather than 100 is onerous but could be undertaken as future work.

As described above for AEWMA, the same detection method is applied for ACUSUM: the test statistic is reset to its initial state, that is the adaptive CUSUM is reset to zero, and the process is started again each time the adaptive statistic exceeds the value of 1. A detection occurs when two or more observations in one day exceeds the value of 1. An example plot is shown on the right in Figure 1. Potential outbreaks for flu are detected on 1 August, 29 August, 8 September and 23 September.

The AEWMA was very close to flagging an outbreak on 1 August, on the 28 August the EWMA was set at the lower limit and therefore finds it difficult to flag the next day. This generally will not happen unless the increases are very unusual. There are no obvious explanations for the other misses, mostly because the CUSUM is dependent on when it is reset to the initial value.

### **3. THE INTERACTIVE DEMONSTRATOR**

The demonstrator is a map-based interactive web site to allow users to explore the findings of this investigation. Example screen shots are shown in the figures below. The aim is for the user to investigate the

results in terms of plots that represent the statistical findings as well as providing access to the individual tweets upon which they are based. The demonstrator, available at <https://smm.csiro.au/ss1/tweets.html><sup>1</sup>, consists of two main web pages, the Tweet Explorer and Statistical Analysis, each described below.

### 3.1. Tweet Explorer

An example of the tweet explorer is shown in Figure 2. The user can enter the required date range and select up to seven target conditions. The tweets matching these criteria are retrieved when the ‘Get Symptom Tweets’ button is activated. This shows the tweets on the right which are colour coded with dots to indicate the matching condition. A summary of the tweets obtained is shown as a header to the list of tweets: how many tweets, for which conditions and the date range. Note that a maximum of 2000 tweets can be obtained for performance reasons. In Figure 2, all conditions have been chosen for the date range 2 – 8 July 2014 which finds 1247 tweets. We have also built text classifiers to identify flu and fever conditions and tweets considered positive examples by the classifiers are also highlighted in the interface; two have been found in this example.

The map in Figure 2 shows where the tweets originate from and can be adjusted to show tweets by volume. This provides a visual ‘condition intensity’ which is obviously biased by the number of people tweeting and so normally corresponds to the areas of higher population. Selecting an icon on the map shows the tweets at this location. Alternatively, for geo-coded tweets, an icon is displayed in the list of tweets: when selected, only this icon is shown on the map. An example of this icon can be seen in Figure 2 for user ‘Gordy m8’. It is interesting to note that this user’s profile location is listed as ‘Glorious Sydney’ and this tweet has a geo-coded coordinate in the Sydney area. To again show all the icons for the geo-coded tweets, select the ‘Map All Tweets’ link on the top right.

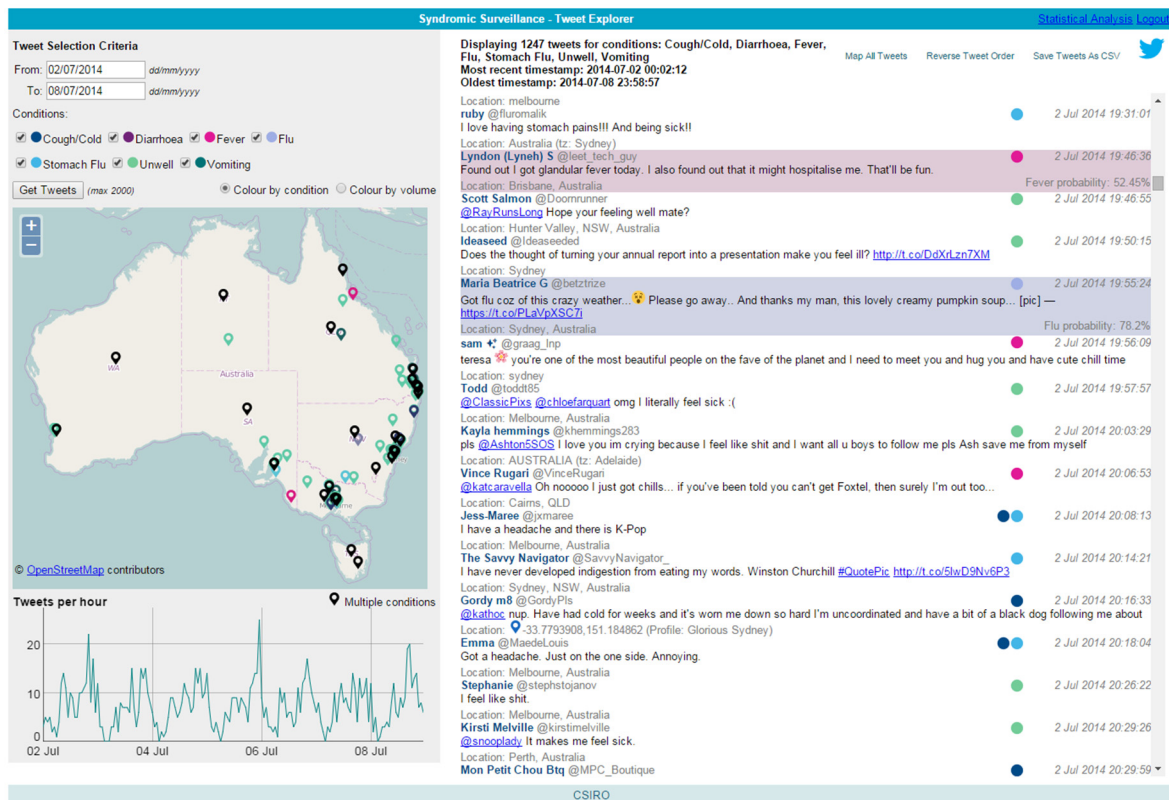


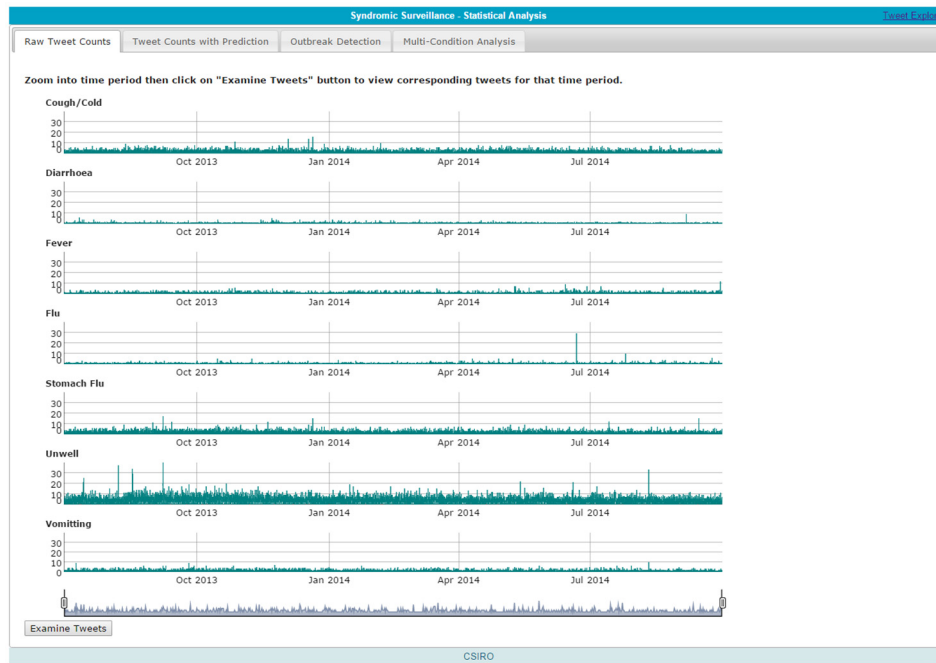
Figure 2. Syndromic Surveillance Demonstrator: Tweet Explorer.

A tweet timeline is displayed below the map showing the number of tweets in hourly intervals, or daily if the number of days chosen is greater than seven. Individual tweet counts can be found by hovering the mouse over the timeline which shows the date and time along with the number of tweets found in that time interval.

<sup>1</sup> Please contact the authors for access credentials.

### 3.2. Statistical Analysis

An example of the statistical analysis web page is shown in Figure 3, showing the first of the four tabs available. The ‘Raw Tweet Counts’ tab provides an overview of the hourly tweet counts for each condition spanning the full data collection period 1 July 2013 – 30 September 2014. The plot on the bottom of the page is used to focus on a period of interest by using the sliders on the left and right. Note that these plots are linked such that zooming the time period effects all plots together.



**Figure 3.** Statistical Analysis: Raw symptom tweet counts.

These plots can also be navigated by dragging the mouse to the time period of interest. This is especially useful when a small time interval is needed since the slider at the bottom is cumbersome to operate in such situations. To quickly reset the original full time period, double click on any of the top six plots. The actual tweets corresponding to a period of interest can be found by navigating back to the tweet explorer using the ‘Examine Tweets’ button in the lower left hand corner. Moving the mouse over any one of the seven condition plots shows the tweet count for that hour on the right.

The ‘Tweet Counts with Prediction’ tab show the raw hourly or daily tweet counts for the selected condition during the time period 1 July 2014 to 30 September 2014 along with the predicted values in red. The predicted values are calculated using the 1 July 2013 to 30 June 2014 counts as input to a generalized linear model. The hourly counts are low, usually ranging from 0 to 2, and the predicted values oscillate between 0 and 1 depending on the time of day. The ‘Outbreak Detection’ tab provides access to the univariate AEWMA and ACUSUM control charts for the testing period 1 July 2014 to 30 September 2014. Options are available for the user to select one of the seven conditions and to choose the preferred detection method, AEWMA and ACUSUM. The plots in Figure 1 are available using this interface. The ‘Multi-Condition Analysis’ tab provides access to the multivariate dynamic bi-plots for the testing period 1 July 2014 to 30 September 2014. Options are available for the user to choose different outbreak detection approaches, as explained in Sparks et al. (2015).

## 4. CONCLUSIONS AND FUTURE WORK

This paper describes an investigation of using statistical process control methods to identify potential disease outbreaks from people self reporting their health status on Twitter. We have applied these methods to detect the onset of unusual public health events by comparing consecutive winter periods as well as comparisons about departures from a model that could be based on several years of data. Thus it has the ability to describe what symptoms reported in tweets are higher than expected. We have developed a statistical model that can be used to identify unusual tweets potentially corresponding to a disease outbreak. This can be determined using both univariate and multivariate techniques with the results available online. The evidence supporting these likely outbreaks can be closely examined using the interactive web interface.

This work should be integrated with other data sources, such as Google flu trends, emergency department statistics, syndrome databases and pathology reports. These datasets each have different characteristics that complement each other in ways that remain open research topics. The main benefit of our approach is that it offers a near-real-time insight into people self-reporting their health conditions. Also, these tweets can be tracked to form an understanding of the burden of disease which is not possible with other data sources. This is an area of potential future work.

This work can be used to establish a near-real-time detector that notifies interested users of potential disease outbreaks. This information might be valuable for other international initiatives such as the US Biosurveillance Portal and the Biosurveillance Ecosystem. Also, our syndromic surveillance work can be applied to other national security social media monitoring tasks such as terrorism and criminal behaviour. These activities are likely to focus on the communication network, understanding who is communicating with who, identifying persons of interest and their friends and followers (in Twitter) and understanding when there is unusual communication activity such as bursts of communication between individuals or groups.

The selection bias of using Twitter data is acknowledged; although no authoritative data is available to verify this, it is expected that the predominant users of Twitter in Australia are a particular cohort of the population, generally younger people living in metropolitan areas. However, all health-related data sets will have a selection bias of one form or another, for example some doctors are more likely to request pathology tests than others. Future work could investigate and characterise these data issues to provide insight into the groups of people various data sets represent.

#### ACKNOWLEDGEMENTS

This work has been financially supported by the Commonwealth of Australia Defence Science and Technology Organisation National Security Science and Technology Centre.

#### REFERENCES

- Aramaki E, Maskawa S, Morita M (2011). Twitter catches the flu: Detecting influenza epidemics using twitter. In: *Proceedings of the Conference on EMNLP '11*. pp. 1568–1576 (2011)
- Power R, Robinson B, Colton J, Cameron M (2014) Emergency Situation Awareness: Twitter Case Studies. *Information Systems for Crisis Response and Management in Mediterranean Countries - First International Conference*. ISCRAM-med 2014, Toulouse, France, October 15-17, 2014. pp. 218-231.
- Power R, Robinson B, Ratcliffe D (2013) Finding Fires with Twitter. *Proceedings of the Australasian Language Technology Association Workshop 2013 (ALTA 2013)* December, 2013, Brisbane, Australia. pp 80-89.
- Robinson B, Power R, Cameron M (2013) An evidence based earthquake detector using Twitter. *Proceedings of the Workshop on Language Processing and Crisis Information (LPCI) 2013*. October 14, 2013, Nagoya, Japan. pp 1-9.
- Sadilek A, Kautz HA, Silenzio V (2012). Predicting Disease Transmission from Geo-Tagged Micro-Blog Data. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*. 2012.
- Signorini A, Segre AM, Polgreen PM (2011). The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS one*. 2011; 6(5):e19467.
- Sparks R, Adolphson A, Phatak A (1997), Multivariate Process Monitoring Using the Dynamic Biplot. *International Statistical Review*, 65: 325–349. doi: 10.1111/j.1751-5823.1997.tb00312.x
- Sparks R, Carter C, Graham P, Muscatello D, Churches T, Kaldor J, Turner R, Zheng W, Ryan L (2010) Understanding sources of variation in syndromic surveillance for early warning of natural or intentional disease outbreaks. *IIE Transactions*, 42:9, 613-631, DOI: 10.1080/07408170902942667
- Sparks R, Cameron M, Woolford S, Robinson B, Power R (2015). Syndromic Surveillance using Twitter Data. *Emergency Medicine (Los Angel)* vol. 5 issue 3. 9 pages. doi: 10.4172/2165-7548.1000254.
- Yin J, Lampert A, Cameron M, Robinson B and Power R (2012). Using Social Media to Enhance Emergency Situation Awareness. *IEEE Intelligent Systems*, vol. 27, no. 6, pp. 52-59, Nov.-Dec., 2012
- Zucco G, Khanna S, Nguyen A, Boyle J, Hamlet M, Cameron M (2015). Automatic detection of tweets reporting cases of influenza like illnesses in Australia. *Health Information Science and Systems*. 2015; 3(Suppl 1):S4. doi:10.1186/2047-2501-3-S1-S4.